

# Internationalized Europe vs. Continent Oriented Americas

A Network Analysis Study on the Internationalization of Higher Education Research

Aliakbar Akbaritabar (Ali)

[akbaritabar@gmail.com](mailto:akbaritabar@gmail.com)

@ akbaritabar

[akbaritabar.netlify.app](http://akbaritabar.netlify.app)

June 3<sup>rd</sup> 2020 - MPIDR

# Introduction

- I see myself as a *computational social scientist* who is interested in using his skills to answer a diverse array of research questions
- I use *computational tools* to carry out *social research* with these main steps: *ask questions, gather data, clean it, model it, interpret, report and publish* it
- Please see my personal webpage [akbaritabar.netlify.app](http://akbaritabar.netlify.app) for a detailed description and links of each step

## Overlap with *Digital and Computational Demography Lab*

- IMHO, as an outside observer
- My 1<sup>st</sup> PhD on **cyber social capital** and online survey on Facebook and Google Plus  $\Rightarrow$  *digital trace data* and *Technology Access over the Life Course*
- My work with large **bibliometric databases** (e.g., Scopus, Web of Science, Dimensions, Crossref, Unpaywall and PLOS full-text corpus) studying different subjects with a focus on collaboration networks in the past 5 years, during my 2<sup>nd</sup> PhD and Postdoc in Berlin  $\Rightarrow$  *Academic mobility* and *High-Skilled Migration*
- I am working on an **Agent Based Model (ABM)** of micro motivations of scientists to choose collaborators and how the duality of performance oriented work with voluntary community work gives rise to a macro level hyper-competitive science system by evaluating different macro level funding schemes, performance evaluations and international collaboration policies  $\Rightarrow$  a methodology used by your group

## Why this research project?

- Collaboration with **Giovanni Barbato** from UNIMI
- Maybe other projects with higher technical complexity or focus on sociology would be more attractive!
- But this is an example of an *international* collaboration, complementing my skills, one I hope to repeat
- Because, while being from the same academic cohort, Giovanni is closer to the known ideal type of a social scientist
- He trusted me to handle the computational side
- I trusted him to handle the literature review, questions and theoretical framework
  - An extended version of the same literature I had reviewed before. But he proved me there is much for me to learn on how to better articulate previous works
- We wrote using Git/Gitlab, he learned to use git and Github desktop in 10 minutes!
  - I had to give reminders on: pull, edit and push (reasonable cost)
- I feel we arrived at an acceptable result in a short time, not perfect, in need of improvement, but still a nice achievement, IMHO

# Introduction: Object of Study

- Internationalization of Higher Education (HE) research, empirically analyzed through international co-authorship
- **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 as a multidisciplinary and emergent field of study characterized by a **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

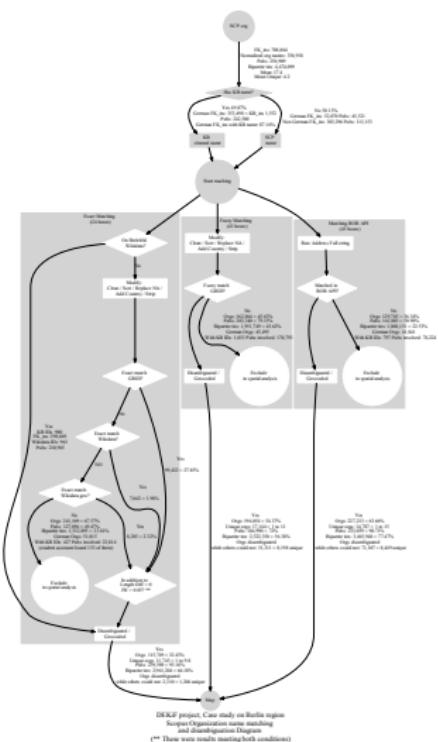
# Hypotheses

- **H1: Cumulative advantage (preferential attachment):** Scholars with higher previous collaboration have higher probability to be chosen for new collaborations.
- **H2: Difference in total publications' count:** Scholars with similar levels of *research productivity* have higher tendency to collaborate with each other.
- **H3: Difference in first publication year:** Scholars with similar academic status and seniority have higher tendency to collaborate with each other.
- **H4: Difference in last publication year:** Scholars with similar length of publication activity have higher tendency to collaborate with each other.
- **H5: Social closure among groups:** There are cohesive sub-groups in co-authorship network with higher tendency to collaborate *within* 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) 10<sup>th</sup> December 2019
- **Research Organization Registry (ROR)** [local] API 18<sup>th</sup> December 2019
- Scopus author numeric IDs

# Evaluation of Organization Name Disambiguation



# Bipartite Network Modeling, Why?

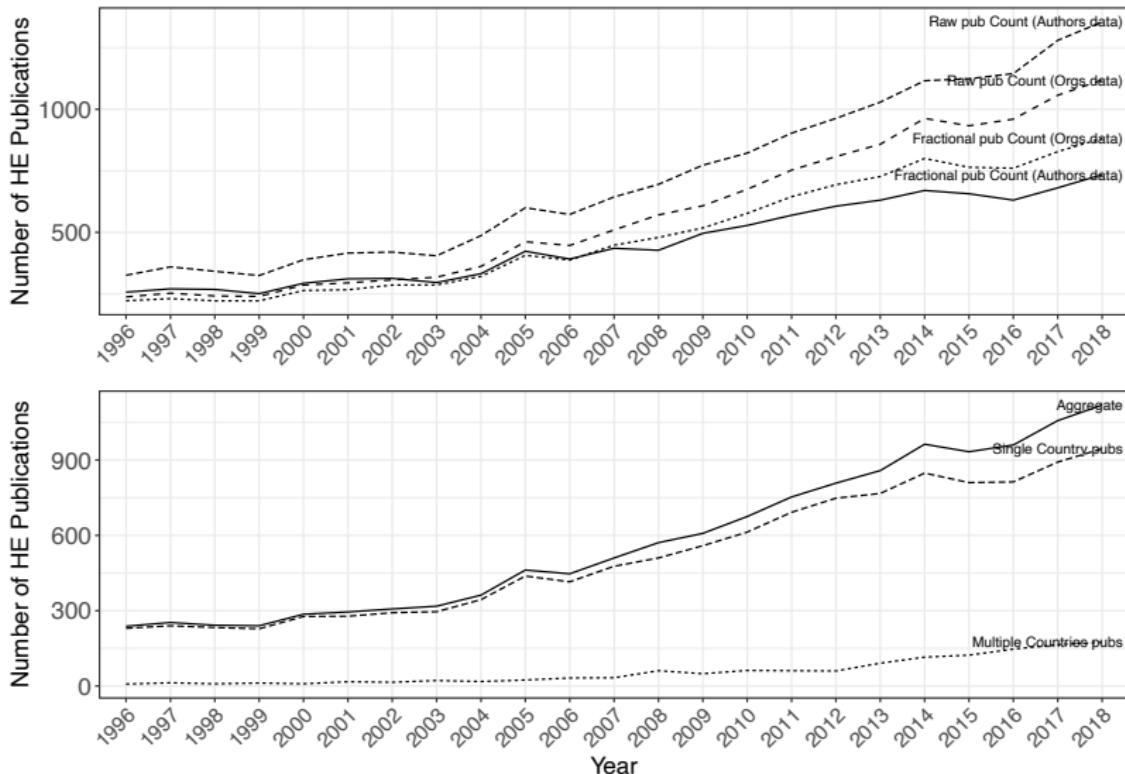
- Two node types: Organization/author and papers
- Two issues with *one-mode projection*
  1. Different structures yield same projected structure  $\Rightarrow$  information loss
  2. Articles with multiple authors  $\Rightarrow$  artificially high cliquish behavior
- Bipartite community detection with Constant Pots 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  $\gamma$  and the link density within communities (*internal density*) is higher than  $\gamma$
- $\gamma = 6 \times 10^{-5}$ , yield 36 communities with a uniform distribution 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)

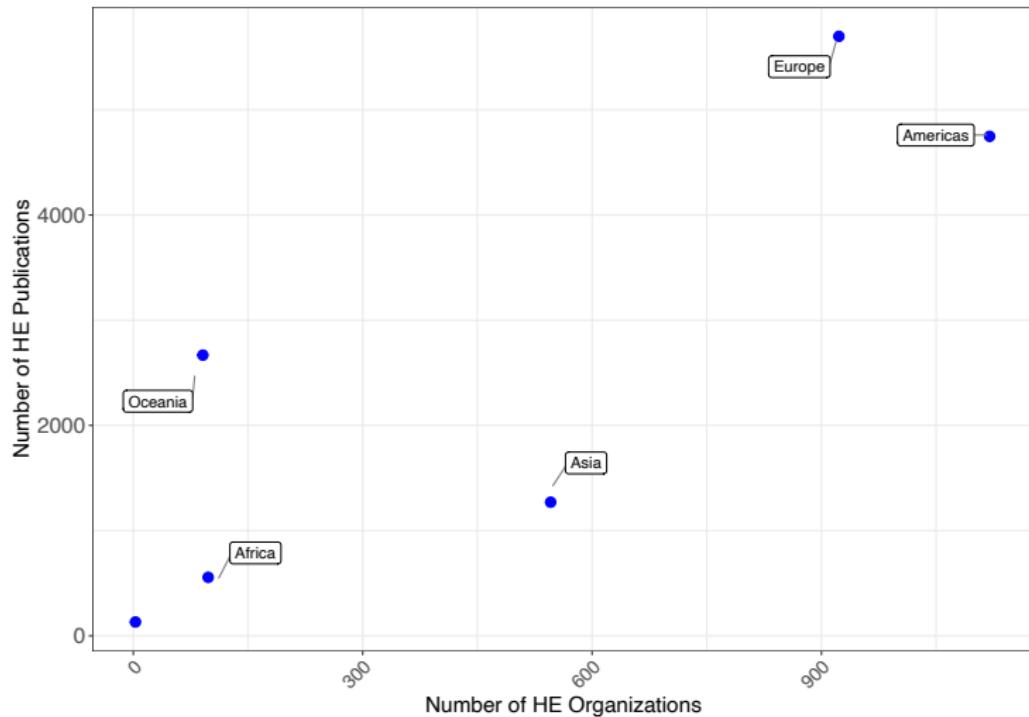
Metrics	Non disamb orgs	Disamb orgs	Non disamb authors	Disamb authors SID
Number of connected components	12,876	786	8,703	7,998
Number of bipartite nodes	43,122	16,663	55,546	41,997
Number of bipartite edges	30,549	19,098	53,024	37,382
% of bipartite nodes in G	11	88	19	23
% of bipartite edges in G	15	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?



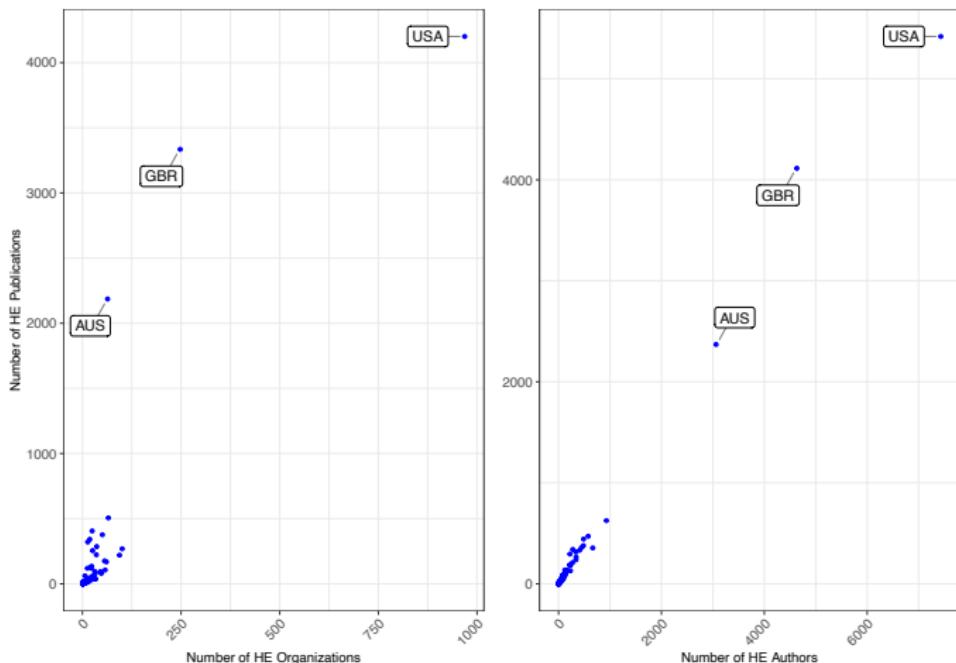
# HE organizations vs. HE publications (continents)

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



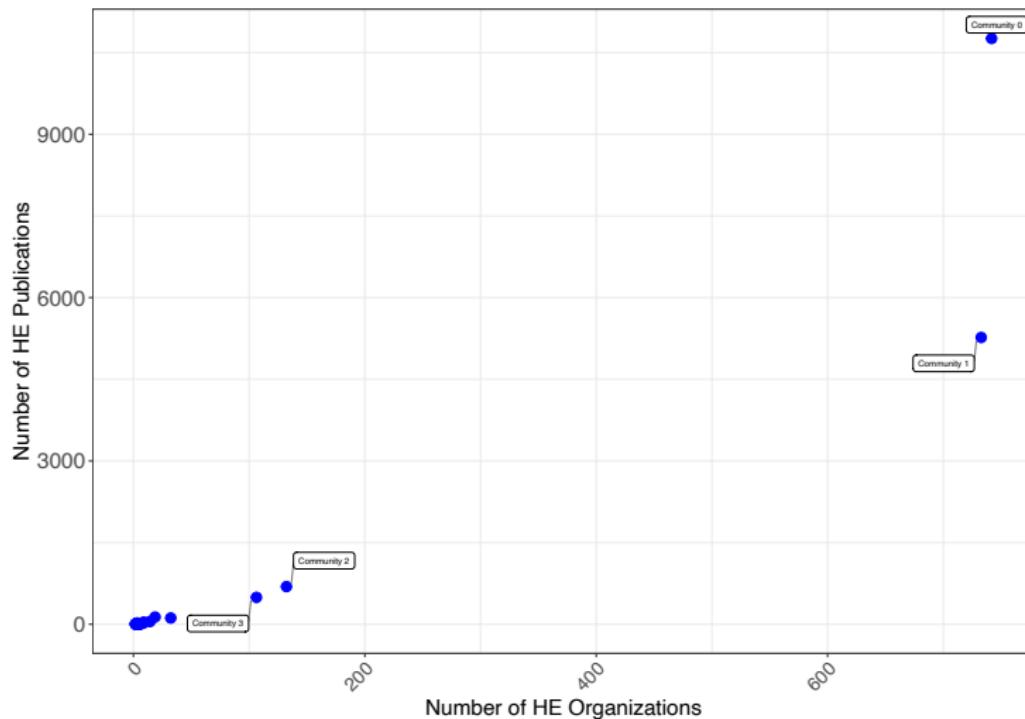
# HE organizations/authors vs. HE publications (countries)

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



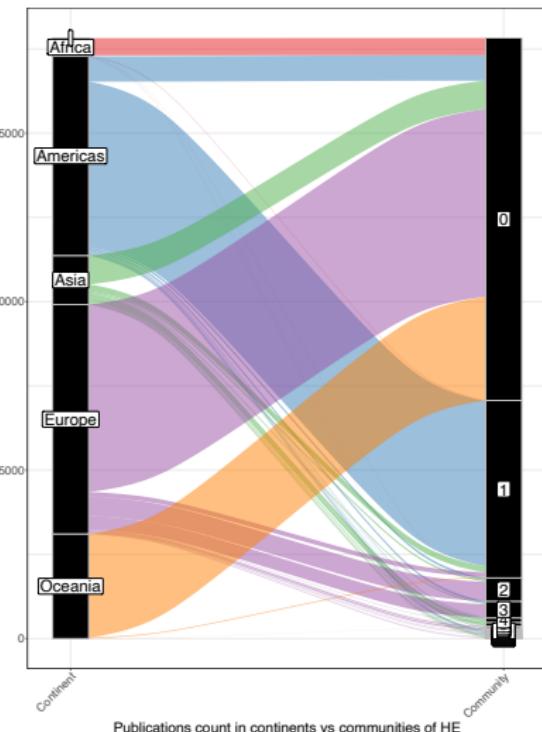
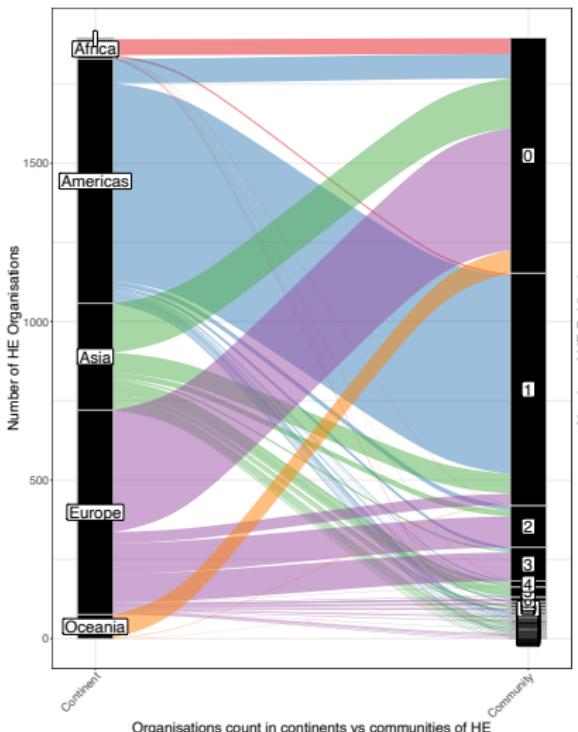
## Composition of HE communities (1/2)

- 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



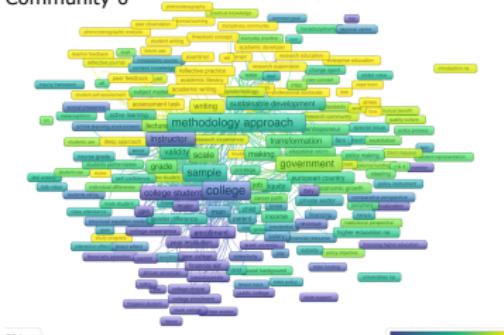
# HE communities' substantial focus

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

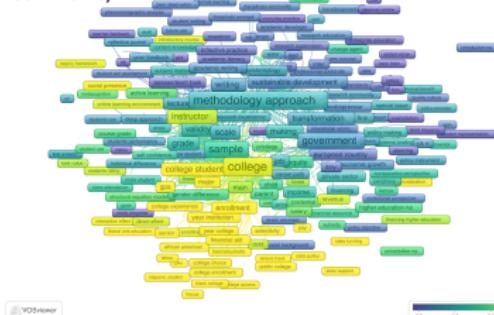
Terms Clusters



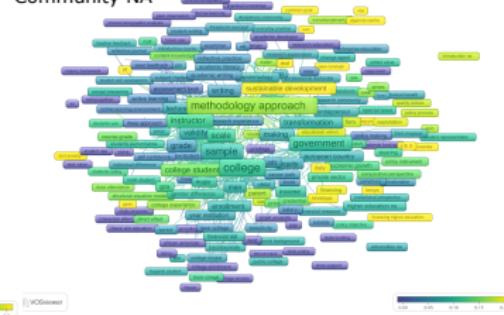
Community 0



Community 1



Community NA



# Bipartite ERGM Results

- H1: Tendency towards higher preferential attachment in *authors* and lower preferential attachment in *organizations*
- H2: Homophily and higher tendency of collaboration within most prolific organizations and authors (rich club)
- H3: Homophily in similar date of first publication (e.g., higher tendency of collaboration within 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 reoccur until recently

Table 4: Bipartite ERGMs results explaining effect of organization and author attributes and structural variables on co-authorship tie existence

	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)		2.996*** (0.003)	-5.913*** (0.081)		-5.917*** (0.079)
Difference in total pubs (H2)		-0.010** (0.004)	-0.014** (0.005)		-0.001 (0.001)	-0.002 (0.002)
Difference in first pub (H3)		-0.007** (0.003)	-0.010*** (0.004)		-0.004 (0.003)	-0.011** (0.005)
Difference in last pub (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 RQ:** *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 RQ:** *how network mechanisms contribute to explain evolution of the international HE co-authorship network?*
- Preferential attachment (H1): is an explanatory mechanism at the author level, while its effect is lower at organization level; confirms literature on other discipline (Wagner and Leydesdorff 2005; Zhang et al. 2018)
- There are minority of *key authors* while in the *organization* level there is not
- 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  $\implies$  atomization of the HE field (confirms the literature & methodological nationalism?)

## Discussion and Conclusions 3/3

- **THIRD RQ:** 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

## What, I have done vs. What, I hope to do

- In my first study, I was focused on informal, *friendship ties* formed online and *rational support* provided and received through them.
- In my second study, I was focused on *rational ties* formed through scientific collaboration and how it is affected by a social, *informal aspect*.
- I am hoping to continue this line of research and how social network tie formation and dissolution is a mixture of rational and irrational decisions and processes despite their context.
- Germany is the sixth European country and academic context in which I work.
- My high mobility enabled me to experience similarities and differences of these contexts.
- I am keen on studying *how migrant scientists do academic work* and *how the migrant background affects their academic work and integration in host country*.

## What, I have done vs. What, I hope to do

- Do more *solution-oriented social science* (Watts 2017). **How?**
- **Substantive:** By focusing on social network tie formation and dissolution process
  - It is a mixture of rational and irrational decisions despite the context
- **Empirical:** By focusing on [some] aspects of [*migrant*] scientists' academic life:
  - **Cultural**
  - **Social**
  - **Health**

# Cultural aspect of [*migrant*] scientists' academic life

- International scientific collaboration
  - Government policies *attacking* vs. *facilitating* academic mobility and collaboration
  - **Example:** Effect of US travel ban on scientific collaboration
- Migrant scientists' integration in host countries
  - **Example:** Does case of sociologists differ significantly? (Unity of subject/object)
  - **Example:** Language barriers for migrant scientist's [team] work and integration in collaborating groups
  - **Example:** Unwritten rules of power, collaboration and division of labor

# Social aspect of [*migrant*] scientists' academic life

- Duality of academic work
  - Selfless activity (e.g., peer review, editorial tasks) driven by unwritten rules of community vs. publish or perish and [quantitative] evaluation oriented action
  - **Example:** An ABM simulation of micro motivations and individual actions vs. top-down policies vs. middle range organizational embeddedness to reconstruct the science system's past and project its dynamics to future
- Evolution of sociology in different national/regional cases
  - **Example:** Tracing *contextual* effects on theoretical, methodological and epistemic aspects of the field
- **Conflict** in scientific collaboration
  - Science as a zero sum game
  - **Example:** Signed networks of scientific collaboration going beyond success stories of [published] academic work

# Health aspect of [*migrant*] scientists' academic life

- Personal networks' effect
  - **Example:** Effects on mental and physical well-being and quality of academic work
  - **Example:** Interaction of social and health aspects of doing science
- FOMO, Burnout and work-life balance
  - **Example:** Based on trends of social media use of academics
- Depression, stress and emotional deprivation during short research visits *abroad*

My background



Internationalization of HE



Future [Research] Plans



# Thank you!

- akbaritabar@gmail.com
- @akbaritabar
- akbaritabar.netlify.app

- 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.