Quantitative sociology of academic work in an era of hypercompetition and rankings

PhD Candidate:
Aliakbar Akbaritabar (Ali)

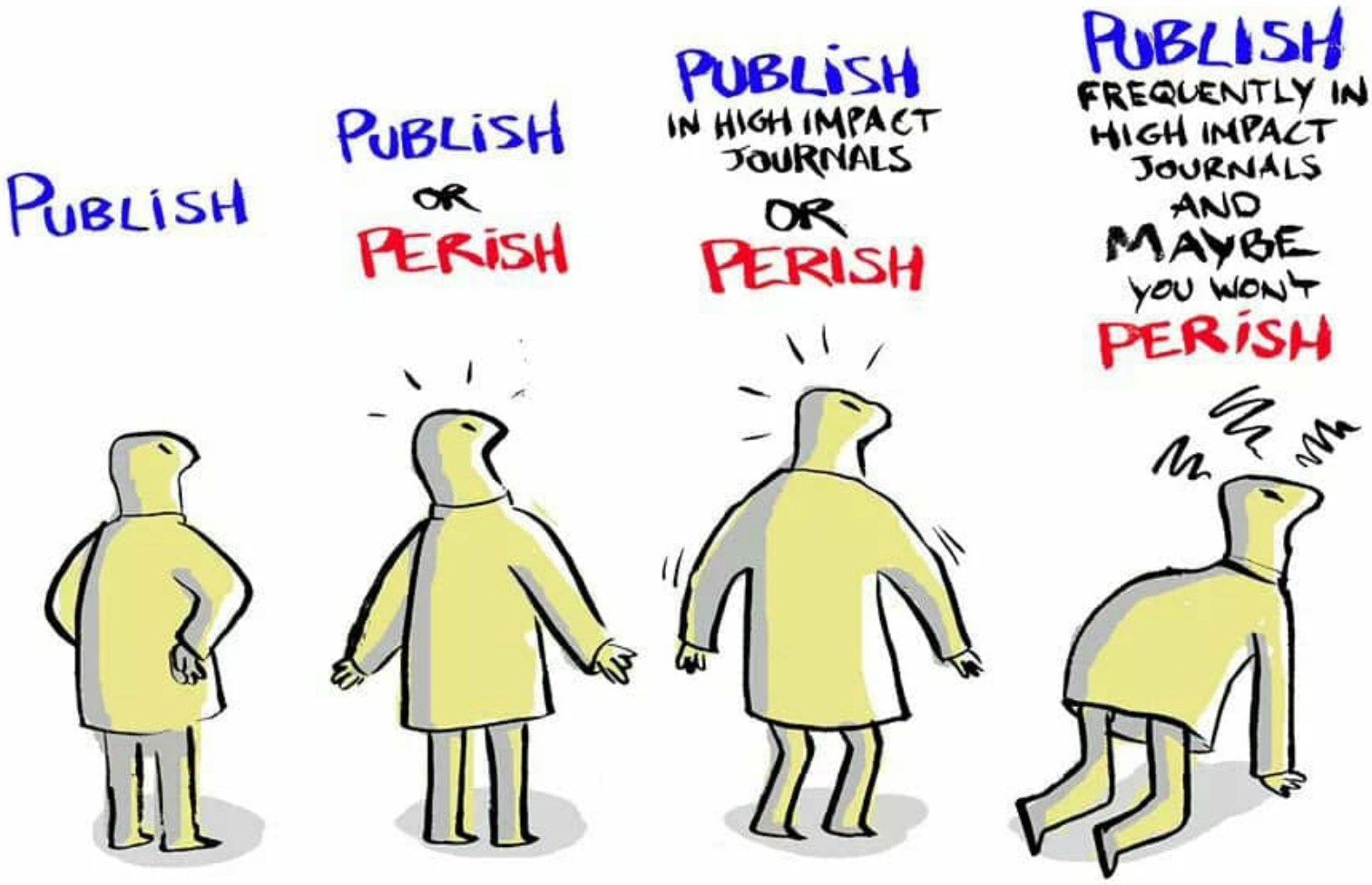
Supervisor:
Flaminio Squazzoni

UNIMI - UNIBS

16/4/2019

A hyper-competitive academic landscape, an introduction!

THE EVOLUTION OF ACADEMIA

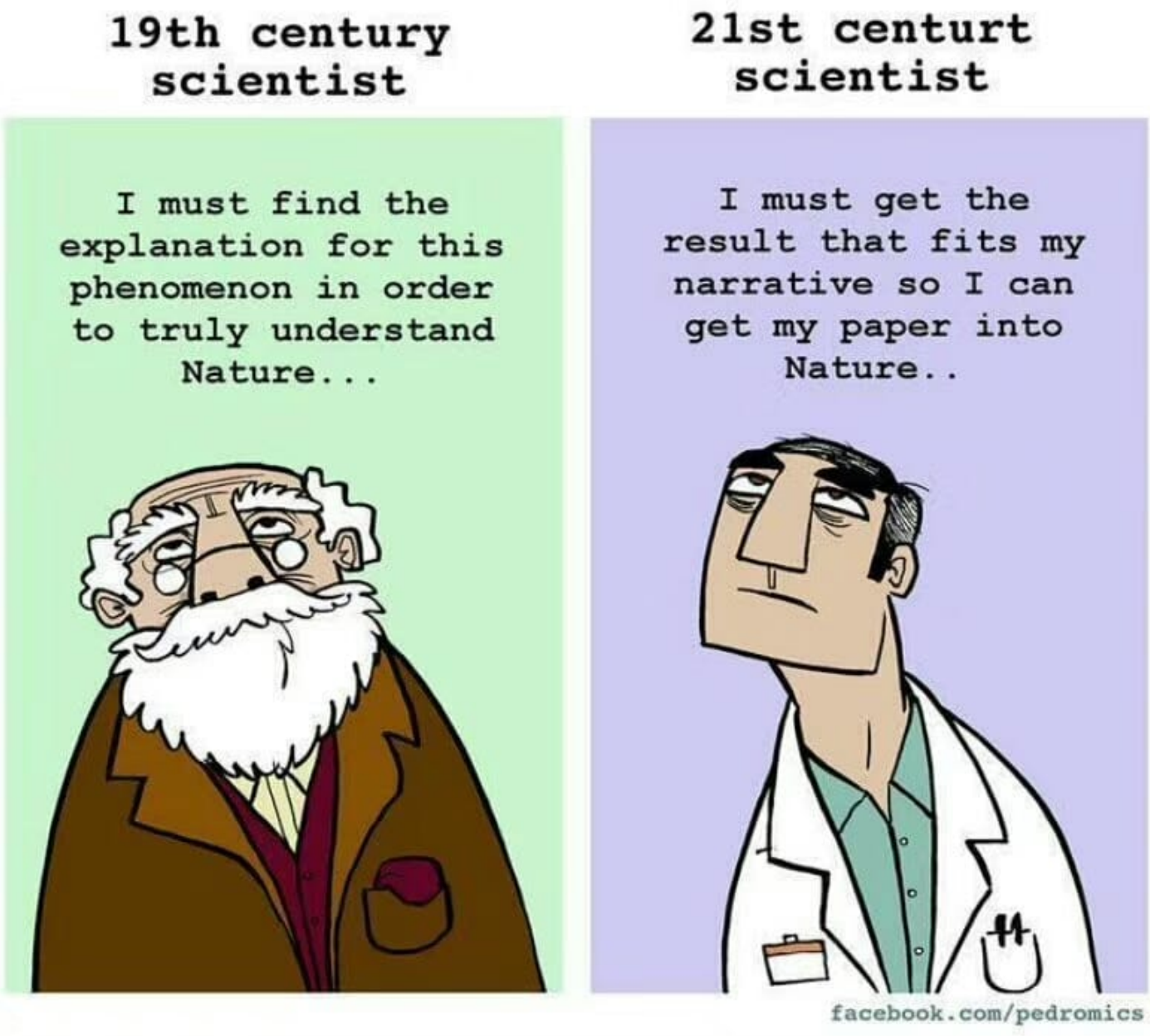


facebook.com/pedromics

A hyper-competitive academic landscape (Chapter 2)

- Quantity vs. quality and the hegemony of quantitative research evaluation
- Funding, promotion and reputation circulation in 21st century
- Responsible [use of] metrics
  - Peer review vs. bibliometrics, informed peer review and contextualized scientometrics

Goal displacement?!



Sociological theories

- **Matteo effect**, winner takes all?
  - Highly prolific scientists attract higher collaborations from other scientists
  - Attaching preferably to a few **star scientists**/leaders?
- **Fragmentation** of ideas, sociology as a *interstitial science*!
  - Methodologists bridging the islands?
- Sociological **small world** of disconnected islands?
- Embeddedness and **organizational ambiguity** at work?
- **Core** of leaders and **periphery** of followers?
- **Sum up**: Some farther away, some closer to our quantitative focus

Brief description of data

- **Two** data sets, one **national**, one **international**
  - All Italian sociologists (chapters 3, 4 and 6)
    - 1,029 professors & 198 postdocs
    - 3,168 papers, 1973-2016
    - 55% Male, 45% Female
    - Only 64% of 1,227 had at least 1 article
  - Whoever published in AJS & ASR (chapter 5)
    - 4,709 authors
    - 2,593 papers, 1946-2016
    - ASA members: 47% Male, 53% Female
    - AJS & ASR: 70% Male, 30% Female
    - 40% of papers, at least one female author(s)
    - 84% of papers, at least one male author(s)
    - 80% solo or team of Americans

Brief on methods

- Computational social science at work (come to my talk in IC2S2)
  - Cross-fertilization between *computational sociology* & *science studies*
  - API calls & web scrapping for data gathering
  - Hierarchical linear models (nested & mixed membership)
  - Repeated measurement models
  - Text analysis, structural topic models, VOS algorithm
  - Network analysis, community detection and ERGM
  - Code, data, analysis and report are all reproducible (hosted on Github)

Why sociologists? Why Italy?

- A community that is between **humanities** scholars, who are predominantly qualitative, anti-bibliometric and "**hard**" scientists, who are quantitative
- The **co-existence** of different epistemic communities makes sociologists an interesting case to examine contrasting forces towards internationalization
- Research on scientists' productivity in Italy have looked mainly at the case of *hard sciences* stating that research in humanities and social sciences cannot be examined quantitatively

A brief look at 4 case studies in chapters 3, 4, 5 and 6

Individual level research productivity (Chapter 3)

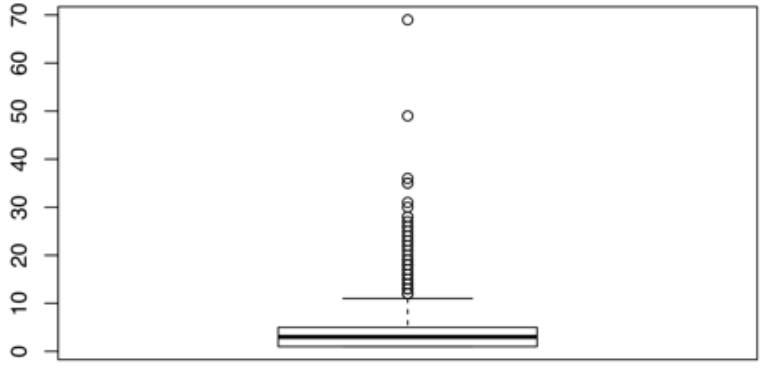
- **Looked at**:
  - Internationalization
  - Co-authors similarity
  - Individuals embedded in organizational settings (*University*, *department* and *sectors*)
- **Found that**:
  - **Male** scientists
  - Those working more **internationally**
  - Those working with a *similar* group of coauthors *are more productive*
  - But **not** necessarily more cited by other members of the community
- On Scientometrics, DOI=10.1007/s11192-017-2606-5

A glimpse of Chapter results

Table 3.2: Comparative table of multilevel regression models

	log(FSS)	Total Publications
Constant	−4.03 (0.22)***	0.02 (0.01)
Internationalisation	0.91 (0.34)**	0.10 (0.02)***
Coauthors Stability	0.12 (0.69)	0.46 (0.05)***
Gender (male)	0.39 (0.14)**	0.03 (0.01)**
Associate professor	−0.27 (0.16)	0.02 (0.01)
Full professor	−0.38 (0.19)*	0.03 (0.01)*
Postdoc	0.99 (0.43)*	0.10 (0.02)***
AIC	1652.31	-702.13
BIC	1697.66	-656.38
Log Likelihood	-815.15	362.07
Num. obs.	456	473
Num. groups: university	60	61
Num. groups: sector	7	7
Num. groups: department	7	7
Var: university (Intercept)	0.29	0.00
Var: sector (Intercept)	0.11	0.00
Var: department (Intercept)	0.00	0.00
Var: Residual	1.89	0.01

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$



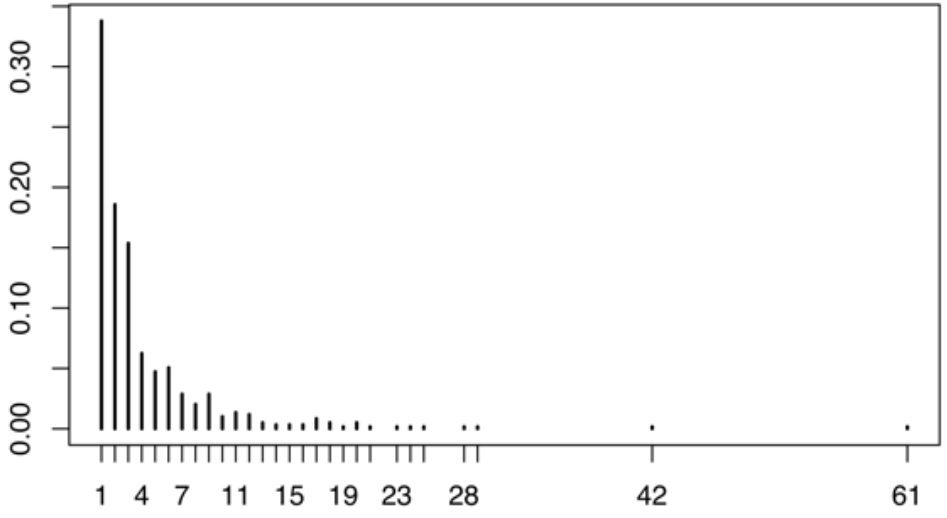
Top-down process of inspiring research productivity (Chapter 4)

- **Looked at**:
  - ANVUR, VQR 2004-2010 effect on research productivity
  - Five full years **before** and **after** assessment 2006-2015
- **Found that**:
  - ANVUR had a **limited** influence on research productivity
  - **No clear pattern** of increase or decrease inspired by top-down process
  - Most differences were due to individual characteristics

A glimpse of Chapter results

	Model 1	Model 2
Constant	0.13 (0.02)***	0.05 (0.01)***
Pub in Fascia journals post-ANVUR	−0.09 (0.02)***	
Pub in non-Fascia journals post-ANVUR		0.01 (0.01)
AIC	-236.60	-1277.24
BIC	-215.84	-1250.24
Log Likelihood	124.30	644.62
Num. obs.	235	665
Num. groups: id	214	508
Num. groups: university	49	64
Num. groups: department	7	7
Var: id (Intercept)	0.01	0.00
Var: university (Intercept)	0.00	0.00
Var: department (Intercept)	0.00	0.00
Var: Residual	0.01	0.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$



Diversity in research productivity (Chapter 5)

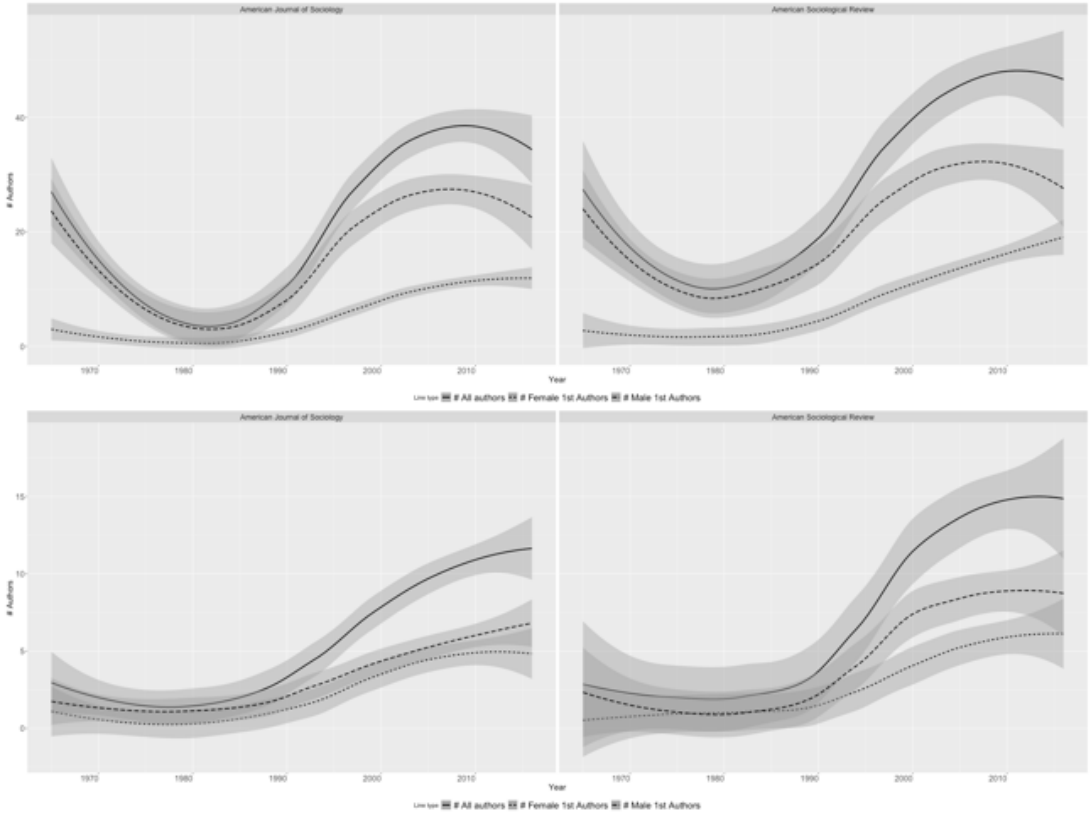
- **Looked at**:
  - **Gender** diversity in research productivity & impact
  - *Ivy-League* effect in research productivity & impact
- **Found that**:
  - These prestigious journals especially considered solo/teams of **male** authors (60% papers)
  - These gender penalties persist even when looking at *citations* and after controlling *affiliation* (both PhD and current affiliation)
  - The "**Ivy-League**" effect greatly benefits *only male* authors
- On ST&HV, DOI=in-press

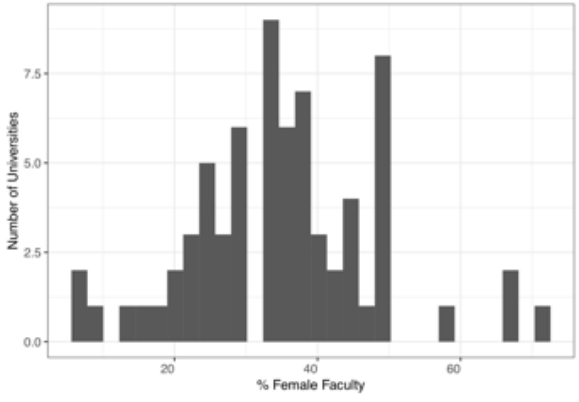
A glimpse of Chapter results

Table 5.11: Multilevel regression models on star sociologists

	Total Publications	Publications before 2000	Publications after 2000	Total Citations
Constant	0.12 (0.03)***	0.04 (0.05)	0.12 (0.04)***	4.20 (0.07)***
Gender Male	0.09 (0.03)**	0.10 (0.05)	0.06 (0.04)	0.04 (0.06)
Star sociologist	1.16 (0.03)***	0.79 (0.05)***	0.91 (0.04)***	1.35 (0.07)***
AIC	6610.52	2751.34	4310.07	28380.56
BIC	6645.37	2781.28	4342.54	28415.42
Log Likelihood	-3299.26	-1369.67	-2149.03	-14184.28
Num. obs.	2463	1086	1655	2463
Num. groups: latest_uni	444	256	336	444
Num. groups: phd_awarded_university	329	195	250	329
Var: latest_uni (Intercept)	0.00	0.00	0.00	0.16
Var: phd_awarded_university (Intercept)	0.00	0.00	0.00	0.05

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

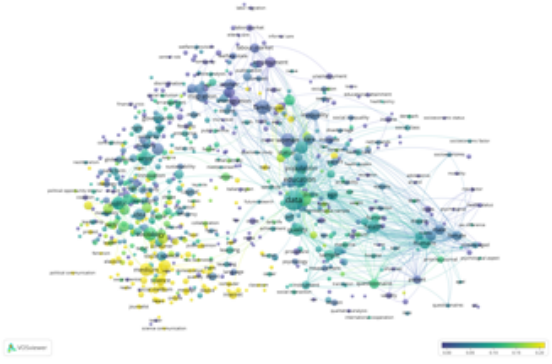


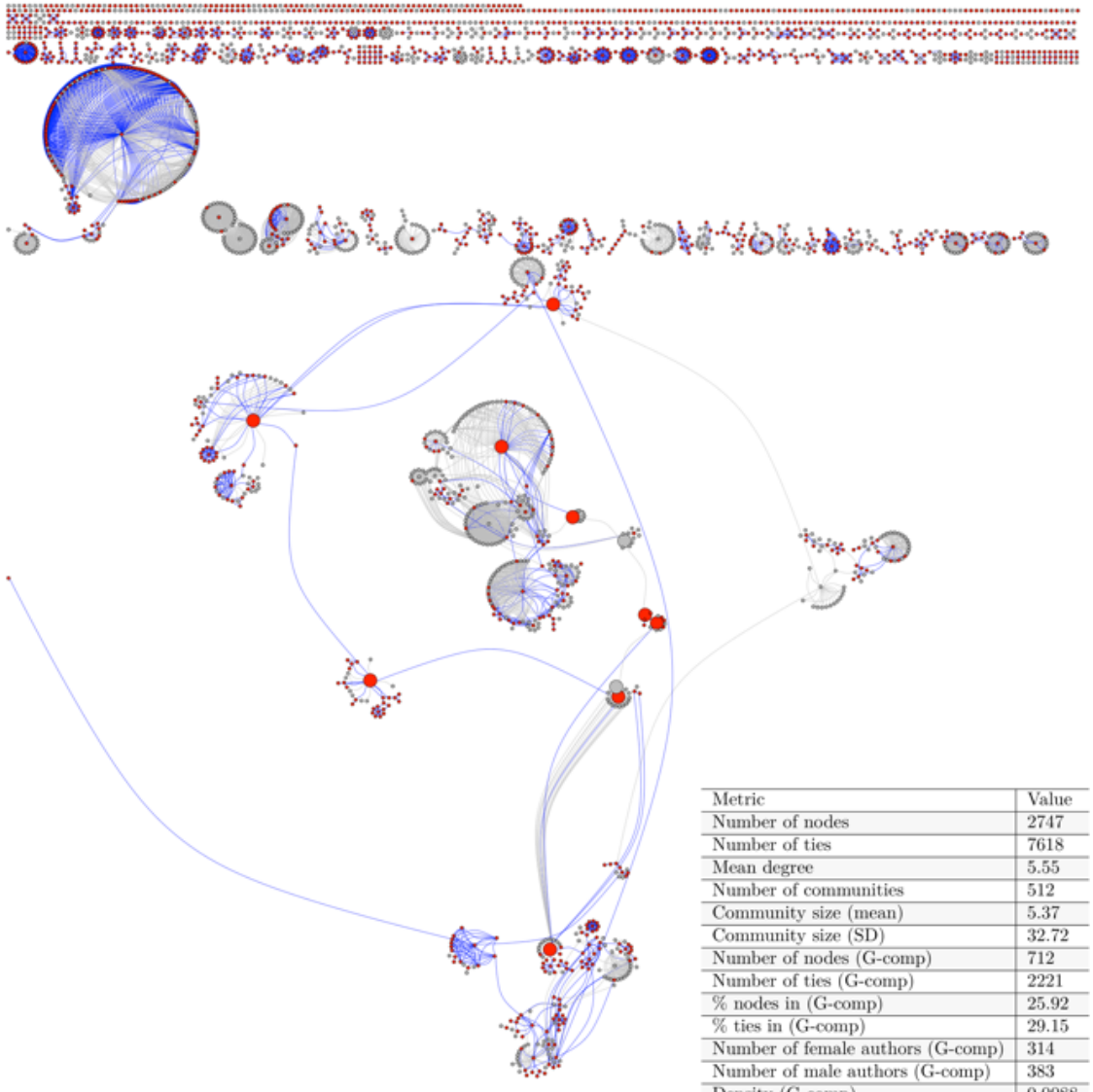


Networks effect in research productivity (Chapter 6)

- **Looked at**:
  - *Community* membership & evolution
  - *Substantive* similarity & research communities
  - ERGM with *homophily*, structural and community effects
- **Found that**:
  - High **disconnectedness** vs. mathematical simulations (512 Comp., 25.92% (29.15%) nodes (ties) in G-comp)
  - Relatively **high** rate of change in communities (even G-comp) members, compared to all sociology & other cases
  - **Male** *newcomers* are more likely to stay, joine the core and continue in academia
  - Two *largest* and *most stable* research communities in Italian sociology are **economic** and **political** sociologists
  - Collaboration ties were mainly driven by the *research focus*
  - Other factors, such as *preferential attachment*, *gender* and *affiliation* homophily were also important
  - Political sociologists tend to be *more international*

A glimpse of Chapter results 1/2





Metric	Value
Number of nodes	2747
Number of ties	7618
Mean degree	5.55
Number of communities	512
Community size (mean)	5.37
Community size (SD)	32.72
Number of nodes (G-comp)	712
Number of ties (G-comp)	2221
% nodes in (G-comp)	25.92
% ties in (G-comp)	29.15
Number of female authors (G-comp)	314
Number of male authors (G-comp)	383
Density (G-comp)	0.0088
Diameter (G-comp)	32

A glimpse of Chapter results 2/2

Table 6.8: ERGMs results explaining effect of author attributes and structural variables on coauthorship tie existence

	The Giant component of Italian sociologists and their coauthors			
	ERGM Models			
	(1)	(2)	(3)	(4)
Ties	-4.549^{***} (0.022)	-4.354^{***} (0.050)	-11.348^{***} (0.997)	-11.254^{***} (1.012)
Preferential attachment	15.095^{***} (4.421)			4.063^{***} (0.976)
Homophily Females		0.002 (0.059)	0.049 (0.060)	0.067 (0.060)
Homophily Males		0.281^{***} (0.049)	0.267^{***} (0.049)	0.258^{***} (0.054)
Community 0			7.207^{***} (0.996)	7.136^{***} (1.015)
Community 1			8.318^{***} (0.996)	8.216^{***} (1.015)
Europe		1.226^{***} (0.051)	1.099^{***} (0.053)	1.113^{***} (0.059)
Italy		0.600^{***} (0.057)	0.718^{***} (0.059)	0.716^{***} (0.067)
Other countries		1.546^{***} (0.170)	1.188^{***} (0.177)	1.170^{***} (0.216)
Homophily Total Pubs		0.060^{***} (0.002)	0.063^{***} (0.002)	0.063^{***} (0.003)
Homophily First Pub		-0.092^{***} (0.006)	-0.095^{***} (0.006)	-0.095^{***} (0.007)
Homophily Last Pub		-0.372^{***} (0.014)	-0.367^{***} (0.014)	-0.369^{***} (0.016)
Akaike Inf. Crit.	25,237.220	22,732.670	19,277.840	19,154.370
Bayesian Inf. Crit.	25,268.540	22,826.640	19,392.700	19,290.110

Note: *p<0.1; **p<0.05; ***p<0.01

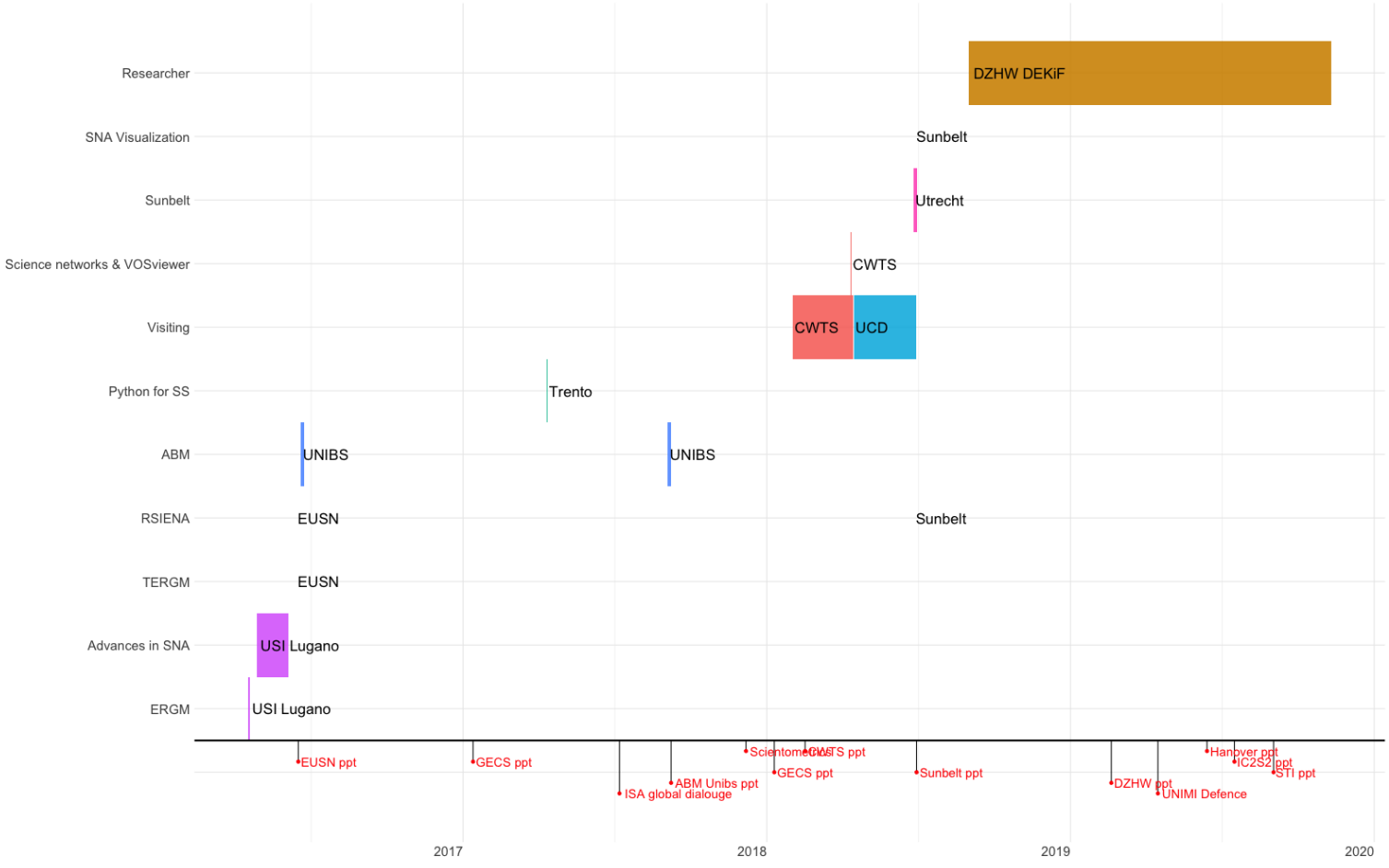
Discussion and conclusions (Chapter 7)

- We aimed to provide a quantitative look at academic work in 21st century
- Academics today are embedded in a dual context as if they were living a double life
- We were focused on the tension between "**publish or perish**" from one hand and **being part of the "scholarly community"** on the other
- In different chapters, we studied a variety of **embeddedness scenarios** to see how sociologists reacted to this hyper-competitive academic landscape
- We found that the **quantitative evaluation** mantra in sociology is **not** inspiring a clear stream of behavior, mixture of ambigous signals with disconnect from reward system causes sociologists to continue in an unknown state
- They are *not* yet as close to hard sciences to be competition driven, thus causing them to present humanities like behaviors (e.g. lack of care for citations)
- There are many criticisms of **quantitative evaluation** while not much alternatives are introduced
- I feel it is time to *give voice* to those under evaluation to tell us why they don't comply with the currently introduced motivations

Limitations

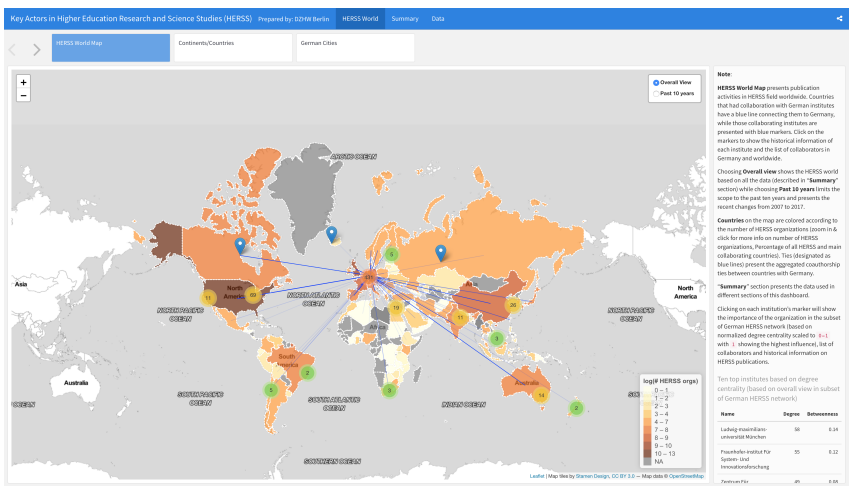
- Success stories only!
- Positive collaboration only!
- Scopus indexed publications only!
- One mode projection of bipartite *paper-author* ties which causes artificially high cliquish behavior and could be better studied with bipartite modelling frameworks
- We couldn't account for motivations (more on this in next slides)

Activities NOV-2015:NOV-2019



Next steps. Work, still in progress!

- I joined DZHW (German Center for Higher Education Research and Science Studies) on September 2018 for DEKiF (3 years) project
- **DEKiF**: Determinanten und Effekte von Kooperation in homogenen und heterogenen Forschungsverbunden; Determinants and effects of cooperation in homogeneous and heterogeneous scientific networks
- I am involved in 3 mixed methods (quantitative, qualitative and bibliometric) case studies of scientific collaborations (example pilot study below)
- I hope to resolve some of the limitations that exist in this dissertation
- I hope to look at *decision making process* of scientists before any scientific collaboration has taken place, to see **why they do what they do**!



Thanks a lot for your attention

Email: akbaritabar@DZHW.eu

Twitter: @akbaritabar