

A Network Analysis Study on the Internationalization of Higher Education Research

Aliakbar Akbaritabar, Giovanni Barbato

akbaritabar@DZHW.eu, giovanni.barbato@UNIMI.it

INSNA Sunbelt Virtual Conference, July 2020

Introduction: Object of Study

- ▶ Internationalization of Higher Education (HE) research
- ▶ **International co-authorship:** when two or more scholars from different countries have co-authored one or more papers (Katz and Martin 1997; Kwiek 2018): steady growth + increasing authors, organizations and countries involved (Wagner, Park, and Leydesdorff 2015). However... social sciences less studied;
- ▶ HE research: multidisciplinary; emergent field; **high fragmentation**
- ▶ Marked division between *policy-based* and *learning* and *teaching* studies and scholars; low communication between them; lots of **part-time researchers** (Tight 2004; Horta and Jung 2014; Kim, Horta, and Jung 2017)

Three Research Goals

- ▶ **DESCRIPTIVE:** to investigate the main features of the international co-authorship network in HE research: how much HE research publications are international and collaborative
⇒ descriptive network analysis
- ▶ **EXPLANATORY:** to study how specific network-based mechanisms (e.g. preferential attachment and homophily) contribute to explain the growth and evolution of the international HE co-authorship network ⇒ Exponential Random Graph Model (ERGM), bipartite community detection: rarely applied to study HE research
- ▶ **METHODOLOGICAL:** to highlight what is the level of improvement in accuracy of network construction obtained by disambiguation of organization and author names ⇒ disambiguation procedures often overlooked

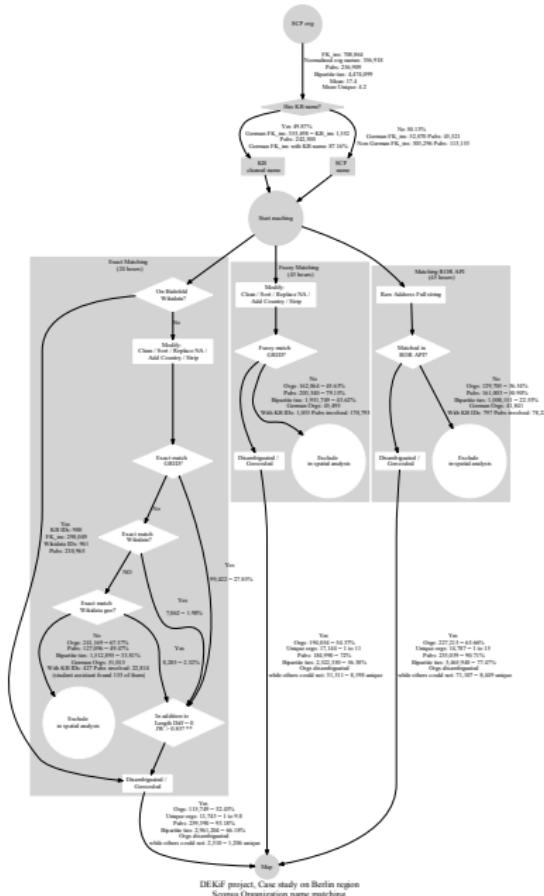
Hypotheses

- ▶ **H1 (Preferential attachment)**: scholars with several previous collaborations present a higher probability to establish new co-authorships.
- ▶ **H2 (Homophily for research productivity)**: Scholars with similar levels of research productivity have higher tendency to collaborate with each other.
- ▶ **H3 (Homophily for academic seniority)**: Scholars with similar academic seniority have higher tendency to collaborate with each other.
- ▶ **H4 (Homophily for publication activity)**: Scholars with similar length of publication activity have higher tendency to collaborate with each other.
- ▶ **H5 (Social closure among groups)**: There are cohesive communities in co-authorship network with higher tendency to collaborate among themselves.

Data Sources

- ▶ **33** main HE journals (as evaluated by experts of the field)
- ▶ *Article and Review*, 1996-2018, from in-house Scopus
- ▶ **GRID** (Global Research Identifier Database) 10th December 2019
- ▶ **Research Organization Registry (ROR)** [local] API 18th December 2019
- ▶ Scopus author numeric IDs

Evaluation of Organization Name Disambiguation



Bipartite Network Modeling, Why?

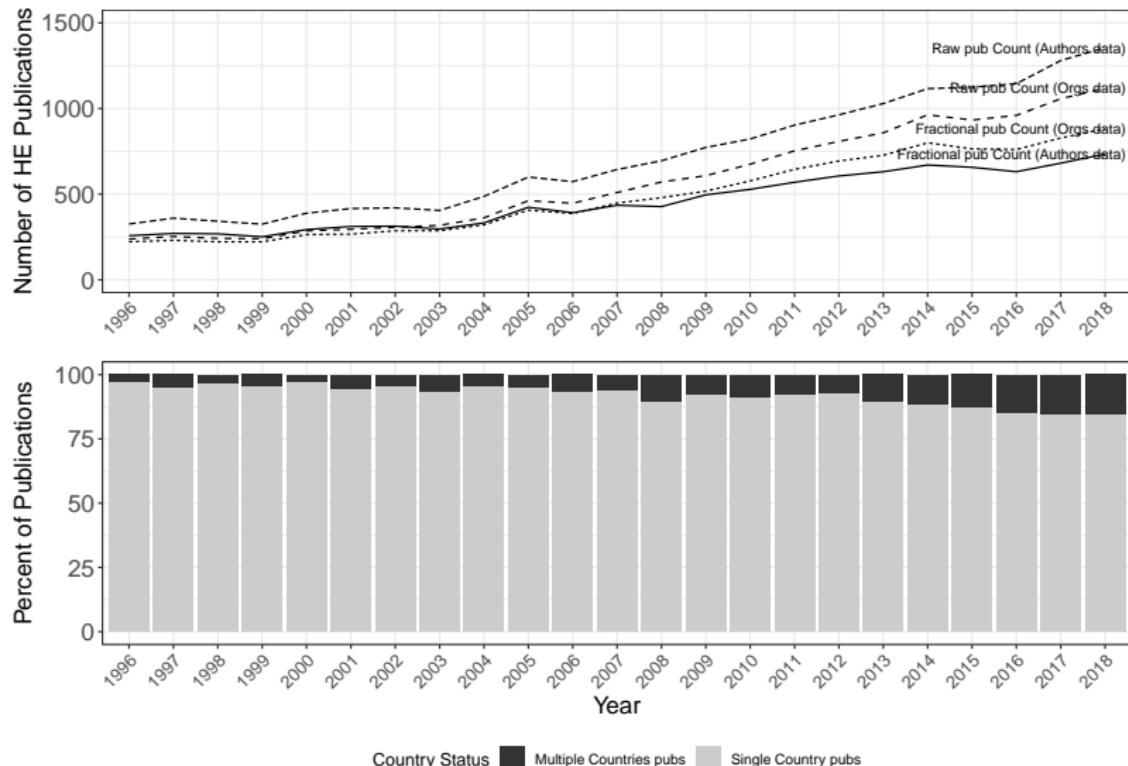
- ▶ Two node types: Organization/author and papers
- ▶ Two issues with *one-mode projection*
 1. Different structures yield same projected structure \Rightarrow information loss
 2. Articles with multiple authors \Rightarrow artificially high cliquish behavior
- ▶ Bipartite community detection with Constant Potts Model in Leidenalg library
- ▶ Emphasizes the importance of links *within* communities vs. *between*
- ▶ Communities such that the link density between the communities (*external density*) is lower than γ and the link density within communities (*internal density*) is higher than γ
- ▶ $\gamma = 6 \times 10^{-5}$, yield 36 communities of org/pubs
- ▶ Bipartite ERGMs: models probability of tie existence as dependent variable

Findings: Description of networks

- ▶ High improvement in network accuracy with **disambiguation** (1/9 orgs)

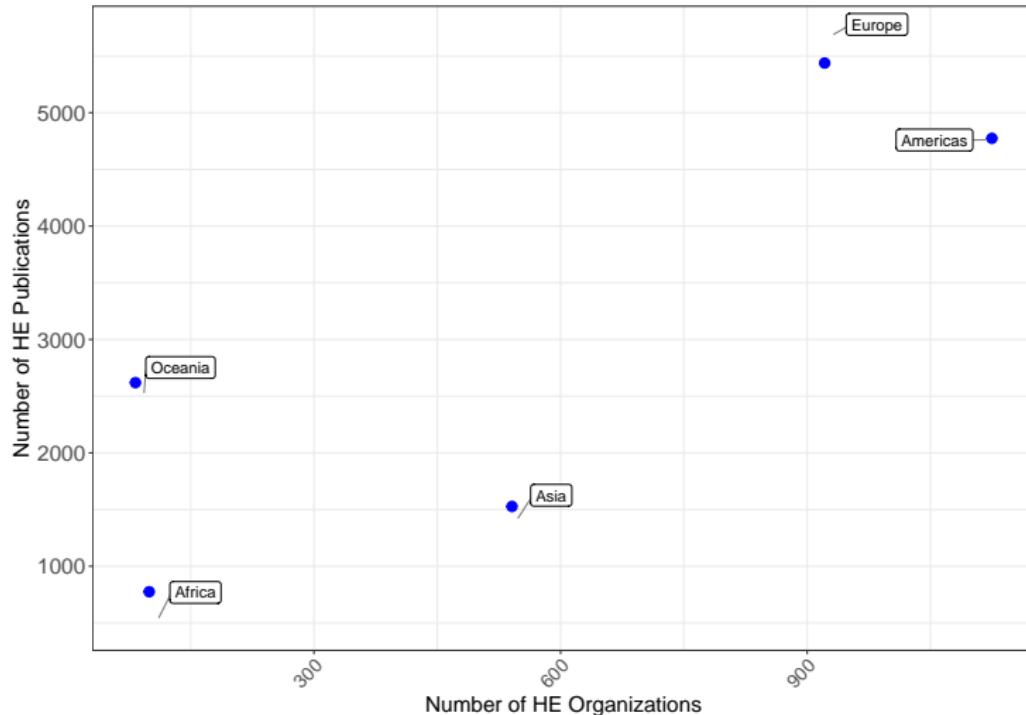
Metrics	Non disambiguated orgs	Disambiguated orgs	Non disambiguated authors	Disambiguated authors SID
Number of connected components	12,876	786	8,703	7,998
Number of bipartite nodes	43,122	16,663	55,546	41,997
Number of bipartite edges	30,549	19,098	53,024	37,382
% of bipartite nodes in G	11	88	19	23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

Higher collaboration. Internationalization?



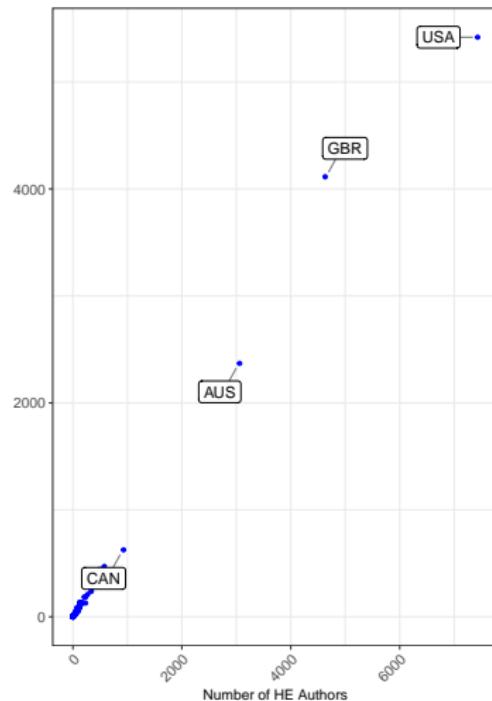
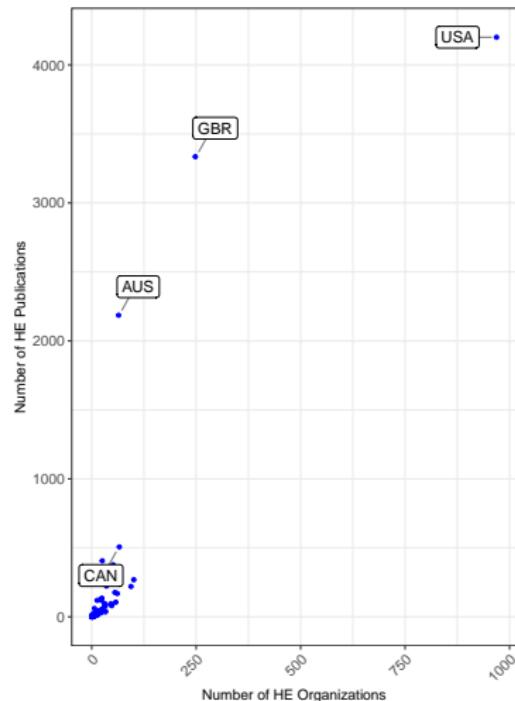
HE organizations vs. HE publications (continents)

- ▶ Europe is more prolific in the aggregate picture; Asia vs. Oceania.



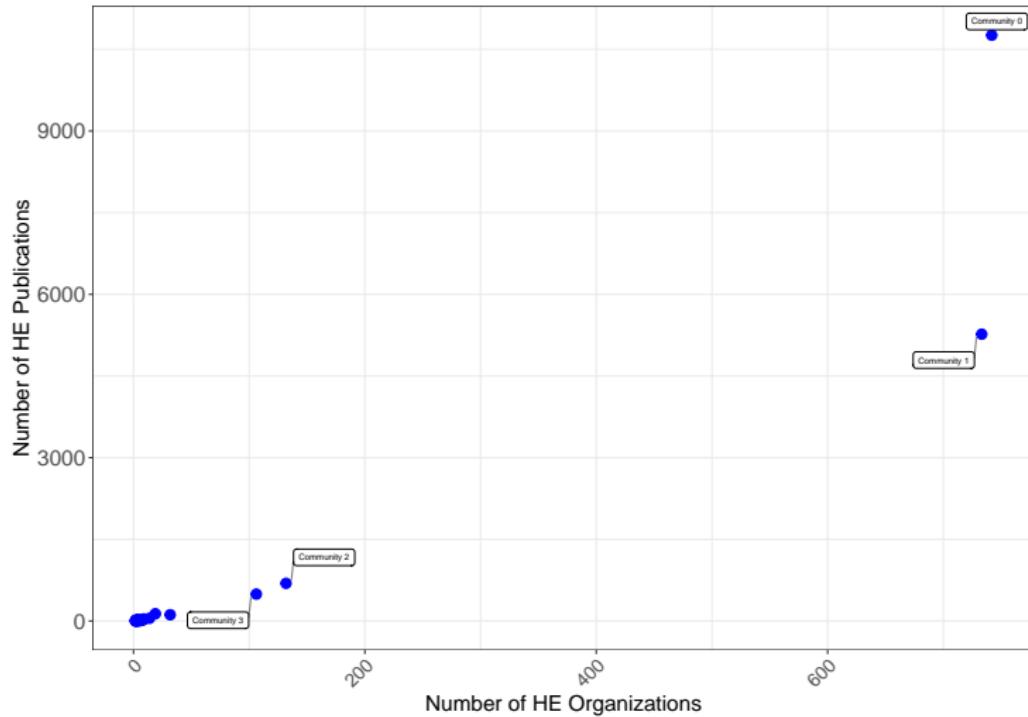
HE organizations/authors vs. HE publications (countries)

- ▶ In single country view, US is the most prolific
- ▶ GBR and AUS, smaller HE population, relatively prolific



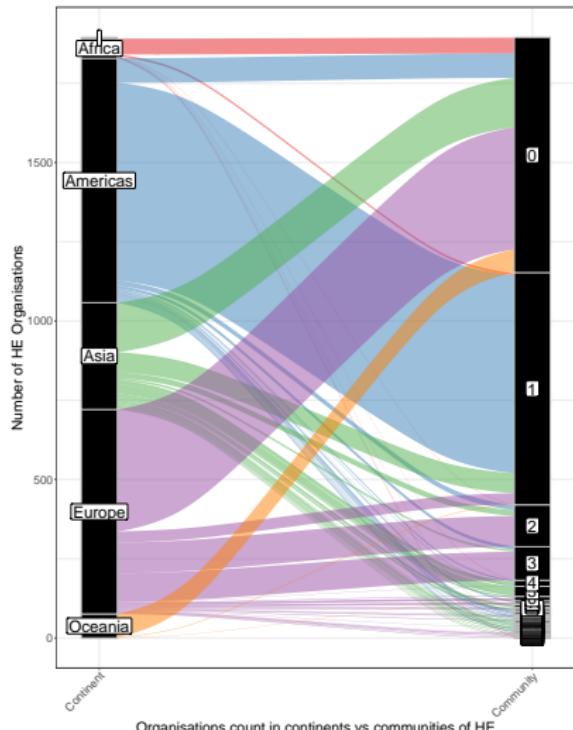
Composition of HE communities (1/2)

- ▶ H5: Larger communities not necessarily more prolific (see 1)

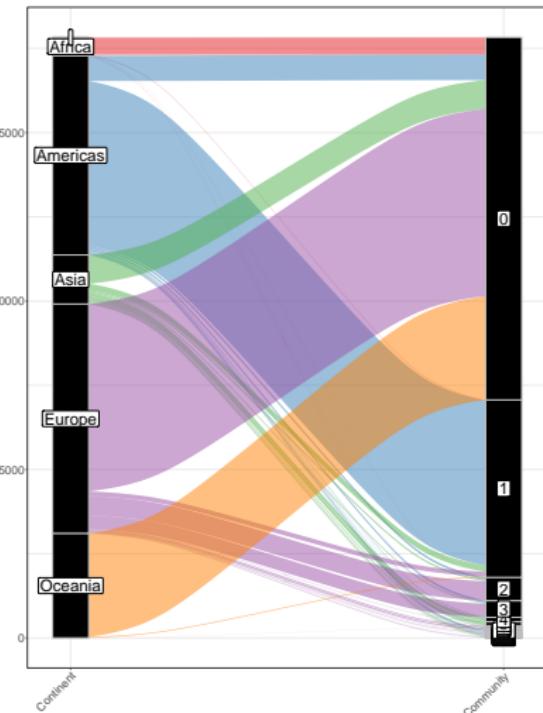


Composition of HE communities (2/2)

- H5: Fragmentation of HE network. Internationalized Europe vs. Continent Oriented Americas



Organisations count in continents vs communities of HE



Publications count in continents vs communities of HE

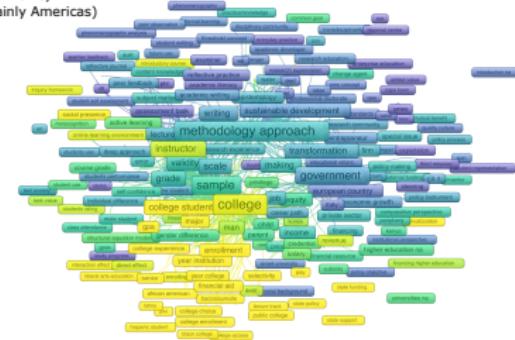
HE communities' substantial focus

- H5: Fragmentation of HE network. Clear divide in substantial focuses (substantional nationalism?)

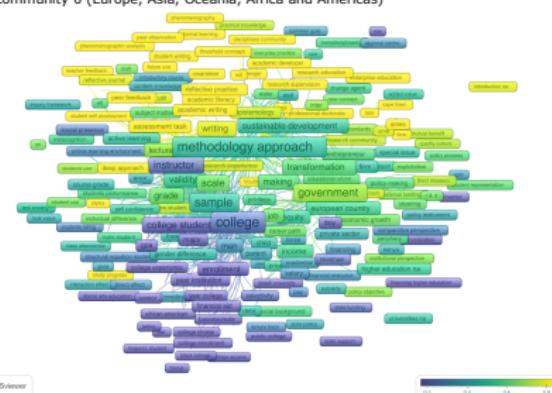
HE Research



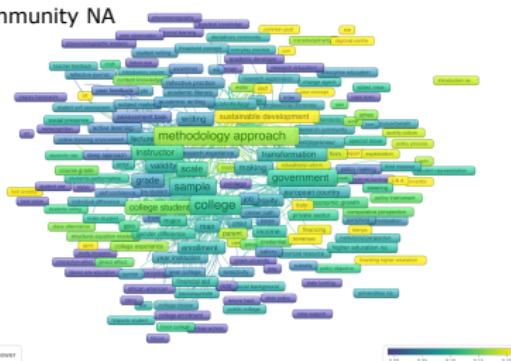
Community 1
(mainly Americas)



Community 0 (Europe, Asia, Oceania, Africa and Americas)



Community NA



Bipartite ERGM Results

- ▶ H1: Tendency towards lower preferential attachment in *authors* and higher preferential attachment in *organizations*
- ▶ Former could be due to disambiguation and multiple representation of actors
- ▶ H2: Homophily and higher tendency of collaboration among most prolific organizations and authors (rich club)
- ▶ H3: Homophily in similar date of first publication (e.g., higher tendency of collaboration among closer academic cohorts)
- ▶ H4: Similar date of latest publication (e.g., publishing until recently) decreased the probability of ties in *authors* and increased it in *organization*. Co-authorship, a **one-off incident** that does not occur regularly

Table 3: Bipartite ERGMs results

	Giant component of HE co-authorship					
	Author level			Organization 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 (H1)	1.934*** (0.120)		2.990*** (0.003)	-5.913*** (0.081)		-5.917*** (0.079)
Homophily for research productivity (H2)		-0.010** (0.004)	-0.014** (0.005)		-0.001 (0.001)	-0.002 (0.002)
Homophily for academic seniority (H3)		-0.007** (0.003)	-0.010*** (0.004)		-0.004 (0.003)	-0.011** (0.005)
Homophily for publication activity (H4)		0.007** (0.003)	0.009** (0.004)		-0.0004 (0.005)	-0.001 (0.005)
Akaike Inf. Crit.	196,515.700	198,000.600	196,577.700	45,834.690	55,290.900	44,737.440
Bayesian Inf. Crit.	196,560.500	198,060.300	196,667.300	45,879.680	55,350.890	44,827.430

Note:

*p<0.1; **p<0.05; ***p<0.01

Discussion and Conclusions 1/3

- ▶ **FIRST RG:** *how much are HE research publications international and collaborative?*
- ▶ Increasing degree of *team science* (Wuchty, Jones, and Uzzi 2007): from single-author to multi-authors publications
- ▶ But still a highly fragmented field: 768 disconnected components, 36 sub-groups of giant component
- ▶ Low degree of internationalization: publications by authors from a single country are more prevalent than multiple country ones (10%)
- ▶ Majority of internationally co-authored papers come from western countries: Europe and Americas have 3 times more HE publications than others (Kuzhabekova, Hendel, and Chapman 2015)

Discussion and Conclusions 2/3

- ▶ **SECOND RG:** *how network mechanisms contribute to explain evolution of the international HE co-authorship network?*
- ▶ Preferential attachment (H1): is an explanatory mechanism at the organization level, while its effect is lower at authors level; confirms literature on other discipline (Wagner and Leydesdorff 2005; Zhang et al. 2018)
- ▶ There are minority of *key organizations* while in the *authors* level there is not (disambiguation!)
- ▶ Homophily effects: homophily of research productivity (rich club) and similar academic cohorts.
- ▶ Community detection (H5): several cohesive sub-groups that present a higher number of co-authorship ties within themselves rather than between them ⇒ atomization of the HE field (confirms the literature & methodological nationalism?)

Discussion and Conclusions 3/3

- ▶ **THIRD RG:** what is the level of improvement in accuracy of network construction obtained by disambiguation of organization and author names?
- ▶ Out of each 9 organizations, 1 is reliable
- ▶ Scopus IDs, out of each 1.5 authors, 1 is reliable

Limitations

- ▶ In the data collection process: publications on HE outside HE journals are not included
- ▶ Journals not indexed in Scopus are not included
- ▶ We only cover **successful** scientific collaboration that results in a *publication* while as outlined by Katz and Martin (1997), scientific collaboration can take multiple forms (e.g., funding proposals co-authorship, patents)
- ▶ Further homophily effects: Gender, academic level, continents, countries, etc. not implemented due to non-converging models

Thanks for your attention!

- ▶ akbaritabar@DZHW.eu
- ▶ giovanni.barbato@UNIMI.it

Part of bibliography

- Horta, Hugo, and Jisun Jung. 2014. "Higher Education Research in Asia: An Archipelago, Two Continents or Merely Atomization?" *Higher Education* 68 (1). Springer: 117–34.
- Katz, J. S., and B. R. Martin. 1997. "What Is Research Collaboration?" *Research Policy* 26 (1): 1–18.
- Kim, Yangson, Hugo Horta, and Jisun Jung. 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). Taylor & Francis: 149–68.
- Kuzhabekova, Aliya, Darwin D Hendel, and David W Chapman. 2015. "Mapping Global Research on International Higher Education." *Research in Higher Education* 56 (8). Springer: 861–82.
- Kwiek, Marek. 2018. "International Research Collaboration and International Research Orientation: Comparative Findings About European Academics." *Journal of Studies in International Education* 22 (2). SAGE Publications Sage CA: Los Angeles, CA: 136–60.
- Tight, Malcolm. 2004. "Research into Higher Education: An a-Theoretical Community of Practice?" *Higher Education Research & Development* 23 (4). Taylor & Francis: 395–411.
- Wagner, Caroline S, and Loet Leydesdorff. 2005. "Network Structure, Self-Organization, and the Growth of International Collaboration in Science." *Research Policy* 34 (10). Elsevier: 1608–18.
- Wagner, Caroline S, Han Woo Park, and Loet Leydesdorff. 2015. "The Continuing Growth of Global Cooperation Networks in Research: A Conundrum for National Governments." *PLoS One* 10 (7). Public Library of Science.
- Wuchty, Stefan, Benjamin F Jones, and Brian Uzzi. 2007. "The Increasing Dominance of Teams in Production of Knowledge." *Science* 316 (5827). American Association for the Advancement of Science: 1036–9.
- Zhang, Chenwei, Yi Bu, Ying Ding, and Jian Xu. 2018. "Understanding Scientific Collaboration: Homophily, Transitivity, and Preferential Attachment." *Journal of the Association for Information Science and Technology* 69 (1). Wiley Online Library: 72–86.