Higher Education (2021) 81:1079–1095
https://doi.org/10.1007/s10734-020-00600-8

The emergence of the higher education research field (1976–2018): preferential attachment, smallworldness and fragmentation in its collaboration networks

Jef Vlegels[1] · Jeroen Huisman[1]

Published online: 13 August 2020

Abstract
The field of higher education research is fast-growing, both in number of publications and in geographical reach. There is however limited evidence on how this growth in publication influences the structure of the underlying co-authorship network. This is important as structural network parameters can change quickly in a fast-growing network, leading to fundamental different network structures, e.g., in terms of hierarchy, fragmentation, and inequality. Ultimately, these network structures can influence the current and future innovation and knowledge production in the field. Empirically, we construct 34 different co-authorship networks of all authors published in 28 higher education journals listed in Web of Science between 1976 and 2018 and perform bibliometric network analyses. We find that the growth of publications and authors in the higher education research field leads to increased clustering among authors, creating a dense core of well-connected author clusters. At the same time, we observe an increasing inequality in the network. The co-authorship network is characterized by high fragmentation and reveals a core-periphery structure. Our analysis shows that co-authorship is a selective process, driven by a Matthew effect based on previous publications. As a result, core authors are unlikely to co-author with newer, less established authors. Moreover, we also detect a growing inequality in the average impact of an article. We conclude the paper by discussing possible explanations and by offering some suggestions for future research.

Keywords Higher education research · Social network analysis · Bibliometrics · Collaboration networks

✉ Jef Vlegels
jef.vlegels@ugent.be

[1] Ghent University, Ghent, Belgium

Introduction

The field of higher education research has gained popularity since the massification of tertiary education spread throughout the world. This massification resulted in more complex and, more recently, globalized higher education systems (Cantwell et al. 2018) and consequently to new challenges for governments, universities, and other stakeholders. At the same time, the relevance of formal and organized learning in sustaining competitiveness among countries grew substantially in the era of globalization and the emergence of knowledge economies. This led to an increasing concern about the quality of higher education and spurred the need for more scientific knowledge on these topics (Santos and Horta 2018; Tight 2004).

The growth of the higher education research field has inspired several researchers to analyze its bibliographic structure (e.g., see: Kosmützky and Krücken 2014; Kuzhabekova et al. 2015; Tight 2014). Bibliographical studies of research fields help to gain insight in important changes in its field dynamics or (relations between) key contributors. This is particularly relevant in research fields that experience relatively big evolutions, like higher education research (Tight 2004, 2018). Initially, these bibliographical analyses in the field of higher education were limited to US-based publications and focused on very restricted subsets of journals or articles (e.g., Budd and Magnuson 2010; Calma and Davies 2014, 2017; Kandlbinder 2012; Milam 1991; Tight 2008). More recently, however, bibliographical analysis included larger sets of journals (e.g., Tight 2004, 2018). Additionally, researchers set up internationally comparative studies across continents, involving not only research in the USA and Europe but also in Asia (Jung and Horta 2013, 2013; Kim et al. 2017; Tight 2014).

These studies have offered valuable insights in identifying the most important actors, e.g., journals, authors, or articles, within a given geographical and time scope. And it is well documented how themes evolve over time (see e.g. Kosmützky and Putty 2016; Daenekindt and Huisman 2020). However, there is still very little knowledge on how macro-level structure of these bibliographical networks in higher education is evolving (Kuzhabekova et al. 2015; Tight 2014). Macro-level structures are captured by general network level properties such as size, cohesion, hierarchy, and fragmentation. Nevertheless, insights from theoretical and empirical developments in network analysis have proven repeatedly that macro-structural properties of a fast-growing network tend to shift dramatically (see, e.g., Barabási et al. 2002; Perc 2010; Wagner and Leydesdorff 2005). More specifically, there is a risk of growing fragmentation, leading to a network of disconnected components. This is particularly relevant for bibliographical studies of research fields, as fragmentation may hamper future collaboration, innovation, and knowledge production (Cowan and Jonard 2003; Fleming et al. 2007a; Fleming et al. 2007b). The fast-growing field of higher education research is thus potentially at risk of losing it innovational potential.

Therefore, in this study, we contribute to the bibliometric knowledge on the higher education research field by, first, analyzing the evolution in the network structure from a very large set of publications in the field of higher education research, unlimited by geographical boundaries and for a considerably long time range (1976–2018). Secondly, we contribute by taking a macro-bibliographical approach, instead of focusing on identifying individual actors. More specifically, we focus on how the structural properties of the network, in terms of the average reachability of an author, the tendency for experienced authors to co-author with each other, and the fragmentation of the network, have evolved over time. Finally, we contribute by discussing how this observed structure relates to potential opportunities for future research in the higher education field.

Collaboration structures and knowledge production

Performing scientific research is often thought to be an individual or at maximum a small team endeavor of people striving for new academic insights or progress. However, in reality, an individual researcher is a relatively small actor in a vast and complex system of people, institutions, and publications, linked by affiliations, citations, and co-authorship (Merton 1973). The individual production of knowledge, and more specifically its documented output, is deeply embedded in and dependent on this complex network of relations. Furthermore, the particular structure of a social network, as well as the position of individuals within this network, is known to have a substantial impact on its outcomes. As such, using theoretical and methodological insights from social network analysis to study scientific research is self-evident. Bibliometrics is one of the most widely used methods to study network aspects of science. It uses quantitative methods to provide insights into the structure and dynamics of a scientific field, analyzing information gathered in connected databases such as citation networks, co-authorship networks, keywords networks, reference networks, or affiliation networks (Pritchard 1969). The growing discipline of bibliometrics has facilitated a better understanding of the scientific landscape in recent years. It helped, e.g., predicting collaboration patterns, identifying important cliques, hubs and authorities of authors and journals, and provided insight in the dynamics of scientific networks over time (e.g., Barabási et al. 2002; Luukkonen et al. 1992; Newman 2001a, 2001b, 2001c; Watts and Strogatz 1998). Furthermore, it also spurred the emergence of studies focusing on mapping the structure and dynamics of a specific thematic scientific field (e.g., Gurzki and Woisetschlaeger (2017) for the luxury research field).

One of the most thriving fields in bibliographical research is focusing on collaboration patterns. Collaboration has become a dominant way of producing knowledge in science (Wuchty et al. 2007). This is not surprising, given the fact that collaborating can lead to higher productivity, as well as improved research impact (Gazni and Didegah 2011; Katz and Martin 1997; Lee and Bozeman 2005; Sooryamoorthy 2009). Co-authorship is a specific form of collaborating, which results in a shared paper or any other form of scientific output as explicit product (Li et al. 2013). By linking authors based on shared publications, a co-authorship network can be constructed. This network of connected and unconnected authors is an important source of information, shared understanding, and knowledge transfer (Li et al. 2013). It contributes to the social capital of a researcher, which allows her or him to share resources with others, and provides opportunities for improving output and impact of current and future work.

It is well established that co-authorship links are not randomly dispersed in a network. Based on the groundbreaking work of Watts and Strogatz (1998) and Barabási and Albert (1999), two general network mechanisms have been proposed that influence tie formation in co-authorship networks: *smallworldness* and *preferential attachment.*

(1) *Smallworldness* refers to a specific relationship between local clustering and the average distance between two actors in a network (Watts and Strogatz 1998). A network is considered a small world when the local clustering, which is the tendency of two of an author's connections to be connected as well, is relatively high, and the average number of steps between actors is small. This smallworldness leads to a network where very well internally connected clusters are formed and at the same time few between-cluster ties exist. Overall cohesion of a small world network is thus rather low, but there are some

very densely connected clusters, and a few between-cluster ties ensure that the average distance is relatively low across the network.

(2) *Preferential attachment* (also known as the Matthew effect) points to a tendency for cumulative advantage (Barabási and Albert 1999). More specifically, more popular nodes in the network will be more attractive for future cooperation, thus leading to a preferential attachment mechanism based on previous popularity of an author. Formally, this translates into two co-occurring mechanisms. First, the observation that networks expand continuously because new authors emerge and, second, the new authors preferably attach to already well-connected authors.

Both network mechanisms have been shown to be important structural determinants of co-authorship networks in different (sub-)disciplines across different social and geographical settings (Barabási et al. 2002; Kronegger et al. 2011, 2012; Moody 2004; Newman 2000, 2001c; Perc 2010). However, they are not universally present in all (sub-)fields of science (Uzzi et al. 2007). Some disciplines are more strongly characterized by single-authorship publications, and the maturity of a field also relates to the level of smallworldness and preferential attachment (Ebadi and Schiffauerova 2015).

Nevertheless, in the context of collaboration networks in science, the consequence of smallworldness and preferential attachment as potential driving determinants for network dynamics cannot be underestimated. In literature on innovation and knowledge production, the question on how the structure of collaboration networks influences the diffusion of knowledge is one of the most fundamental ones (Scherngell 2013). This has led several authors to conclude that the balance between clustering and between-cluster ties is very important for the innovative potential of a collaboration network. This means that, on the one side, there needs to be a sufficient amount of "cliquishness" in the network, and on the other side, sufficient bridging ties are needed between those different cliques (Cowan and Jonard 2003; Fleming et al. 2007a; Fleming et al. 2007b). Dense cliques, or clusters, create trust between its actors; they foster cooperation, knowledge integration, and decrease opportunistic behavior. Arguably, this has positive effect on the creation of innovative knowledge. On the other hand, dense cliques can lead to conformity, risk-avoidance, and less disruptive thinking and contain a lot of redundant information. Those elements contribute to a situation of lock-in, where innovation and knowledge production are potentially hampered (Cowan and Jonard 2003; Crespo et al. 2013; Fleming et al. 2007a; Fleming et al. 2007b; Scherngell 2013; Wagner and Leydesdorff 2005).

Therefore, to maximize the potential of knowledge production, bridging ties between those dense cliques are essential to avoid lock-in. Bridging ties ensure access to new knowledge, the import of non-redundant information, and the dissemination of ideas to other cliques and allows for the recombination of existing ideas and thus improves innovation capabilities (Burt 1992; Cowan and Jonard 2003; Fleming et al. 2007a; Fleming et al. 2007b; Granovetter 1983; Wagner and Leydesdorff 2005). Smallworldness is actually a formalization of this lock-in versus lock-out balance, and hence, it is important to study the existence and evolution of small world properties in co-authorship networks.

Besides the lock-in versus lock-out balance, literature on the importance of collaboration networks for innovation and knowledge production also points to the relevance of studying hierarchy and inequality in those networks. In collaboration networks, a tendency of actors to seek contact with someone who is highly connected can be beneficial, as this increases access to resources. Moreover, working together with previously successful people increases the likelihood to enhance productivity and credibility. This negative assortativity can thus act as a self-organizing system, where newcomers or less experienced actors can benefit from the reputation and resources from the more highly connected ones (Crespo et al. 2013; Newman 2001c; Wagner and Leydesdorff 2005). However, on the other side, this assortativity effect also increases competition for collaboration, especially among newcomers. Furthermore, it potentially increases inequality in the collaboration network and reduces accountability for the experienced nodes (Crespo et al. 2013; Merton and Merton 1968; Newman 2001c; Wagner and Leydesdorff 2005). Hierarchy and inequality, formalized by the definition of preferential attachment, are thus also a very important factor influencing innovation and knowledge creation. Preferential attachment effects are natural to occur and can be beneficial for knowledge creation. However, it should not be at the cost of equity and accountability, as this hampers innovation and knowledge creation.

Given the importance of these two macro-level co-authorship network properties for the current and future innovational potential of higher education research, and the current lack of available statistics for these macro-level network properties, we document in this paper the evolution of smallworldness and preferential attachment in the co-authorship networks in the higher education research field. And, secondly, we discuss how this influences the future of knowledge production in the field.

Data

To gain insight in co-authorship networks, a fairly comprehensive set of journal contributions is used. Peer-reviewed journals that (almost) solely focused on higher education are selected from the Web of Science (WoS). A restriction of WoS is that only papers written in English papers could be included, but given the focus on collaboration, it is acceptable to only consider papers written in the academic lingua franca. WoS is preferred above the Scopus database, for the latter contains many more peripheral (less cited) journals. Within the WoS database, all journals are selected that had "higher education" in its title or—in their mission or objectives—clearly alluded to a prime focus on higher education. This leads to the inclusion of some disciplinary journals that deal, e.g., with teaching and learning in higher education (e.g., *Academy of Management Learning and Education*), but journals that primarily focused on the discipline-related professions are excluded (e.g., *Journal of Social Work Education*). Table 1 lists the 28 journals included in the analysis. This set largely coincides with the list proposed by Tight (2018) and is exactly the same as in Daenekindt and Huisman (2020).

This selection of journals allows us to construct co-authorship networks of all authors published in the abovementioned journals between 1976 and 2018. We start from 1976 as WoS produces reliable information on authors from then on. Indeed, we found only 22 articles with an anonymous author, which we decided to exclude from our analysis. Furthermore, we used a fuzzy-matching procedure using restricted Damerau-Levenshtein distance to detect whether slightly different names actually belonged to the same author. This procedure allowed us to use background information, such as affiliation, country, and previous co-authors, to get a reliable estimation of similarity. Doubtful cases were manually checked. The nodes, x, in the networks are the authors, and a tie, x_{ij}, between two nodes express a co-authorship between two actors. Each of the yearly networks ($n = 34$) is thus undirected and valued, with only whole numbers possible on the ties. Furthermore, we construct the networks as cumulative over time. This means that each new publication year adds articles, and consequently its

Table 1 List of 28 journals included in the analysis

Generic	*Studies in Higher Education*; *The Journal of Higher Education; Review of Higher Education*; *Research in Higher Education*; *Higher Education Research & Development*; *Higher Education*
Topic specific	*Active Learning in Higher Education*; *Assessment & Evaluation in Higher Education*; *Higher Education Policy*; *International Journal of Sustainability in Higher Education*; *Internet and Higher Education*; *Journal of College Student Development*; *Journal of Computing in Higher Education*; *Journal of Diversity in Higher Education*; *Journal of Higher Education Policy and Management*; *Journal of Studies in International Education*; *Teaching in Higher Education*
Discipline specific	*Academy of Management Learning & Education*; *Journal of American College Health; Journal of English for Academic Purposes*; *Journal of Geography in Higher Education*; *Teaching Psychology*; *Teaching Sociology; Physical Review Physics Education Research*; *Journal of Legal Education; Journal of Engineering Education*; *Journal of Hospitality Leisure Sport & Tourism Education*; *Journal of Economic Education*

authors and co-authorship links, to the previous years. The networks can thus be considered to have a "memory," as each new year adds to the previous one. This cumulative construction of the network is a logical consequence of the fact that an existing co-authorship does not disappear when a new publication year is added.

Results

Growth

First of all, we look into the growth of the higher education research field over time. The first plot of Fig. 1 visualizes the number of published articles by year, with a LOESS smoother and 95% confidence interval to reveal the trend. The plot clearly shows how the higher education field started in 1976 as a relatively small field, growing at a relatively modest and constant rate until the first years of this century. From around 2005 on, however, the number of published articles per year seems to increase exponentially. It suggests that the higher education research field really took off in the 2000s, and this exponential trend is still present today.

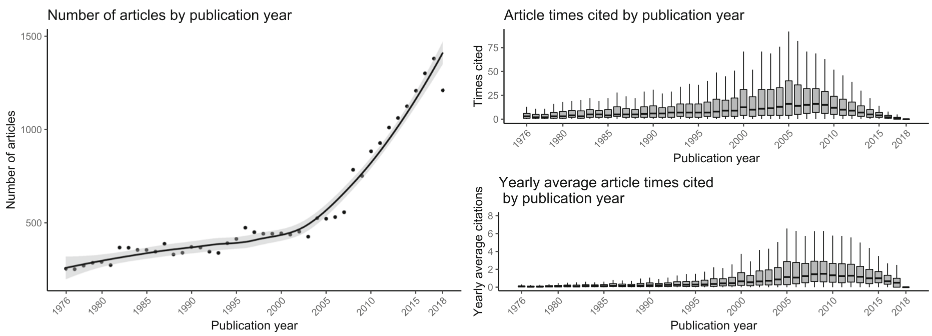


Fig. 1 Article and citation growth

The next two plots in Fig. 1 offer insight in the impact of this growth. The upper plot on the right side of Fig. 1 shows for each publication year a boxplot for the distribution of the number of times an article is cited. The lower plot shows boxplots of the yearly average times an article is cited (by publication year). Immediately, it is obvious that the distribution of citations in both plots is not following the same trend as the growth in number of articles. Both plots show a linear growth, at a moderate rate. Unsurprisingly, article times cited is decreasing for the most recent articles, as these articles are only recently published. However, and perhaps less expected, yearly average article times cited is also decreasing slightly for the most recent years. This means that more recent publications receive on average fewer citations each year. Furthermore, both multiple boxplots also show an increase in variation in citations for more recent years. Particularly in yearly average article citations, there are more outliers with high values. All this leads us to conclude that, although the number of articles in the higher education research field is growing fast the last two decades, the impact, in terms of citations, of an average article decreases in the last decade. Additionally, the right tail of these yearly citation distributions increases in recent years, suggesting a growing inequality in citation impact in recent years compared with the early years.

Figure 2 then shows the growth in number of unique authors by year on the left side plot and the cumulative number of unique authors by year on the right side. Results show that, like the number of publications, the number of unique authors by year grew relatively modest until the start of the twenty-first century. From around 2005 on, the number of unique authors started to grow in an exponential way. This trend is still present in 2018. The cumulative number of unique authors by year shows a more steep but still linear curve until the turn of the century. After which, it also starts to grow in an exponential way. The cumulative number of unique contributors at the end of 2018 has grown to 29,057 authors, an impressive number given the fact that we identified only 246 unique authors in 1976. Furthermore, the graphs in Fig. 2 show that the set of authors contributing to the higher education research field was more stable in the period before the turn of the century than afterward.

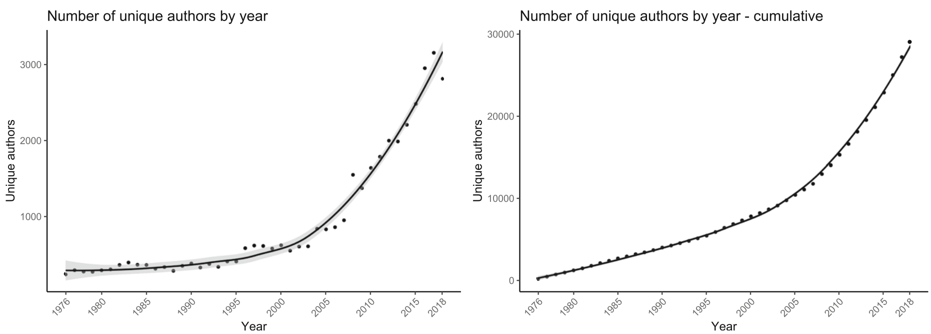


Fig. 2 Author growth patterns

Smallworldness

Next, we focus on the smallworldness properties of the cumulative co-authorship networks over time. As smallworldness points to the relationship between clustering and average path length in a graph, we construct plots on clustering coefficients and average path length by year (Fig. 3).

Formally, the *clustering coefficient*, also called transitivity, is the number of closed triplets (three authors that have co-authored with each other) over the number of total triplets (all combinations of three authors in the network, co-authored or not) in a network (Wasserman and Faust 1994). We also calculate the average of this same coefficient on 500 random networks with the same number of nodes and edges as in our observed networks (Humphries and Gurney 2008; Watts and Strogatz 1998). This allows us to compare the observed clustering coefficients with expected clustering in random networks. In a small world network, we expect observed clustering coefficients to be substantially higher than in random networks.

Indeed, in all years, the observed clustering coefficient is higher than for random networks. However, from 1976 to the turn of the century, the clustering coefficient was decreasing every year. From the beginning of the twenty-first century, this trend changed completely, and clustering started to increase, leading to a big difference in clustering coefficients between the randomly simulated and the observed networks by 2018. As long as the co-authorship network is still relatively small, clustering seems to not really be a concern for the involved authors. However, the exponential growth in the field of higher education research in the last two decades seems to have fostered researchers to create very dense clusters of cooperation. As the co-authorship network is expanding, authors are relying on closed triangles based on trust, re-occurring cooperation, and knowledge integration. However, as discussed above, there is a danger of lock-in if there are not enough ties between those different cliques. Therefore, we also measure average path length. The *average path length* in our cumulative observed networks is calculated as the arithmetic mean of the shortest paths between all pairs of nodes. It can be considered as a measurement of reachability in a network, as it represents the average number of steps needed to reach any other node in the network. In the context of our co-authorship network, it represents how many steps an author would need on average to connect indirectly with another. A lower number points to a better reachability. Again, we compare our observed networks with the mean of the path lengths in 500 random networks. An observed average path length substantially lower than in random networks points to smallworldness, as

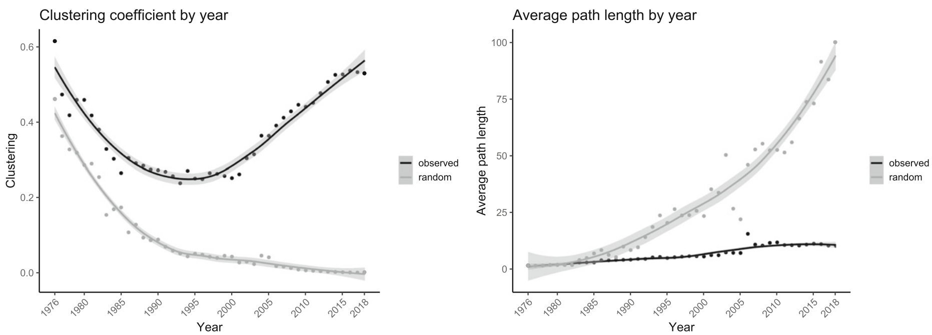


Fig. 3 Smallworldness

this is an indication that there are enough bridging ties between the different cliques in the network to avoid lock-in.

Our results show that until the late 1980s, average path length in our observed networks is very similar to what we expect randomly. However, from the beginning of the 1990s, the observed average path length is deviating more from the randomly simulated networks. While path length starts increasing substantially in the random networks reaching a maximum of 100.16, the increase of average path length in the observed networks is relatively limited, reaching a maximum value of 15.68. This means that, on average, it takes almost 16 steps before any random author connects to another. This number is significantly lower than what one would expect in a random network of similar size (around 100 steps). The results of the clustering and average path length analyses reveal that, although the growing network of co-authorship in the higher education research field shows a lot of cliquishness, the average path length is still relatively low. This is the result of ties that bridge the existing clusters in the network, which has a positive effect on the general reachability. The danger of lock-in, at least from the 1990s on, seems thus relatively limited.

Overall, we can conclude that our observed networks show clear characteristics of smallworldness, certainly from the turn of the century on. There is a tendency for clustering among authors in higher education, while the number of bridging ties is sufficiently high to ensure information flow between those different cliques.

Preferential attachment

After looking into the growth and smallworldness in our observed networks, we focus on preferential attachment. We analyze preferential attachment by studying degree assortativity over time and by analyzing the scale-freeness of the degree distributions (Barabási and Albert 1999). Degree in a co-authorship network equals to the number of co-authored publications of an author, and degree assortativity is the preference of nodes in a network to attach to other nodes that have a similar degree. Preferential attachment, thus, is expected to show dissassortativity, as low-degree (inexperienced) authors tend to co-author with high-degree (experienced) nodes. We calculate the degree assortativity coefficient using Newman's (2002) approach, based on the person correlation coefficient of degree between pairs of linked authors. Positive values indicate degree assortativity, zero points to non-assortativity, and negative values to disassortative degree relations.

The first plot in Fig. 4 shows the evolution of degree assortativity by publication year in our cumulative co-authorship networks. In all years, assortativity is positive, meaning that experienced authors tend to co-author with each other, but significant changes happened between 1976 and 2018. In the first years of our observations, the degree assortativity is very high, reaching a maximum of 0.89. However, notwithstanding a short revival between 1995 and 2005, the coefficient decreased significantly over the years to a minimum of 0.29 in 2017. This means that in general, the co-authorship network shows a tendency for degree similarity between two co-authors. Even though this tendency has decreased over the years, the overall coefficient is still positive. Consequently, it is unlikely for a (not yet) successful author to co-author an article with an experienced author and vice versa. Establishing co-authorship in the higher education field, according to the degree assortativity coefficient, is thus fairly selective.

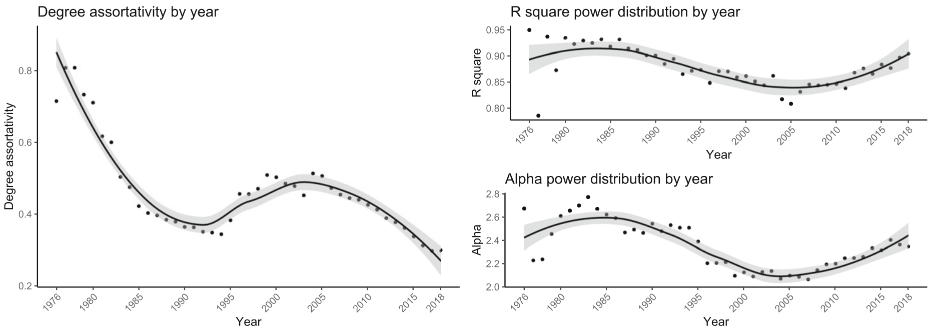


Fig. 4 Preferential attachment

Alternatively, preferential attachment can be detected by focusing on the degree distribution of a network. If newly added nodes to a network preferentially attach to nodes with higher degrees, the experienced become more experienced. This means that the degree distribution will show a very long right tail: most authors will have relatively few co-authored papers, while a few hubs of authors will show very high degrees. The presence of these hubs that are several orders of magnitudes larger in degree than most nodes is a characteristic of so-called power law networks. These power law networks are called scale-free, because power laws have the same functional form at all scales (Albert and Barabási 2002; Barabási and Albert 1999). We can test whether the degree follows a power law distribution by plotting the proportion of each degree value on a log-log scale. Next, we try to fit a power law distribution on this scatterplot. R-square can be used as a measurement of fit of the theoretical power law distribution on the observations (second plot in Fig. 4). And alpha, the exponent of the power law, is an indicator for the slope of the line, which is expected to be at least higher than one in a scale-free network (third plot in Fig. 4).

The R-square and alpha of the fitter power distributions by year follow a similar pattern. In the first 10 years, when the networks are still relatively small, the results show a lot of variance. However, from the 1990s on, there is a decreasing trend in both R-square and alpha of the power distributions. This changes shortly after the turn of century. The fit and the slope starts to increase from around 2005, and this trend is still present in 2018. Generally speaking, we can conclude that the differences over the years are relatively small. In all years, a power distribution seems to be a relatively good fit on the log-log degree distributions with R-square values higher than 0.8 and alpha never lower than 2. This confirms the scale-freeness of the co-authorship networks in the higher education field, leading to very skewed degree distributions. Hence, experienced authors tend to co-author with each other, and it is very difficult for inexperienced authors to set up a co-authorship with more experienced authors.

Fragmentation

Finally, we include an analysis on fragmentation over time in the co-authorship networks. The combination of our results on smallworldness and preferential attachment so far shows a high tendency for clustering, a relatively low path length, and a very pronounced inequality in degree. This situation can, but not necessarily, lead to a fragmented network with a pronounced core-periphery structure. A relatively dense core of well-connected clusters can possibly account for most of the connections between the authors and explain the results on smallworldness indicators. On the other side, a relatively large periphery of authors might struggle to connect to any of those core authors, explaining our preferential attachment results.

We look into the general fragmentation of the networks over time using two metrics: the percentage of the largest component and modularity (Fig. 5). The largest component in a network is the largest subgraph of this network in which any two actors are connected by a path (Wasserman and Faust 1994). It is thus a fully connected subgraph, without isolates. The first plot in Fig. 5 shows that, although the proportion of the largest component has grown substantially from 1976 to 2018, it is still relatively low with a maximum value of 0.33 in 2018. The largest connected component includes about a third of all authors in the network in 2018, which indicates that most actors are not connected to the most important component of the co-authorship networks.

Secondly, modularity is a measurement of general fragmentation in the network. It is the proportion of within-cluster ties divided by the number of ties between clusters. High modularity thus points to a very fragmented network. We use the Louvain clustering algorithm to calculate modularity in the different years (Blondel et al. 2008). The second plot in Fig. 5 shows that modularity is decreasing from the 1990s on. However, in general, modularity is very high across all years and never lower than 0.98. Again, this points to a very fragmented network. The dense core might lead to low average path lengths in general, but these very high modularity values show that this is probably only true for authors in the core, and not for the periphery. This leads us to conclude that the positive effects of bridging ties, e.g., better dissemination and access to new non-redundant information, is only accessible for a relatively limited group of core authors.

Finally, a visual representation in five time slices of the networks (see Fig. 6) confirms these findings. Over the years, the smallworldness and preferential attachment mechanisms in the co-authorship network have established a strong core-periphery structure. In general, core-periphery networks are considered to hamper

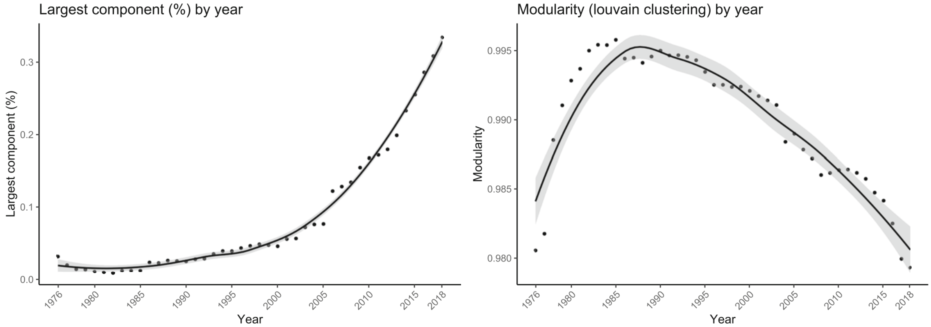
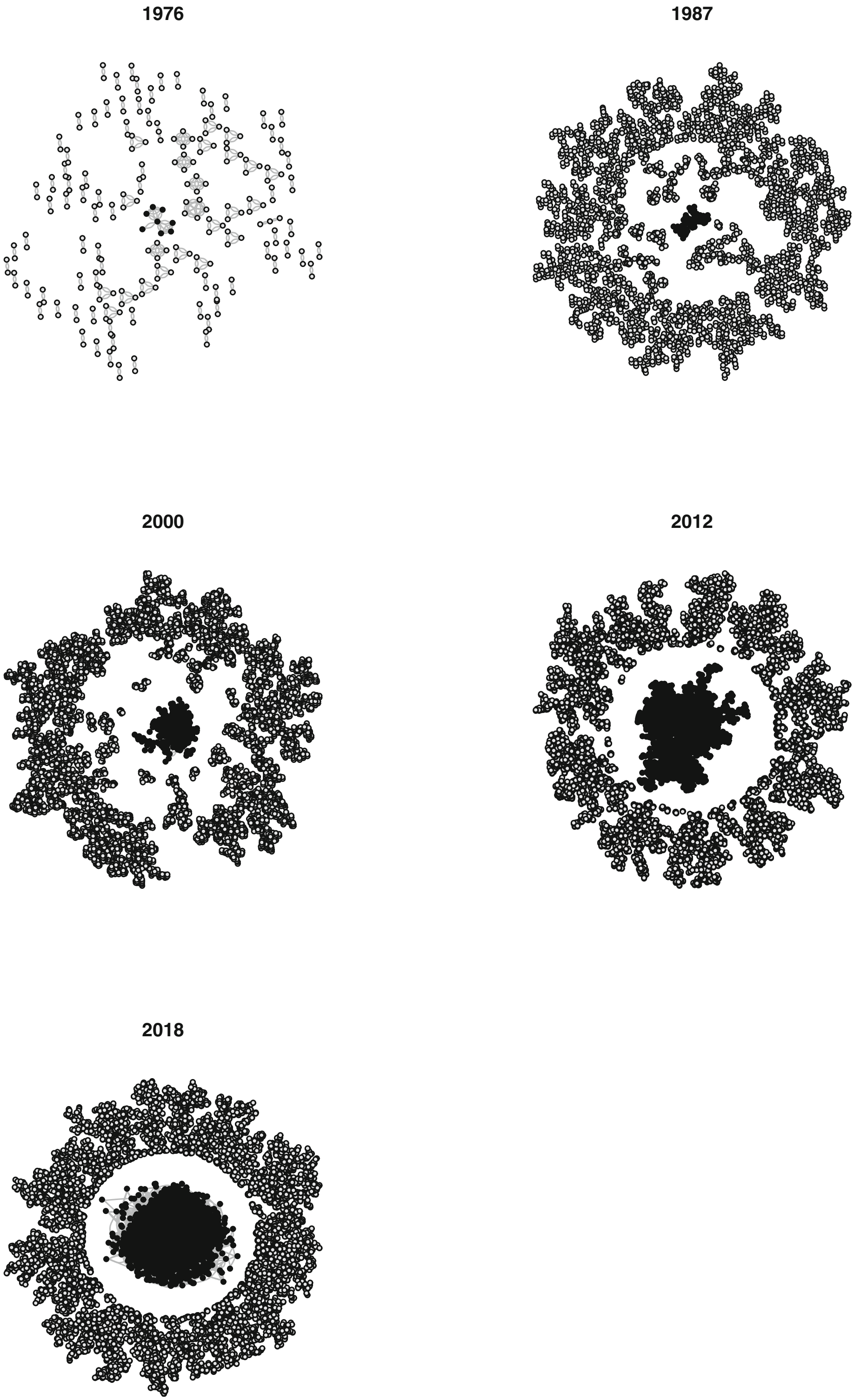


Fig. 5 Fragmentation


1976
1987
2000
2012
2018

◀ **Fig. 6** Network plots of 5 time slices (1976, 1987, 2000, 2012, 2018), members of the largest component are colored black, others gray. Unconnected authors are excluded from these plots

information flow as they fail to bridge between the core and the periphery actors. This also points to strong inequalities.

Conclusion and discussion

Co-authorship networks constitute the foundation of a scientific community. It has been demonstrated in many contexts that collaboration is the prevailing way of performing research in contemporary academics (Wuchty et al. 2007). It is also well understood that the particular structure of co-authorship networks is of central importance for the performance of its individual actors but also for the general outcomes of the research community (Barabási et al. 2002; Watts and Strogatz 1998). There is however limited evidence of the structure of co-authorship networks in the higher education research field. This is particularly pertinent given the fact that several authors argue that the field of higher education research is fast-growing, both in number of publications and in geographical reach (Santos and Horta 2018; Tight 2004). Structural network parameters can change quickly in a fast-growing network, leading to fundamental different network structures, e.g., in terms of hierarchy, fragmentation, and specialization. This paper therefore looked into the evolution of the general structure of co-authorship networks in the field of higher education research. More specific, we focused in our analysis on growth, smallworldness, preferential attachment, and fragmentation of the co-authorship networks over time. As shown in the theoretical overview of this paper, pertinent literature expects that these parameters potentially influence current and future innovation and opportunities in higher education research. Empirically, we started from 34 co-authorship networks of all authors published in a higher education journal listed in Web of Science (WoS) between 1976 and 2018.

The results show, first of all, that both the number of authors and the number of articles in the higher education research field grew substantially since 1976. Especially in the last two decades, when this growth turned from linear to exponential. Second, our analyses of smallworldness show that the growing co-authorship network leads to increased clustering among authors in the field. At the same time, the average path length between authors remains remarkably low given the very fast growth in size. In terms of innovational potential of the co-authorship network in the higher education field, this is good news. First of all, there is a sufficient level of cliquishness in the network. This is important, as these dense clusters are known to foster trust, cooperation, and knowledge integration. Second, we detect a large number of bridging ties between clusters which avoids lock-in and enables the import of non-redundant information, the dissemination of ideas, and the recombination of existing knowledge.

The results are less positive, however, when we look into indicators of inequality. We observe that more recent publications receive on average fewer citations every year and that the variation in the number of citations for each article is increasing in recent years. So, although the field is growing fast in size, its average impact—in terms of citations per article—is not. Moreover, the increasing variation in citations per article is an indication that the growth is leading to more inequality. We observe a field in which some authors benefit from the larger community by receiving more citations, while other authors remain largely unnoticed. This finding is confirmed by the results of the analyses on preferential attachment. Our analysis shows that co-authorship in the field of higher education research is very selective. There is a clearly a Matthew effect at play, based on a preferential attachment by degree. It is very unlikely for a new or inexperienced author to co-author with an established author. Finally, an additional analysis on the fragmentation in the field points to very high levels of fragmentation and reveals the emergence of a strong core-periphery structure in recent years.

Possible explanations for these outcomes can be found in literature on predicting scientific collaboration. A first explanation is the fact that search frictions and communication costs limit and shape collaborations between scientists. Indeed, scientists looking for a co-author tend to focus on minimizing search and communication efforts (Boudreau et al. 2017; Kraut et al. 2014). A successful strategy in this sense is a focus on similar partners. Communicating with like authors is more efficient for there is a lower risk of misunderstandings. This is in line with our findings on preferential attachment and positive degree assortativity. Authors with a similar degree can be expected to share comparable levels of experience, status, and popularity. However, it is worth nothing that we intended to reveal generic patterns. That is, within subfields of higher education or other specific contexts (countries, national, or institutional cultures), patterns of collaboration may deviate from this macro-level pattern. Secondly, degree assortativity effects can also be explained as a strategy to lower risk in productivity, visibility, and recognition (Katz and Martin 1997). As scientists strive for recognition of their published work, it is considered a risk-reducing strategy to co-author with someone who has previously proven his competences. This explains the observed Matthew effect, the formation of the core-periphery structure, and, in a second step, the strong clustering effects. Finally, the importance of spatial proximity offers another possible explanation. Research on determinants of scientific collaboration continues to stress the importance of spatial proximity. Despite some scholars predicting the “death of distance,” information-rich face-to-face interactions are still of fundamental importance for setting op academic collaborations (Boudreau et al. 2017, p. 575; Spithoven et al. 2019). The growing number of authors in the field of higher education also increased the internationalization of the authors involved. Many of the new authors originate from regions and countries that were previously not represented in the field (Jung and Horta 2013; Kosmützky and Putty 2016; Kuzhabekova et al. 2015; Tight 2012, 2014). As a result of this spatial distance, they might struggle to connect with the more established authors. This leads to more inequality in the field, i.e., core-periphery structures, general fragmentation, and preferential attachment mechanisms.

The findings from this paper can be an inspiration for authors and other actors, e.g., publishers, funding institutions, or conference organizers, in their future actions and plans. We believe, for example, that the higher education research field would benefit from more informal opportunities for researchers to connect with each other. Network events at conferences, online fora, or even an introduction from an editor might offer the opportunity for less established scholars to connect with more central authors. These direct contacts might lower future communication costs and thus form the basis for future collaboration. Approachability is the key here, as this is a necessary condition for a periphery author to connect with a core author. In the end, this could lead to a less elitist core which is beneficial for the innovative potential of the field. Funding agencies can also play an important role. Dedicated programs could encourage collaborations between established scholars and newcomers. This external incentive might help reducing the risk of setting up a collaboration with a relatively new author. Again, we expect this to have a positive influence on equality and ultimately the innovational potential of the higher education field.

Inevitably, this research has some limitations which might be addressed in future research. First, we limit our co-author networks to publications in a selection of higher education journals listed in the Web of Science. While we thoughtfully constructed this selection of journals, scholars might be interested in adding other journals or deleting some. Second, particularly the fact that we focused on publications in English can be a source of concern. Many scholars use—for good reasons—their own (non-English) languages to report on their research findings. This seems to be particularly the case for practitioners from Asia (Jung et al. 2018). Third, while we believe our journal selection reflects the most important internationally recognized outlets, future researchers might be interested in taking a different perspective. They can, for example, come up with another research design to construct a specific network of co-authors involved in more localized higher education research (e.g., using snowball sampling).

It is stressed again that the network analysis in this paper focuses on global descriptives of the networks over time. This is a commonly used and published method, and it produces very interesting results for a first paper. However, in future, researchers can consider using inductive network analyses techniques to predict the tie between two authors, using a series of network and author characteristics simultaneously, i.e., by using exponential random graph models, possibly even with a longitudinal research design. This might offer some additional insights in the relative importance of author characteristics (e.g., seniority, gender, etc.) and network effects (as, e.g., preferential attachment, triadic closure, etc.) for predicting collaboration in the higher education research field. Moreover, this would allow the introduction of more contextual variables. Bearing in mind that collaboration patterns may differ by subfield, by country, or even by institution (e.g., embeddedness of higher education researchers in disciplinary units), it is worthwhile to zoom in on more specific national or geographical patterns.

Acknowledgments The authors like to thank Dr. Stijn Daenekindt for his help in collecting the data for this paper.

Funding information This study received financial support of the Research Foundation – Flanders (project G.OC42.13 N).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47.

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

Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications, 311*(3), 590–614. https://doi.org/10.1016/S0378-4371(02)00736-7.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Boudreau, K. J., Brady, T., Ganguli, I., Gaule, P., Guinan, E., Hollenberg, A., & Lakhani, K. R. (2017). A field experiment on search costs and the formation of scientific collaborations. *Review of Economics and Statistics, 99*(4), 565–576.

Budd, J. M., & Magnuson, L. (2010). Higher education literature revisited: Citation patterns examined. *Research in Higher Education, 51*(3), 294–304. https://doi.org/10.1007/s11162-009-9155-6.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge: Harvard University Press.

Calma, A., & Davies, M. (2014). Studies in higher education 1976–2013: A retrospective using citation network analysis. *Studies in Higher Education, 40*(1), 4–21. https://doi.org/10.1080/03075079.2014.977858.

Calma, A., & Davies, M. (2017). Geographies of influence: A citation network analysis of *Higher Education* 1972–2014. *Scientometrics, 110*, 1579–1599. https://doi.org/10.1007/s11192-016-2228-3.

Cantwell, B., Marginson, S., & Smolentseva, A. (2018). *High participation systems of higher education*. Oxford University Press.

Cowan, R., & Jonard, N. (2003). The dynamics of collective invention. *Journal of Economic Behavior & Organization, 52*(4), 513–532. https://doi.org/10.1016/S0167-2681(03)00091-X.

Crespo, J., Suire, R., & Vicente, J. (2013). Assortativity and hierarchy in localized R&D collaboration networks. In T. Scherngell (Ed.), *The geography of networks and R&D collaborations* (pp. 115–128). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-02699-2_7.

Daenekindt, S., & Huisman, J. (2020) Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991-2018. *High Education*. https://doi.org/10.1007/s10734-020-00500-x.

Ebadi, A., & Schiffauerova, A. (2015). On the relation between the small world structure and scientific activities. *PLoS One, 10*(3), e0121129–e0121129. https://doi.org/10.1371/journal.pone.0121129.

Fleming, L., King, C., & Juda, A. I. (2007a). Small worlds and regional innovation. *Organization Science, 18*(6), 938–954. https://doi.org/10.1287/orsc.1070.0289.

Fleming, L., Mingo, S., & Chen, D. (2007b). Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly, 52*(3), 443–475. https://doi.org/10.2189/asqu.52.3.443.

Gazni, A., & Didegah, F. (2011). Investigating different types of research collaboration and citation impact: A case study of Harvard University's publications. *Scientometrics, 87*(2), 251–265.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory, 1*, 201–233.

Gurzki, H., & Woisetschlaeger, D. M. (2017). Mapping the luxury research landscape: A bibliometric citation analysis. *Journal of Business Research, 77*, 147–166.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS One, 3*(4), e0002051. https://doi.org/10.1371/journal.pone.0002051.

Jung, J., & Horta, H. (2013). Higher education research in Asia: A publication and co-publication analysis. *Higher Education Quarterly, 67*(4), 398–419.

Jung, J., Horta, H., & Yonezawa, A. (2018). *Introduction: Higher education research as a field of study in Asia*, *In Researching Higher Education in Asia (pp. 1–11)*. Singapore: Springer.

Kandlbinder, P. (2012). Recognition and influence: The evolution of higher education research and development. *Higher Education Research and Development, 31*(1), 5–13. https://doi.org/10.1080/07294360.2012.642836.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*(1), 1–18.

Kim, Y., Horta, H., & Jung, J. (2017). Higher education research in Hong Kong, Japan, China, and Malaysia: Exploring research community cohesion and the integration of thematic approaches. *Studies in Higher Education, 42*(1), 149–168.

Kosmützky, A., & Krücken, G. (2014). Growth or steady state? A bibliometric focus on international comparative higher education research. *Higher Education, 67*(4), 457–472.

Kosmützky, A., & Putty, R. (2016). Transcending borders and traversing boundaries: A systematic review of the literature on transnational, offshore, cross-border, and borderless higher education. *Journal of Studies in International Education, 20*(1), 8–33.

Kraut, R. E., Egido, C., & Galegher, J. (2014). Patterns of contact and communication in scientific research collaborations. In *Intellectual teamwork* (pp. 163–186). Psychology Press.

Kronegger, L., Ferligoj, A., & Doreian, P. (2011). On the dynamics of national scientific systems. *Quality & Quantity, 45*(5), 989–1015.

Kronegger, L., Mali, F., Ferligoj, A., & Doreian, P. (2012). Collaboration structures in Slovenian scientific communities. *Scientometrics, 90*(2), 631–647.

Kuzhabekova, A., Hendel, D. D., & Chapman, D. W. (2015). Mapping global research on international higher education. *Research in Higher Education, 56*(8), 861–882.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673–702.

Li, E. Y., Liao, C. H., & Yen, H. R. (2013). Co-authorship networks and research impact: A social capital perspective. *Research Policy, 42*(9), 1515–1530.

Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology & Human Values, 17*(1), 101–126.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Merton, R. K., & Merton, R. C. (1968). *Social theory and social structure*. Simon and Schuster.

Milam, J. H. (1991). The presence of paradigms in the core higher education journal literature. *Research in Higher Education, 32*(6), 651–668.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review, 69*(2), 213–238.

Newman, M. E. J. (2000). Models of the small world. *Journal of Statistical Physics, 101*(3–4), 819–841.

Newman, M. E. J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E, 64*(1), 016131. https://doi.org/10.1103/PhysRevE.64.016131.

Newman, M. E. J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E, 64*(1), 016132. https://doi.org/10.1103/PhysRevE.64.016132.

Newman, M. E. J. (2001c). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences, 98*(2), 404–409. https://doi.org/10.1073/pnas.98.2.404.

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters, 89*(20), 208701.

Perc, M. (2010). Growth and structure of Slovenia's scientific collaboration network. *Journal of Informetrics, 4*(4), 475–482.

Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation, 25*(4), 348–349.

Santos, J. M., & Horta, H. (2018). The research agenda setting of higher education researchers. *Higher Education, 76*(4), 649–668.

Scherngell, T. (2013). The networked nature of R&D in a spatial context. In *Advances in Spatial Science* (pp. 3–11). https://doi.org/10.1007/978-3-319-02699-2_1.

Sooryamoorthy, R. (2009). Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. *Scientometrics, 81*(1), 177.

Spithoven, A., Vlegels, J., & Ysebaert, W. (2019). Commercializing academic research: A social network approach exploring the role of regions and distance. *The Journal of Technology Transfer*, 1–36.

Tight, M. (2004). Research into higher education: An a-theoretical community of practice? *Higher Education Research and Development, 23*(4), 395–411.

Tight, M. (2008). Higher education research as tribe, territory and/or community: A co-citation analysis. *Higher Education, 55*(5), 593–605. https://doi.org/10.1007/s10734-007-9077-1.

Tight, M. (2012). Higher education research 2000–2010: Changing journal publication patterns. *Higher Education Research and Development, 31*(5), 723–740. https://doi.org/10.1080/07294360.2012.692361.

Tight, M. (2014). Working in separate silos? What citation patterns reveal about higher education research internationally. *Higher Education, 68*(3), 379–395. https://doi.org/10.1007/s10734-014-9718-0.

Tight, M. (2018). Higher education journals: Their characteristics and contribution. *Higher Education Research and Development, 37*(3), 607–619. https://doi.org/10.1080/07294360.2017.1389858.

Uzzi, B., Amaral, L. A., & Reed-Tsochas, F. (2007). Small-world networks and management science research: A review. *European Management Review, 4*(2), 77–91.

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy, 34*(10), 1608–1618. https://doi.org/10.1016/j.respol.2005.08.002.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge university press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science, 316*(5827), 1036–1039.