Speech notes on each slide

Title slide

- Hello everyone!
- Thank you for coming.
- I am Ali.
- This is a research by academics, on academics, for a target audience, who are also academics.
- I hope you will enjoy the talk. *(next slide)*

A hyper-competitive academic landscape, an introduction!

- My thesis is a puzzle composed of multiple pieces.
- All pieces revolve around the general topic of "how is academic work today"
- In the age of hyper-competition
- [**pause**]
- What is **hyper-competition**? *(next slide)*

Image slide (Publish or perish)

- There was a recent article published in Nature on "hyper-prolific scientists".
- The scientists who publish a paper **every five days**.
- That is a rare case, though. *(next slide)*

A hyper-competitive academic landscape (Chapter 2)

- Trends are toward larger scientific teams, higher number of publications, higher citations, higher number of students and faculty members
- Salami slicing publications, building invisible colleges of thoughts and citation cartels are observed.
- This in turn leads to higher levels of stress, *burnout* and the *fear of missing out* on newest scientific outcomes.
- This type of *over-competitiveness* increasingly demotivates scientists.
- Vicious circle of funding on **quantitative basis**, inspiring more quantity over quality and discouraging breakthrough and innovative scientific works are the sad outcomes recorded for the current trends.
- With increasing move toward quantification, there are initiatives to attract attention toward *responsible [use of] metrics*. Research on comparison between *peer review*, *bibliometrics* and more recently, *contextualized bibliometrics* are examples of academics trying to question, if academia is headed toward the wrong place. *(next slide)*

Image slide (Goal displacement?!)

- Some observers are talking about a *goal displacement* that has happened.
- Research evaluation has become *a goal in itself.*
- 21st century scientists do **publishable, AKA, evaluable research**.
- As they call it **REF worthy research** in the UK.
- Examples of effort to game the system in terms of publication fraud, CV manipulation and self-citations have been recorded *(next slide)*

Sociological theories

- There is a general divide in "**Science and Technology Studies**", a group of scientometricians, bibliometricians and information scientists versus those doing sociology of science.
- Where the former have been more successful to institutionalize their efforts in form of institutions, centers and departments. e.g. CWTS
- In sociology of science, two main streams of theoretical efforts can be differentiated.
- **One**, quantitative, sometimes with macro approach, with streams of *structural functionalism* in American sociology, inherited from Merton and his students' works.
- **Second**, more qualitative approach by Sociology of Scientific Knowledge (SSK) and Actor Network Theory (ANT), which is farther from our *quantitative sociological* approach here.
- **Matteo effect** has been measured mainly with *preferential attachment* in network analysis terms
- **Fragmentation of ideas** have been looked at with *small-world* and *transitivity* in network terms plus by evaluating content of scientific publications
- **Organizational Embeddedness** has been evaluated by *homophily effects* and including different affiliation variables
- **Core and periphery** structures have been extensively studied by *community detection*, *blockmodelling* and *hub and authority* in network analysis terms which all stem from *social cohesion* idea from Durkheim. *(next slide)*

Brief description of data

- We gathered, cleaned, reshaped and used 2 sets of data, one national, and one international.
  - In the national case, we looked at all Italian sociologists and all of their publications indexed on Scopus
  - We extracted their list from MIUR website, which includes demographic information on academic level, gender, current affiliation and the like.
  - It included 1,029 professors & 198 post-docs.
  - 55% of them are Male, and they have published 3,168 papers, from 1973 to 2016
  - Only 64% of 1,227 Italian sociologists had at least 1 article indexed in Scopus
- **our international dataset** includes all those who published in American Journal of Sociology and American Sociological Review
  - Including 4,709 authors who published 2,593 papers, from 1946 till 2016
  - While women are the majority of members in American Sociological Association, equal to 53% in AJS & ASR authors, we observed a 70 to 30 percent ratio
  - While 40% of papers had at least one or more female authors, the case for male authors was 84% (so double)
  - 80% of the articles were written by single American authors or teams of American authors *(next slide)*

Brief on methods

- This was a research in **computational social science**.
- We walked on the border of *computational sociology* and *science studies*
- From data gathering, cleaning to analysis, we used state of the art techniques from machine learning in gender and ethnicity imputations, to web scrapping to API calls to complement our traditional sources of data
- To answer some of our questions, I had to learn and apply some statistical models like *Hierarchical linear models* (with nested and mixed membership in random effects) to *Repeated measurement models*
- We looked into content of articles' abstracts, titles and keywords, by text mining methods, structural topic models and VOS algorithm

- Of course, **doing science is a social practice**, so we needed to take the relations into account, we modeled the coauthorship ties as an undirected network
- Community detection and exponential random graph models from network analysis field helped us in this aspect
- All scripts in each of these stages are maintained on Github and could be shared with you for replication purposes *(next slide)*

Why sociologists? Why Italy?

- Sociology has multiple, diverse sub-communities.
- They range from *quantitative* to *qualitative*
- While they have different research focuses and they might use completely different methods and techniques, interestingly, they sometimes use similar sociological concepts to interpret their results, example of social capital, in purely quantitative studies, compared to qualitative, in-depth studies, both use similar concepts, while their research belong to two different worlds
- The co-existence of these diverse communities in a field is highly interesting
- But, this community, its research productivity and collaboration network, has been understudied in both Italian and international cases in comparison to economics and other neighboring fields of science *(next slide)*

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

- Of course, it is *close to impossible* to go over all details of each case study in 30 minutes, so I will give you a glimpse of each *(next slide)*

Individual level research productivity (Chapter 3)

- We looked at correlates of *Italian sociologists' research productivity* in **individual level**
- We took factors like *how international they work*, the group of coauthors each scientist has, while we embedded each individual into his/her *university*, *department*, and *sector*
- At first this embeddedness followed a nested structure, on suggestion of one of our referees, we tried with mixed membership structure, which didn't change the main trend of results
- We found that male scientists, those working more internationally, and those working with a similar group of coauthors had higher rates in publications, while not necessarily in citation and impact *(next slide)*

A glimpse of Chapter results

- A heavily long-tailed distribution of number of publications with the median close to 1 article was observed
- Sector accounted in moderate level for inter-group variations (percentages below the table), while it was not significant once included as a *fixed effect*. *(next slide)*

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

- After our probe into individual research productivity and its correlates
- We then asked, if a top-down process of inspiration is at work to motivate Italian sociologists to publish more

- After first exercises at research evaluation in Italy, like VTR, from 2010, ANVUR started to carry out VQR, which is more structured and closer to international examples like REF in the UK
- We **could not** find a clear pattern of influence from ANVUR and more specifically VQR 2004-2010 on publication behavior or Italian sociologists
- Some of our observations were counter-intuitive. *(next slide)*

A glimpse of Chapter results

- We observed a similar long-tailed distribution of publications on 2006-2015
- Counter-intuitive to what we expected, there was a decreasing post-ANVUR trend of publications in *fascia-A* journals, which have higher impact in evaluation.
- Articles in *non-Fascia-A* journals were increased post-ANVUR, but was not significant.

Diversity in research productivity (Chapter 5)

- Sociology discusses inequality and diversity to a high extent
- Observing these discussions in sociological literature, gives an idea that the whole field would pay a higher attention to composition of its members and to the issue of gender and gender equality among its members
- Well, our observations proved otherwise
- As briefly mentioned in data description, only 40% of papers in two top sociology journals had at least one female author(s), while this share is 84% for male authors
- 60% of articles were written by solo males or teams of male authors
- We controlled for **Ivy-League** effect. Females didn't have a better situation in those institutes either
- Our first looks at content of research showed that male and female scientists tend to have different research focuses. *(next slide)*

A glimpse of Chapter results

- A referee asked us if the gender difference in first authorship in AJS and ASR would be different in **cross gendered authored articles**, it was not.
- At the top, the aggregate trend of all publications, and at the bottom the specific trend of cross-gender coauthored articles is presented.
- The plot on the right shows the % of female, sociology faculty members, in top 100 universities based on Shanghai ranking. Majority of these universities have less than 50% females faculties.
- A few *male star scientists* dominate the field
- When we added them to our models, other factors became less or non-significant *(next slide)*

Networks effect in research productivity (Chapter 6)

- In limitations section of each of the previous case studies, we were pointing to the fact that research and *academic work* doesn't happen in a social and organizational vacuum
- We were sure on the need to take network of relations, groups, communities, and community evolution into account while looking at *how academics do science*.
- So, in chapter 6, we looked at coauthorship networks of Italian sociologists.
- We tried to find answers to the questions raised in *sociological theories* reviewed briefly.
- We took similarity in substantive focus of research into account.
- We used sophisticated temporal community detection model to find cohesive research communities, we then explored their evolution over the time and asked *what brings these communities together and detaches them from other parts of the network?*

- In our ERGM, we took individual attributes, homophily, structural, covariate, and community effects into account.
- We found:
  - A 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 continue in academia than females
  - 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 *(next slide)*

A glimpse of Chapter results 1/2

- You can observe the 512 disconnected components on the graph on top right.
- Colors: Affiliated to Italy = Red, Affiliated elsewhere = Gray, Ties within Italians = Blue, Other ties = Gray, Node size = Betweenness centrality
- We observed that Italian sociologists were collaborating with inter-mediation of other scientists affiliated to a diverse set of institutes which is further explored in our homophily effects
- Three plots on left column shows the substantive focus of the members of community one, i.e. *economic sociologists*, community 2, i.e. *political sociologists* and those who were not member of the giant component. In each case, yellow part of the graph shows the main substantive focus. *(next slide)*

A glimpse of Chapter results 2/2

- We built four ERGMs.
- One exploring only *Matteo effect* with preferential attachment and tie existence which is the usual base line in ERGM and usually is negative (showing lower tie existence compared to what can be expected from random networks of the same size)
- One with only node attributes (**gender, affiliation country, academic senirity and number of publication homophily**) *excluding* the community membership (that defines the research focus)
- Next one included these communities
- And the final one, with all of these effects showing that only homophily between female sociologists was not significant *(next slide)*

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
- On the one hand, they have to keep up with quantitative evaluation standards.
- They have to publish as much as possible under the publish or perish imperative.
- On the other hand, they are evaluated by their disciplinary community and their peers *not merely quantitatively*
- In different chapters, we studied a variety of **embeddedness scenarios** to see how sociologists reacted to this hyper-competitive academic landscape
- We found that Italian sociology same as international sociology is rather male dominated and it favors male new comers to join and stay in the core of the community
- At the same time, it is driven by substantive focus of research communities and in contrast to some neighboring fields studied in the literature, sociology in both our cases presents a high level of divide (even the two sub-communities of Italian case were only connected by one tie!).

- In both cases we observed a stable core that is growing slowly and takes new members selectively and some groups have much lower chances to join and stay in the publishing/collaborating core.
- It doesn't present the so-called *big science* and *interdisciplinary* sciences' qualities. We feel it can be further probed if we look at the corpus of all sociological literature which we have extracted, but couldn't finish to include in the thesis.
- We found that the **quantitative evaluation** mantra in sociology is **not** inspiring a clear stream of behavior, mixture of ambiguous 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 *(next slide)*

Limitations

- Our research suffered from many limitations, as any research would do, I have some hopes to resolve them, which I will discuss in two slides later.
- We did not have data on submissions, editorial decisions, negative or unsuccessful collaborations, and our data suffers from Scopus's coverage which is the highest among available sources, but still biased toward English language literature while we checked whether it includes main Italian journals, which it does.
- One mode projection of two-mode *author-paper* ties causes artificially high cliquish behavior, we tried to cope with that by using temporal slices of the network over years (since we observed that repeated coauthoring pairs and groups were rare to happen inside a single year) but still there are models which operate on bipartite ties that might give better insights.
- We didn't have data on motivations of scientists to collaborate or not, I will discuss this a bit in next. *(next slide)*

Activities NOV-2015:NOV-2019

- It was a rather busy 3 years for me. Apparently the next three years are going to be busier!
- I tried to participate and learn from different courses, schools and visiting periods.
- Presenting results in academic events, to samples of our target audience, helped me hear interesting feedback, same happened with journal referees and their kind suggestions
- I am grateful to the academic world for that.
- Last year today I arrived to Dublin for my research visit and on April 17th I met Diane Payne for the first time in person, while we had Skype discussions before. It was really sad and shocking for me to receive an email two days ago and hear that she has passed away. This is a great loss for *computational social science community*. *(next slide)*

Next steps. Work, still in progress!

- This might not be the tradition in the final discussion of a PhD thesis, but, I have to say, this work is not yet finished!
- It is a journey I am currently continuing during my Post-doc in *German Center for Higher Education Research and Science Studies* in Berlin
- I am hopeful to tackle some of the main limitations we faced during our quantitative study with mixing a qualitative and quantitative approach in DEKiF project.
- It is a 3 years project looking at *Determinants and effects of cooperation in homogeneous and heterogeneous scientific networks* *(next slide)*

Final slide

- Thank you all for your attention