

# An internationalised Europe and regionally focused Americas: A network analysis of higher education studies

Aliakbar Akbaritabar<sup>1</sup>  | Giovanni Barbato<sup>2</sup> 

<sup>1</sup>Laboratory of Digital and Computational Demography, Max Planck Institute for Demographic Research (MPIDR), Rostock, Germany

<sup>2</sup>Department of Economics, Management and Quantitative Methods (DEMM), University of Milan, Milan, Italy

## Correspondence

Aliakbar Akbaritabar, Laboratory of Digital and Computational Demography, Max Planck Institute for Demographic Research (MPIDR), Konrad-Zuse-Str. 1, Rostock 18057, Germany.  
Email: akbaritabar@demogr.mpg.de

## Funding information

Data was obtained from Kompetenzzentrum Bibliometrie (Competence Centre for Bibliometrics), Germany, which is funded by the Federal Ministry for Education and Research (BMBF), Germany, with grant number 01PQ17001

## Abstract

The study on which this article reports investigated the internationalisation of higher education studies by examining collaborations in the form of international co-authorships. We analyse how network-based mechanisms, related to structural relationship between authors (preferential attachment, i.e., higher tendency to collaborate among the most productive ones) and node level features (homophily, i.e., tendency to collaborate with similar others), affect higher education co-authorship networks. We build a bipartite co-authorship network based on 17,262 publications from 33 specialised higher education journals indexed in Scopus from 1996–2018. Scientific collaboration in higher education mainly occurs within national borders. We found that higher education is not an internationally oriented field of research, with around 90% single-country publications. A geographical divide was observed between the two largest communities (Europe, Asia and Oceania vs. the Americas) which was also reflected in the research themes addressed by these communities, structured around the known divide between (1) learning and teaching, and (2) policy-based studies. Preferential attachment was observed to be a network-based mechanism that contributes to drive the formation of new co-authorships. Similarly, homophily based on academic seniority and research productivity emerged

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *European Journal of Education* published by John Wiley & Sons Ltd.

as a significant explanatory mechanism. We present a methodology for disambiguating the names of higher education institutions. We demonstrate an associated effect on co-authorship networks. Our research and analysis support the effectiveness of the proposed method.

## 1 | INTRODUCTION

This paper investigates the internationalisation of higher education studies by focussing on international research collaborations. *International research collaborations* are structured or occasional scientific activities carried out by scholars and institutions from different countries to achieve a common research goal (Katz & Martin, 1997). One strategy to empirically investigate international research collaborations is through international co-authorships, namely, when two or more scholars from different countries have written one or more papers together (Acedo et al, 2006). Although international research collaborations do not necessarily result in co-authoring a paper, this approach is claimed to be one of the most reliable means of measuring them (Katz & Martin, 1997).

The literature on aspects of internationalisation of higher education such as student and staff mobility, international partnerships and association between universities, has often intrinsically assumed and claimed that such processes would be permanently expanding involving all types of actors and dimensions of higher education (Fumasoli, 2019).

Concerning international research collaborations, the literature has demonstrated a constant growth of international co-authorship along with a rising number of authors and countries involved, reflecting a trend towards *team science* or more highly collaborative work (Kwiek, 2021; Wagner et al, 2015). However, empirical research has mainly been focused on natural sciences, while there are fewer studies on social sciences (Acedo et al, 2006; Akbaritabar et al., 2018, 2020; Moody, 2004) and even fewer in interdisciplinary fields such as higher education studies. This article contributes to the literature by analysing the internationalisation of higher education studies using data on international co-authorships retrieved from 33 higher education journals.

Previous works depict higher education studies as a highly fragmented field characterised by a marked division between *policy-based* and *learning and teaching* approaches (Horta & Jung, 2014; Tight, 2004), leading to Macfarlane's (2012) idea of higher education studies as an *archipelago*. This idea emphasises the low level of collaboration between these two research communities. Daenekindt and Huisman (2020) show how these *islands* of research will drift further apart due to a trend of increasing specialisation. Horta and Jung (2014) and Kim et al (2017) illustrate that almost all collaboration patterns occur *within* the two communities, while cross-fertilisation is low. A large part of higher education studies is carried out by *part-time researchers* (Horta & Jung, 2014; Kim et al, 2017), namely, scholars or practitioners who occasionally used higher education to investigate phenomena anchored in other disciplines and who thus present infrequent publications on higher education, contributing to an increase in the fragmentation of research on higher education. This phenomenon is also a consequence of the weak institutional basis of higher education studies (Teichler, 2000). The number of research institutes, journals and study programmes specifically on higher education is indeed limited, especially outside the United States and the United Kingdom, causing the scholarship on higher education to be spread out across other disciplines (e.g., management and sociology) and thus highly disconnected (Kehm, 2015).

Against this backdrop, few studies have examined the internationalisation of higher education studies by looking at international co-authorships (Avdeev, 2019; Jung & Horta, 2013, 2015; Kuzhabekova et al, 2015; Kwiek, 2020; Lovakov & Yudkevich, 2021; Tight, 2007, 2014). These studies explain internationalisation processes by focussing on elements such as the linguistic and geographical proximity of authors or the maturity of national higher education communities of research (Avdeev, 2019; Kwiek, 2020; Tight, 2014). However, these factors

are exogenous to the co-authorship network, i.e., they do not consider the role of the position or relationships among actors (scholars/universities), which have thus far been mostly overlooked in the higher education literature. Therefore, this article contributes to filling this gap by adopting a social network theory perspective to study the internationalisation of higher education studies.

Social network theory (SNT) states that the set of direct and indirect relationships among actors provide opportunities and constraints that affect their behaviours (Granovetter, 1973). Following SNT, the characteristics and internal dynamics of the co-authorship network—namely, the connections among actors (ties) and the position of actors within the network—should be considered explanatory factors in the internationalisation of science (Acedo et al, 2006; Wagner & Leydesdorff, 2005; Wagner et al, 2015). Based on this theoretical lens, we articulated three research goals, one foundational for our study and two about the internationalisation of higher education studies:

1. Disambiguation of scientific entities' names (e.g., names of authors and institutions), have often been taken for granted, while the association of a scholar with a specific institution has paramount effect on the dis-connectivity of co-authorship networks (D'Angelo & van Eck, 2020; De Stefano et al, 2013). We investigated the level of improvement in the accuracy of descriptions of network constructions obtained by the correct identification of institutions and author names, i.e., disambiguation. We wish by this to contribute to the higher education literature in methodological terms.
2. We have sought to identify the main features of higher education co-authorship networks, specifically, to describe international co-authorship networks.
3. We have investigated how network-based mechanisms explain the dynamics and growth of co-authorship networks for international higher education studies through a bipartite exponential random graph model (ERGM) and bipartite community detection. We have carried out a statistical analysis of hypotheses on specific network mechanisms such as preferential attachment, homophily and the significance of network sub-communities. These techniques have rarely been applied to higher education co-authorship (Biancani & McFarland, 2013; Kezar, 2014).

We first review literature regarding the internationalisation of higher education studies. We then present the main network-based mechanisms to be tested using our statistical model. This model is illustrated in the fourth section along with the data. Findings will then be presented and discussed in the last two sections.

## 2 | LITERATURE ON THE INTERNATIONALISATION OF HIGHER EDUCATION STUDIES

There is a dearth of papers that have empirically examined the internationalisation of higher education studies (Avdeev, 2019; Jung & Horta, 2013, 2015; Kuzhabekova et al, 2015; Kwiek, 2020; Tight, 2007, 2014). While most of them focus on either an entire continent or regional area, few present a global perspective.

Tight's analyses (2007, 2014) focus on a sample of North American and non-North American higher education journals by revealing a remarkable division between the North American community and the rest of the world. North American journals are dominated by authors based in the Americas, both in terms of authors' country of origin (Tight, 2007) and citation patterns (Calma & Davies, 2017; Tight, 2014). By contrast, European or Australian journals such as *Higher Education*, *Studies in HE*, and *Higher Education Research & Development* have been more open towards North American and Asian authors over time (Kwiek, 2020; Tight, 2014). This low level of collaboration is claimed to be the result of several factors, including North American *methodological nationalism* and a longer history of higher education studies in that region. Furthermore, this division is also a consequence of the nature of higher education studies, which is largely driven by emerging specific national and local policy problems (Kehm, 2015; Tight, 2007, 2014).

Liu et al (2019) examined how the research themes of higher education studies vary among countries. The authors found that Australia and the United Kingdom play a central role in the international community because several countries share a significant research theme similarity with them. Similarity is heavily influenced by the policy problems that countries are facing. The authors argue that this high similarity in research themes may be considered a key factor for explaining the probability for collaborations among scholars. The European case is, in this sense, explicative, due to an increasingly shared policy framework and the growth of EU-level funding for collaborative research, as pointed out by Kwiek (2021).

In terms of the degree of international collaboration among higher education scholars, while Kosmützky and Krücken's (2014) analysis reveals a steady trend over time, Kuzhabekova et al (2015) present a slowly increasing tendency. However, this share is still only 11.3% of the publications analysed whereas the majority of articles (80.6%) were produced by authors from a single country. Avdeev (2019) confirms that international higher education collaboration is progressing, even if by a slow increase, since 2000. Similar modest levels of international collaboration were noted in Jung and Horta's (2013, 2015) analysis of Asian higher education research networks, which highlights a substantial isolation of scholars (43% single-author publications). Domestic co-authorship patterns are definitely more frequent (see similar results in Lovakov and Yudkevich (2021) for post-Soviet countries). Jung and Horta (2013) also find that productivity and co-authorship patterns rely on a handful of scholars within a few universities, whereas the remaining publications are authored by *part-time researchers* (Kwiek, 2020). Moreover, by examining a sample of prestigious higher education journals, Kwiek (2020) shows that the distribution of country affiliations is changing over time, with the diminishing influence of American scholars compensated by the growth of Continental Europe and East Asia. Finally, Avdeev (2019) uses spatial scientometric techniques to demonstrate that the intensity of international co-authorship in higher education is still dependent on geographical and linguistic proximity between authors.

These papers explain the internationalisation of higher education studies by a variety of elements that include historical circumstances, geographical proximity and linguistic commonality (Avdeev, 2019). Other explanations address policy-related dynamics like funding mechanisms, academic mobility, maturity of higher education systems and similarities in research themes (Jung & Horta, 2015; Kim et al, 2017; Liu et al, 2019; Tight, 2007, 2014). Among these factors, the nature of the relationships between authors and universities embedded in the co-authorship network—which is discussed in the next section—has been overlooked (Biancani & McFarland, 2013; Kezar, 2014).

### 3 | A SOCIAL NETWORK THEORY PERSPECTIVE ON FACTORS AFFECTING INTERNATIONAL RESEARCH COLLABORATIONS

Studies on international research collaborations have considered a diverse set of explanatory factors (Kwiek, 2021; Wagner & Leydesdorff, 2005). Luukkonen et al (1992) underline the role of geographical proximity and linguistic similarities among countries. The role of supranational policies (e.g., EU funding initiatives on research collaboration) are argued to also support international research collaborations (Kwiek, 2021; Wagner et al, 2015). Countries with less developed scientific capacity lean towards *advanced countries* in terms of collaboration preferences (Wagner & Leydesdorff, 2005). Although these elements contribute to explain the growth of international research collaborations, empirical studies have produced controversial evidence (Wagner & Leydesdorff, 2005; Wagner et al, 2015) suggesting that other explanatory factors can shed light on the dynamics of international research collaborations. Social network theory (SNT) provides, in this sense, a complementary perspective. Dynamics at the level of co-authorship networks are considered explanatory factors. These consist of the connections among actors (ties) and the position of actors (nodes) within the network (Acedo et al, 2006; Biancani & McFarland, 2013; Wagner & Leydesdorff, 2005; Wagner et al, 2015). SNT identifies (1) endogenous and (2) exogenous explanatory factors. Based on these factors we have developed and tested hypotheses using a statistical model.

### 3.1 | Endogenous factors

Endogenous factors are structural recurrent processes based on the presence of ties, which explain the formation of other ties. One main endogenous mechanism is preferential attachment (Biancani & McFarland, 2013). Preferential attachment is the propensity to collaborate with scholars who already present many previous collaborations (co-authorships). (Barabási et al, 2002). This mechanism emphasises the power of popularity in research networks, as also articulated in the theory of cumulative advantage in science, i.e., the *Matthew effect* (Merton, 1968). Several empirical works demonstrated the role of preferential attachment in driving co-authorship networks (Akbaritabar et al, 2020; Dahlander & McFarland, 2013; Moody, 2004; Newman, 2001a; Sciabolazza et al, 2017; Wagner & Leydesdorff, 2005b; Zhang et al, 2018). The first hypothesis (H1) that we tested was the following: *scholars with several previous collaborations present a higher probability of establishing new co-authorships.*

### 3.2 | Exogenous factors

Exogenous factors are the attributes of nodes that influence the formation of new ties. Homophily is a major explanatory mechanism in networks (Akbaritabar et al, 2020; Sciabolazza et al, 2017; Zhang et al, 2018). Homophily is the tendency of nodes to start new ties with whom they present similarities on one or more attributes (Biancani & McFarland, 2013). Empirical studies have demonstrated the effects of different homophily effects in co-authorship networks, including same gender, language, ethnicity (Dahlander & McFarland, 2013) and similarity of research themes (Akbaritabar et al, 2020; Zhang et al, 2018). However, differently from preferential attachment, homophily can be tested in empirical analyses only if a node's attributes can be retrieved from publication data. Consequently, we will concentrate on three homophily effects that can be tested based on our data, and are considered significant exogenous factors in literature.

Firstly, research productivity has been largely associated with scientific collaboration. Studies have proved that collaboration affects productivity positively (Biancani & McFarland, 2013). However, how research productivity influences the formation of new collaborations has been less studied. We include it as a homophily effect, testing whether those with a similar level of research productivity tend to collaborate more with each other (Zhang et al, 2018).

Secondly, we consider academic seniority, intended as the length of time that a scholar has been publishing. This effect tests both whether scholars tend to collaborate with those who present a similar seniority and whether collaboration stems from a relationship such as between students and mentors (e.g., PhD students or early career researchers and supervisors) (Floyd et al, 1994).

Thirdly, scholars might collaborate several times on different occasions leading to multiple publications, while some co-authorship events might happen only once. This repeated collaboration relevantly affects the *turnover* of the authors in the sample. Some may enter the sample, publish only a few times and exit the pool of actively publishing authors, while some authors stay for longer time and continue to publish actively (Akbaritabar et al, 2020; Palla et al, 2007). Controlling for homophily based on the latest year of publication evaluates this effect. We formulated three additional hypotheses based on these homophily effects. Our second hypothesis addressed homophily for research productivity (Hypotheses 2): *scholars with similar levels of research productivity have a higher tendency to collaborate with each other.* The third hypothesis addressed homophily for academic seniority (Hypotheses 3): *scholars with similar academic seniority have a higher tendency to collaborate with each other.* The fourth hypothesis addressed homophily for publication activity (Hypotheses 4): *having a similar date of latest publication (e.g., publishing until recently) increases the tendency to collaborate among scholars.*

Finally, SNT scholars emphasise the presence of communities within networks (Akbaritabar et al, 2020; Palla et al, 2007; Sciabolazza et al, 2017). Algorithms are used to examine whether collaboration patterns persist *between* or *within* denser areas of the network (groupings). These communities significantly influence the overall

connectivity of the network and were for this reason addressed in our fifth hypothesis on social closure among groups (H5): *there are cohesive communities in the co-authorship network with a higher tendency to collaborate among community members.*

## 4 | DATA

There is not an uncontested list of higher education journals, thus, we followed previous studies by adopting as a selection criterion journals whose name contains the terms *higher education* or *tertiary education* (Jung & Horta, 2013; Kim et al, 2017; Liu et al, 2019; Lovakov & Yudkevich, 2021). We did not consider professional journals (e.g., *The Chronicle of Higher Education* and *Times Higher Education*) or discipline-specific journals (e.g., *Arts and Humanities in Higher Education*), instead concentrating the analysis on the core of the higher education studies by avoiding discipline-based influences (Tight, 2018). We selected 33 specialist higher education journals (Table 1) consistent with previous studies (Jung & Horta, 2013; Kim et al, 2017; Liu et al, 2019; Lovakov & Yudkevich, 2021). We then queried Scopus 2019 from the German Bibliometrics competence centre (Kompetenzzentrum Bibliometrie, 2020) for all *articles* and *reviews* published in the 33 journals from 1996 to 2018. The dataset includes various metadata for each publication, including publication year, title, author name(s), affiliation addresses and journal name.

### 4.1 | Disambiguation process and methodology

Data delivered by Scopus are prone to errors, and there is a strong need for disambiguation of institution and author names (D'Angelo & van Eck, 2020) without which the constructed networks will have multiple representations of the same actor and an artificially higher level of dis-connectivity (De Stefano et al, 2013). Therefore, we used affiliation addresses to disambiguate institutions using the Research Organization Registry (ROR) API (Research Organization Registry, 2020) which uses data from the Global Research Identifier Database (Global Research Identifier Database, 2020), International Standard Name Identifier, Crossref and Wikidata. ROR takes the different name spellings and misspelled words, acronyms and multiple languages into account and uses the full address strings that Scopus delivers. To disambiguate author names, we used Scopus's author identification numbers, which offer a somewhat improved data quality though they are still prone to error (Aman, 2018). Using this data, we constructed two-mode (bipartite) co-authorship networks. One mode is publications and for the other mode we constructed two sets of networks for authors and institutions.

Table 2 presents descriptive metrics for non-disambiguated and disambiguated networks at the levels of institutions and authors and illustrates the effect of disambiguation procedures. This is relevant to our first research goal to contribute to methods for disambiguation. The networks without any disambiguation present a higher disconnection, indicated by the *number of connected components* (e.g., 12,876 disconnected groups and 8,703 disconnected groups). Lower numbers of connected components are present in the disambiguated versions of the network (e.g., 786 disconnected groups and 7,998 disconnected groups) and point to the improvement achieved by disambiguation. Similarly, the reduced number of institutions (*Number of institutions/authors*), from 25,860 to 2,788 in the disambiguated version, indicates that out of every nine institution IDs, eight were redundant and only one should be considered. We exclude non-disambiguated institutions and their publications from the sample, which reduces the *Number of papers*, indicating that after disambiguation, our coverage of the publications is reduced from 17,262 to 13,875 (~20%) since we keep in the sample only publications for which all collaborating institutions are successfully disambiguated. The number of papers covered is higher at the *authors* level (99%), as Scopus numeric author IDs have a higher coverage; therefore, fewer publications are excluded (i.e., the publications with non-disambiguated authors without a Scopus author ID). Looking at the *number of connected*

**TABLE 1** Higher education research journals, total number of articles and reviews published 1996–2018

Journal	Number of articles and reviews 1996–2018
<i>Higher Education</i>	1,744
<i>Studies in Higher Education</i>	1,456
<i>Assessment and Evaluation in Higher Education</i>	1,308
<i>Industry and Higher Education</i>	964
<i>Teaching in Higher Education</i>	868
<i>Higher Education Research and Development</i>	857
<i>Research in Higher Education</i>	855
<i>International Journal of Sustainability in Higher Education</i>	641
<i>Higher Education Policy</i>	618
<i>Tertiary Education and Management</i>	608
<i>Journal of Higher Education Policy and Management</i>	599
<i>Internet and Higher Education</i>	574
<i>Journal of Further and Higher Education</i>	569
<i>Journal of Higher Education</i>	546
<i>Innovative Higher Education</i>	531
<i>Higher Education Quarterly</i>	488
<i>European Journal of Education</i>	483
<i>Review of Higher Education</i>	450
<i>Journal of Continuing Higher Education</i>	432
<i>Perspectives: Policy and Practice in Higher Education</i>	391
<i>Journal of Marketing for Higher Education</i>	321
<i>Journal of Applied Research in Higher Education</i>	289
<i>Active Learning in Higher Education</i>	287
<i>Quality in Higher Education</i>	280
<i>Journal of Diversity in Higher Education</i>	244
<i>European Journal of Higher Education</i>	230
<i>Higher Education, Skills and Work-based Learning</i>	222
<i>Journal of Higher Education Outreach and Engagement</i>	134
<i>International Perspectives on Higher Education Research</i>	100
<i>NASPA Journal About Women in Higher Education</i>	89
<i>Higher Education Pedagogies</i>	55
<i>Higher Education in Europe</i>	19
<i>Tuning Journal for Higher Education</i>	10
Total publications studied	17,262

Source: Scopus.

components, it is evident that disambiguation by Scopus author ID does not offer a perfect improvement, since each author can have multiple IDs (Aman, 2018). Networks built using each of these data inherit different qualities (De Stefano et al, 2013). At the institution level, the share of nodes in the giant component, which is the largest connected group of nodes (i.e., % of bipartite nodes in  $G$ ) is increased from 11% to 88%, which is closer to expected

**TABLE 2** Higher education studies co-authorship networks using non-disambiguated and disambiguated data at the institution and author levels (G = giant component of the network, SID = Scopus numeric IDs)

Metric/Type of network	Institution level		Author level	
	Non disambiguated institutions	Disambiguated institutions	Non disambiguated authors	Disambiguated authors SID
Number of connected components	12,876	786	8,703	7,998
Number of bipartite nodes	43,122	16,663	55,546	41,997
Number of bipartite edges	30,549	19,098	53,024	37,382
% of bipartite nodes in G	11	88	19	23
% of bipartite edges in G	15	93	26	31
Number of institutions/authors	25,860	2,788	38,284	24,744
Number of institutions/authors in G	2,395	1,895	6,336	4,725
Number of publications	17,262	13,875	17,262	17,253
Number of publications in G	2,144	12,716	4,167	4,779

Source: Authors.

rates. While Scopus author IDs offer some improvement, indicating an increase in share of nodes in the giant component from 19% to 23%, the result is not comparable to the accuracy achieved through institution name disambiguation.

In order to respond to our second research goal pertaining to features of co-authorship networks, we have computed the shares of single-country and multi-country (international) co-authored higher education publications over time. For our third research goal, pertaining to how network-based mechanisms explain the dynamics and growth of co-authorship networks, we developed Bipartite Exponential Random Graph Models (Bipartite ERGMs) (Wang et al, 2013) to test the aforementioned hypotheses (Hypotheses 1–Hypotheses 4). This statistical model presents the probability of tie existence (dependent variable) resulting from the effects of both endogenous and exogenous factors (Zhang et al, 2018).

In order to identify communities of co-authorship—in line with our third goal and fifth hypothesis—we used bipartite community detection, specifically, the Constant Potts Model (CPM) (Traag et al, 2011, 2019) implemented in *Leidenalg* library in Python. CPM uses a resolution parameter  $\gamma$  (*constant*), leading to communities where the link density between the communities is less than  $\gamma$  and where the link density within communities is greater than  $\gamma$  (Traag et al, 2011). We set  $\gamma$  to  $6 \times 10^{-5}$  in the case of *institutions*. We chose this parameter after comparing the number of communities detected against the number of institutions and publications included in each community to arrive at a consistent distribution without too many or too few communities or community members. Community detection at the level of *authors* was not performed due to lack of disambiguated and high-quality data.



Moreover, we extracted the most frequent research themes and topics from the *abstract* and *titles* of the identified studies on higher education using VOSviewer (Liu et al 2019; van Eck & Waltman, 2010). VOSviewer is a software tool for constructing and visualising bibliometric networks. It uses the VOS algorithm and places terms that more frequently appear together in documents at a closer distance on a term map. Each publication can be the product of a single collaboration or a collaboration between multiple institutions. We then calculated relevance scores based on the composition of these contributing institutions for each publication, which allowed us to investigate whether and how the network communities, detected through our community detection procedure, also vary in terms of research themes addressed and whether there is a thematic focus among different communities. We set the number of term clusters to be detected to two, to check alignment with the distinction between *learning and teaching* versus *policy-based* studies highlighted in Tight (2004). We are aware that this dichotomy might appear reductionist in light of the more recent and detailed classifications of higher education themes (Daenekindt & Huisman, 2020; Liu et al 2019), but some empirical studies have demonstrated its functionality for understanding the thematic focusses of the higher education communities (Horta & Jung, 2014; Kim et al, 2017), which is consistent with our purposes.

## 5 | FINDINGS

Looking at the trend of publications authored by institutions from a single country versus multiple countries we found that most of the publications (90%) were authored by institutions within a single country, which aligns with Kuzhabekova et al (2015) and papers with regional foci (Jung & Horta, 2013; Lovakov & Yudkevich, 2021).

Table 3 presents the results of the bipartite ERG models used for testing our first four hypotheses (Hypotheses 1–Hypotheses 4) on structural (preferential attachment) and homophily effects. It presents two sets of models at the level of authors (Models 1, 2, 3) and institutions (Models 4, 5, 6). Models 3 and 6 are the full models at each level, including all structural and homophily effects.

### 5.1 | Preferential attachment (Hypotheses 1)

Results indicate that there is a significant tendency towards lower preferential attachment at the *authors* level and higher preferential attachment on the *institutions* level (Models 1, 4), confirming our first hypothesis (H1) for institutions while for authors, due to disambiguation issues, we won't expect the results to be reliable. In other words, a few key institutions conducting research on higher education attract several co-authorship ties, while the same cannot be observed for authors. This effect stays similar when considering homophily effects. However, since Model 3 (full model, authors level) did not converge due to the high complexity of effects, we cannot claim that its results are as reliable as those for Model 6 on the *institutions* level.

### 5.2 | Homophily for research productivity (Hypotheses 2)

Similarity in the total number of publications is modelled as a continuous variable in our ERGMs in *absolute difference*, meaning the effect is interpreted in reverse. This similarity increases the probability of co-authorship tie formation at the *authors* and *institutions* levels, while this effect is statistically significant only at the *authors level* (Models 2, 5). These effects stay similar in the full models (Models 3, 6), which indicates the presence of an *exclusive club* (a small group with several co-authorships) of authors and institutions who tend to collaborate more among themselves.

**TABLE 3** Bipartite ERGMs' results (giant component of the higher education studies networks)

	Author level			Institution level		
	1	2	3	4	5	6
Ties	-7.801*** (0.012)	-7.723** (3.750)	-7.721*** (0.092)	-6.966*** (0.033)	-0.546 (9.306)	17.129*** (0.367)
Preferential attachment (Hypotheses 1)	1.934*** (0.120)		2.996*** (0.003)	-5.913*** (0.081)		-5.917*** (0.079)
Homophily for research productivity (Hypotheses 2)		-0.010** (0.004)	-0.014** (0.005)		-0.001 (0.001)	-0.002 (0.002)
Homophily for academic seniority (Hypotheses 3)		-0.007** (0.003)	-0.010*** (0.004)		-0.004 (0.003)	-0.011** (0.005)
Homophily for publication activity (Hypotheses 4)		0.007** (0.003)	0.009*** (0.004)		-0.0004 (0.005)	-0.001 (0.005)
Akaike Inf. Crit. (AIC)	196,515.700	198,000.600	196,577.700	45,834.690	55,290.900	44,737.440
Bayesian Inf. Crit. (BIC)	196,560.500	198,060.300	196,667.300	45,879.680	55,350.890	44,827.430

Note: significance levels: \* =  $p < 0.1$ ; \*\* =  $p < 0.05$ ; \*\*\* =  $p < 0.01$ .

Source: Authors.

### 5.3 | Homophily in academic seniority (Hypotheses 3)

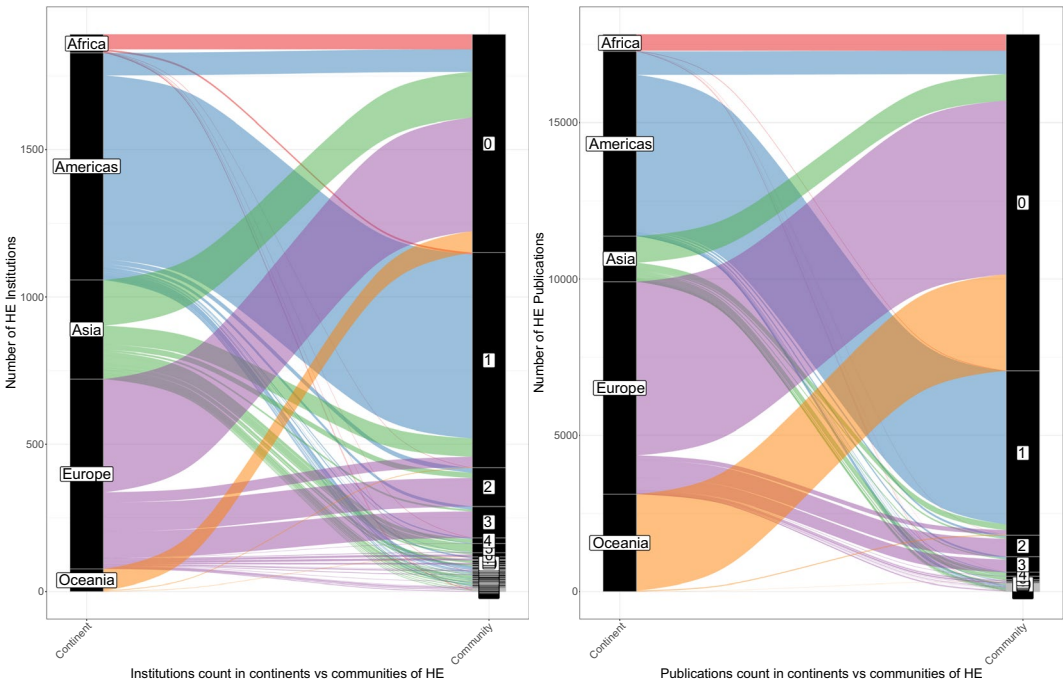
Similar first publication years, here a proxy for closer academic cohorts, present a higher probability of engaging in co-authorship at both the *authors* and *institutions* levels. Although we cannot control for authors' age due to the data, this finding might signal that early career researchers do not regularly publish with their mentors. The homophily in academic seniority effect was not significant at the *institutions* level (Model 5) until when the preferential attachment effect is entered in the full model (6), indicating a higher probability of collaboration among institutions with similar levels of seniority. In other words, this higher probability emphasises that institutions that have been historically active in higher education studies collaborate more among themselves.

### 5.4 | Homophily in publication activity (Hypotheses 4)

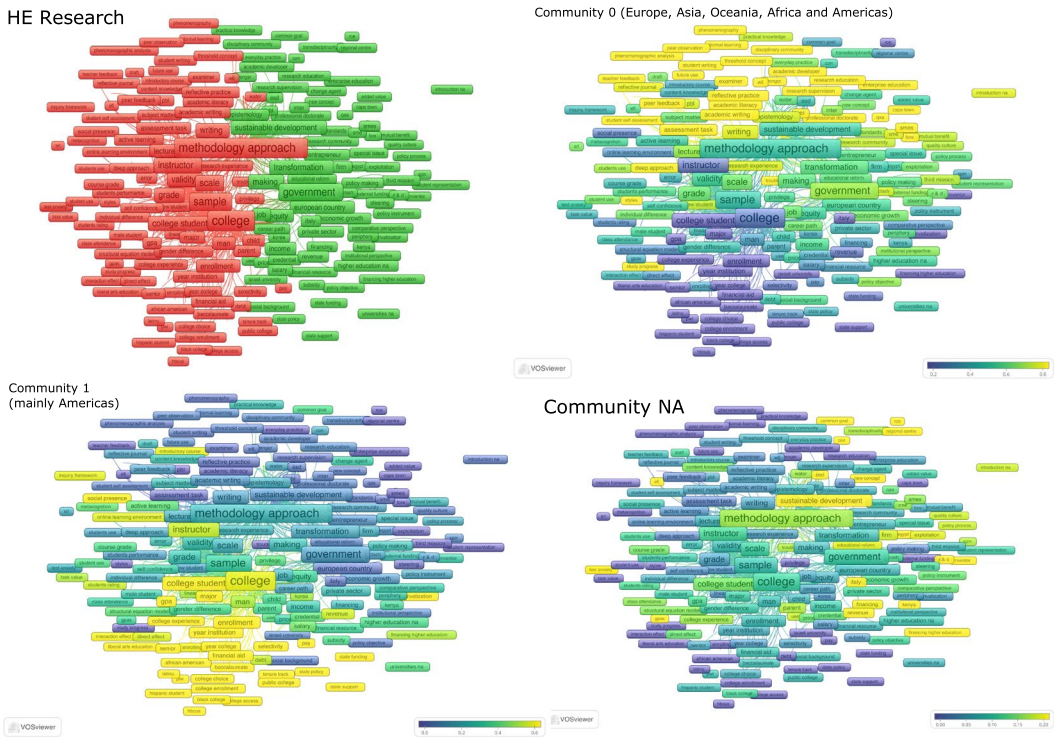
Having a similar date of latest publication (e.g., publishing until recently) decreased the probability of a tie at the *authors* level but increased it at the *institutions* level. However, this effect was significant only at the *authors* level, indicating that co-authorship events among authors with similar publication activity can be one-off incidents that do not occur regularly.

### 5.5 | Social closure among groups (H5)

Our community detection procedure identified 36 communities from the giant component of the co-authorship network at the institutions level which confirms our fifth hypothesis (H5). Although the giant component



**FIGURE 1** Distribution of HE studies institutions from different continents versus their membership in 36 HE studies communities detected from the giant component of network (left panel shows size of communities based on institutions, right panel shows the productivity based on publications). Source: Authors



**FIGURE 2** Term maps showing the research theme focusses of HE studies publications of community 0, community 1 and those institutions not a member of the giant component (community NA). The yellow colour shows the research theme focus (except top left, L&T (red), policy-based studies (green)). Source: Authors

is connected, it consisted of 36 denser areas that are cohesive subgroups presenting a higher number of co-authorship ties *within* themselves than *between* these different groups. We analysed the geographical composition of 36 communities that emerged from the data; specifically, to better understand who were members of the communities. We analysed the communities both in terms of size (number of institutions) and productivity. We observed communities of similar sizes that presented drastically different rates of productivity. Community 0 emerged as a community with members predominantly from *Europe, Asia, Oceania, and Africa* and a small share of the *Americas*. In contrast, higher education research institutions from the *Americas* mostly populated community 1. Another interesting observation is the split of *Asian and European* research institutions on higher education across multiple communities of smaller sizes that have differing productivity rates. *Oceania* is an interesting case where a small number of research institutions conducting studies on higher education, mainly located in community 0, have a high rate of productivity and collaborate with different institutions globally (see Figure 1).

Moreover, we investigated whether the identified communities, especially the two largest (communities 0 and 1), display a particular focus on higher education studies themes. We used term maps for identifying research theme foci; specifically, by identifying research topics in the focus of higher education research publications based on a co-occurrence of terms in *title* and *abstract*. Our analysis highlights two main clusters of research themes, namely, the *learning and teaching approach* and *policy-based studies*. Our findings are consistent with previous studies (Kim et al, 2017; Tight, 2004). Term maps were constructed for each community. Community 0 is dominated by Europe, Asia, Oceania and Africa. Community 1 is dominated by North, Central and South America, i.e., the Americas. Community NA stands for Not Available and represents those institutions who were not members of the giant component of higher education studies network and hence not included in community detection. The

focus of community 1 was largely skewed towards learning and teaching, research themes revolving around topics such as *college choice* and *college experience*, with only the *privatisation* topics coming from the policy-studies domain. The focus of community 0 instead bridges comparatively more the policy themes with the learning and teaching themes. A stark contrast is evident in comparison to themes covered by the American higher education community (i.e., community 1). While the policy versus learning and teaching themes divide is much more evident within the American community, it is equally evident that policy-based studies are comparatively more prominent for higher education authors in community NA who were not part of the giant component (see Figure 2).

## 6 | DISCUSSION AND CONCLUSIONS

The present article investigated the internationalisation of higher education studies through a quantitative analysis of co-authorship networks. We articulated the disambiguation of research networks as a foundational research goal. Two additional goals pertained to developing methods and understanding the growth of co-authorship networks. The latter two goals were specific to higher education studies: in methodological terms, how the accuracy of network construction may improve as a result of disambiguation of institution and author names; to what extent higher education studies are international and how network-based mechanisms explain the growth of higher education co-authorship networks.

In response to our first goal, regarding the differences between disambiguated and non-disambiguated networks, our findings suggest that the method we have used improves the identification of co-authorship networks particularly on the level of institutions (Aman, 2018; D'Angelo & van Eck, 2020; De Stefano et al, 2013). While at the author level, the accuracy was lower and affected the results to a certain extent. This shortcoming applies to empirical research that takes data from bibliometric databases as uncontested truth without probing into problematic aspects of the data.

In response to our second research goal concerning the internationalisation of studies on higher education, we found that studies on higher education are not internationally oriented, as also confirmed by previous research (Avdeev, 2019; Horta & Jung, 2014; Jung & Horta, 2013; Kwiek, 2020; Lovakov & Yudkevich, 2021; Kuzhabekova et al, 2015). Publications by authors from a single country remain more prevalent (90%). Internationalisation in research on higher education is significantly lower compared to internationalisation in the social sciences (Acedo et al, 2006; Akbaritabar et al, 2018). To reverse this trend, a stronger involvement of English-speaking countries (especially the US, Australia and the UK), which are the most productive countries of higher education studies (Calma & Davies, 2017; Liu et al, 2019) seems crucial to build a more integrated and international community.

The second part of the analysis, corresponding to our third research goal, employs network-based mechanisms to investigate the nature of collaborations in higher education studies (Biancani & McFarland, 2013; Wagner & Leydesdorff, 2005). Preferential attachment (Hypotheses 1) emerged as a significant explanatory mechanism at the *institutions* level (this is lower for the authors due to disambiguation issues). Therefore, new co-authorships are formed primarily with institutions who already present many co-authorships, similar to other disciplines (Barabási et al, 2002; Moody, 2004; Sciabolazza et al, 2017; Wagner & Leydesdorff, 2005; Zhang et al, 2018). Regarding homophily effects (Hypotheses 2–Hypotheses 4), seniority among authors and institutions was associated with collaboration among peers. Prolific authors also show a higher tendency to collaborate with other prolific authors, hence forming an exclusive club of highly prolific authors and institutions co-authoring *among* themselves. Other homophily effects are associated with gender, language and PhD granting university, to name a few (Biancani & McFarland, 2013; Zhang et al, 2018). However, due to limitations of the data, we did not test these homophily effects in the study on which this article reports.

Community detection allowed us to confirm our fifth hypothesis regarding the occurrence of cohesive communities. These cohesive subgroups present a higher number of co-authorships among community members than between different groups in the network. These communities differ both in terms of geographical location and

research-focus. The two main communities (0 and 1) highlight a stark division between the Americas (community 1) and the rest of the world (community 0) (Tight, 2007, 2014). Scholars and institutions from Europe are part of different research communities with representation from other continents; this demonstrates a higher degree of collaboration both within Europe and between Europe and other regions worldwide, when compared to the Americas. Similarly, Asia is involved in different communities but is mainly in community 0 with Europe, Oceania and Africa. Oceania was found to be involved mainly in community 0, meaning that it almost exclusively collaborates with European and Asian scholars.

However, higher education studies are not only characterised by a geographical divide. The divide in research focused on *learning and teaching* approaches and research on *policies* is notable (Tight, 2004). Our analysis shows how the geographical division of higher education studies also overlaps with a different research themes focus displayed by the two largest communities. The community associated with the Americas is much more skewed towards specific research themes from the *learning and teaching* domain. Yet, these are different from those covered by other research communities in the world. Other research communities covered comparatively more the traditional divide between *policy studies* and studies on *learning and teaching*, as highlighted in previous works (Horta & Jung, 2014; Kim et al, 2017). In this sense, the comparatively more international composition of community 0 somewhat corroborates Liu et al.'s (2019) insight according to which a higher similarity of research themes also favours international research collaborations. The observed dearth of collaboration between Europe-Asia-Oceania and the Americas confirms the findings from previous studies that identified a clear divide among the more international European higher education journals and the *less open* North American ones (Tight, 2007, 2014). These findings should be taken as an opportunity for higher education journals to reflect on their attitudes towards both authors' nationality and local versus global policy issues. Efforts to balance the observed tendencies can have an impact on the current division that our analysis has observed in research co-authorships between the Europe-Asia-Oceania block and the Americas.

Limitations to our analysis stem in part from the data collection. While we tried to keep the list of journals as international as possible, we did not include articles that deal with higher education but are published in non-specialist higher education journals (Horta & Jung, 2014). Publications in national journals that are not indexed in Scopus were not included. Additionally, our study addressed only scientific collaborations in the form of co-authorships, while collaborations can take multiple forms (funding proposals and co-participation in scientific workshops and projects) (Katz & Martin, 1997). Finally, although ERG models offer a great framework for statistical analysis of networks, using them in large networks similar to the ones in this article causes degeneracy issues, limiting possibilities for exploring further hypotheses.

## DATA AVAILABILITY STATEMENT

The data used in this study is from Kompetenzzentrum Bibliometrie. Restrictions apply to the availability of these data, which were used under license for this study. Data cannot be made publicly available due to the licensing and contract terms of the original data.

## ORCID

Aliakbar Akbaritabar  <https://orcid.org/0000-0003-3828-1533>

Giovanni Barbato  <https://orcid.org/0000-0002-4209-6464>

## REFERENCES

- Acedo, F. J., Barroso, C., Casanueva, C., & Galán, J. L. (2006). Co-authorship in management and organizational studies: An empirical and network analysis. *Journal of Management Studies*, 43(5), 957-983. <https://doi.org/10.1111/j.1467-6486.2006.00625.x>
- Akbaritabar, A., Casnici, N., & Squazzoni, F. (2018). The conundrum of research productivity: A study on sociologists in Italy. *Scientometrics*, 114(3), 859-882. <https://doi.org/10.1007/s11192-017-2606-5>



- Akbaritabar, A., Traag, V. A., Caimo, A., & Squazzoni, F. (2020). Italian sociologists: A community of disconnected groups. *Scientometrics*, 124, 2361–2382. <https://doi.org/10.1007/s11192-020-03555-w>
- Aman, V. (2018). Does the Scopus author id suffice to track scientific international mobility? A case study based on Leibniz laureates. *Scientometrics*, 117(2), 705–720. <https://doi.org/10.1007/s11192-018-2895-3>
- Avdeev, S. (2019). *International collaboration in higher education research: A gravity model approach*. Research Paper No. WP BRP, 54. National Research University—Higher School of Economics, Moscow. <https://doi.org/10.2139/ssrn.3505886>
- Barabási, A.-L., Jeong, H., Nédai, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and Its Applications*, 311(3–4), 590–614. [https://doi.org/10.1016/S0378-4371\(02\)00736-7](https://doi.org/10.1016/S0378-4371(02)00736-7)
- Biancani, S., & McFarland, D. A. (2013). Social networks research in higher education. In M. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 151–215). Springer. [https://doi.org/10.1007/978-94-007-5836-0\\_4](https://doi.org/10.1007/978-94-007-5836-0_4)
- Calma, A., & Davies, M. (2017). Geographies of influence: A citation network analysis of higher education 1972–2014. *Scientometrics*, 110(3), 1579–1599. <https://doi.org/10.1007/s11192-016-2228-3>
- D'Angelo, C. A., & van Eck, N. J. (2020). Collecting large-scale publication data at the level of individual researchers: A practical proposal for author name disambiguation. *Scientometrics*, 123(2), 883–907. <https://doi.org/10.1007/s11192-020-03410-y>
- Daenekindt, S., & Huisman, J. (2020). Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991–2018. *Higher Education*, 80, 1–17. <https://doi.org/10.1007/s10734-020-00500-x>
- Dahlander, L., & McFarland, D. A. (2013). Ties that last: Tie formation and persistence in research collaborations over time. *Administrative Science Quarterly*, 58(1), 69–110. <https://doi.org/10.1177/0001839212474272>
- De Stefano, D., Fuccella, V., Vitale, M. P., & Zaccarin, S. (2013). The use of different data sources in the analysis of co-authorship networks and scientific performance. *Social Networks*, 35(3), 370–381. <https://doi.org/10.1016/j.socnet.2013.04.004>
- Floyd, S. W., Schroeder, D. M., & Finn, D. M. (1994). “Only if I'm first author”: Conflict over credit in management scholarship. *Academy of Management Journal*, 37(3), 734–747. <https://doi.org/10.5465/256709>
- Fumasoli, T. (2019). Learning from internationalisation scholarship in higher education: Commonalities, divergences and possible research directions. In M. Yemini, C. Maxwell, & L. Engel (Eds.), *Beyond the established boundaries: The machinery of school internationalisation in action* (pp. 165–178). Routledge. <https://discovery.ucl.ac.uk/id/eprint/10085108>
- Global Research Identifier Database. (2020). <https://www.grid.ac/pages/policies>
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380. <https://doi.org/10.1086/225469>
- Horta, H., & Jung, J. (2014). Higher education research in Asia: An archipelago, two continents or merely atomization? *Higher Education*, 68(1), 117–134. <https://doi.org/10.1007/s10734-013-9695-8>
- Jung, J., & Horta, H. (2013). Higher education research in Asia: A publication and co-publication analysis. *Higher Education Quarterly*, 67(4), 398–419. <https://doi.org/10.1111/hequ.12015>
- Jung, J., & Horta, H. (2015). The contribution of East Asian countries to internationally published Asian higher education research: The role of system development and internationalisation. *Higher Education Policy*, 28(4), 419–439. <https://doi.org/10.1057/hep.2015.15>
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18. [https://doi.org/10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1)
- Kehm, B. M. (2015). Higher education as a field of study and research in Europe. *European Journal of Education*, 50(1), 60–74. <https://doi.org/10.1111/ejed.12100>
- Kezar, A. (2014). Higher education change and social networks: A review of research. *The Journal of Higher Education*, 85(1), 91–125. <https://doi.org/10.1080/00221546.2014.11777320>
- 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. <https://doi.org/10.1080/03075079.2015.1036850>
- Kompetenzzentrum Bibliometrie. (2020). *German competence centre for bibliometrics*. Website. Retrieved in 2020 from <http://bibliometrie.info/>
- Kosmützky, A., & Krücken, G. (2014). Growth or steady state? A bibliometric focus on international comparative higher education research. *Higher Education*, 67(4), 457–472. <https://doi.org/10.1007/s10734-013-9694-9>
- Kuzhabekova, A., Hendel, D. D., & Chapman, D. W. (2015). Mapping global research on international higher education. *Research in Higher Education*, 56(8), 861–882. <https://doi.org/10.1007/s11162-015-9371-1>
- Kwiek, M. (2020). What large-scale publication and citation data tell us about international research collaboration in Europe: Changing national patterns in global contexts. *Studies in Higher Education*, AHEAD-OF-PRINT, 1–21. <https://doi.org/10.1080/03075079.2020.1749254>
- Kwiek, M. (2021). The prestige economy of higher education journals: A quantitative approach. *Higher Education*, 81, 493–519. <https://doi.org/10.1007/s10734-020-00553-y>

- Liu, S., Liu, M., Jiang, H., Lin, Y., & Xu, K. (2019). International comparisons of themes in higher education research. *Higher Education Research & Development*, 38(7), 1445–1460. <https://doi.org/10.1080/07294360.2019.1654438>
- Lovakov, A., & Yudkevich, M. (2021). The post-Soviet publication landscape for higher education research. *Higher Education*, 81, 273–299. <https://doi.org/10.1007/s10734-020-00541-2>
- Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology, & Human Values*, 17(1), 101–126. <https://doi.org/10.1177/016224399201700106>
- Macfarlane, B. (2012). The higher education research archipelago. *Higher Education Research & Development*, 31(1), 129–131. <https://doi.org/10.1080/07294360.2012.642846>
- Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810), 56–63. <https://doi.org/10.1126/science.159.3810.56>
- Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review*, 69(2), 213–238. <https://doi.org/10.1177/000312240406900204>
- Newman, M. E. (2001a). Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2), 1–4. <https://doi.org/10.1103/PhysRevE.64.025102>
- Palla, G., Barabási, A. L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446, 664–667. <https://doi.org/10.1038/nature05670>
- Research Organization Registry. (2020). ROR API. Website. Retrieved in 2020 from <https://ror.org/about/>
- Sciabolazza, V. L., Vacca, R., Okraku, T. K., & McCarty, C. (2017). Detecting and analyzing research communities in longitudinal scientific networks. *PLoS ONE*, 12(8), e0182516. <https://doi.org/10.1371/journal.pone.0182516>
- Teichler, U. (2000). Higher education research and its institutional basis. In S. Schwarz & U. Teichler (Eds.), *The institutional basis of higher education research* (pp. 13–24). Kluwer Academic Publishers. <https://doi.org/10.1007/0-306-47662-2>
- Tight, M. (2004). Research into higher education: An a-theoretical community of practice? *Higher Education Research & Development*, 23(4), 395–411. <https://doi.org/10.1080/0729436042000276431>
- Tight, M. (2007). Bridging the divide: A comparative analysis of articles in higher education journals published inside and outside north America. *Higher Education*, 53(2), 235–253. <https://doi.org/10.1007/s10734-005-2429-9>
- 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 & Development*, 37(3), 607–619. <https://doi.org/10.1080/07294360.2017.1389858>
- Traag, V. A., Van Dooren, P., & Nesterov, Y. (2011). Narrow scope for resolution-limit-free community detection. *Physical Review E*, 84(1), 1–9. <https://doi.org/10.1103/PhysRevE.84.016114>
- Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9(1), 1–12. <https://doi.org/10.1038/s41598-019-41695-z>
- van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- 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>
- Wagner, C. S., Park, H. W., & Leydesdorff, L. (2015). The continuing growth of global cooperation networks in research: A conundrum for national governments. *PLoS ONE*, 10(7), 1–15. <https://doi.org/10.1371/journal.pone.0131816>
- Wang, P., Pattison, P., & Robins, G. (2013). Exponential random graph model specifications for bipartite networks—A dependence hierarchy. *Social Networks*, 35(2), 211–222. <https://doi.org/10.1016/j.socnet.2011.12.004>
- Zhang, C., Bu, Y., Ding, Y., & Xu, J. (2018). Understanding scientific collaboration: Homophily, transitivity, and preferential attachment. *Journal of the Association for Information Science and Technology*, 69(1), 72–86. <https://doi.org/10.1002/asi.23916>

**How to cite this article:** Akbaritabar A, Barbato G. An internationalised Europe and regionally focused Americas: A network analysis of higher education studies. *Eur J Educ.* 2021;00:1–16. <https://doi.org/10.1111/ejed.12446>