

Publication patterns and scientific productivity in sociology in an era of academic hypercompetition

Supervisor: Flaminio Squazzoni . Aliakbar Akbaritabar

June 29th, 2017

Outline of presentation 1/2

- ▶ **First paper:** “The conundrum of research productivity. A study on sociologists in Italy”, by: Aliakbar Akbaritabar; Niccolo Casnici; Flaminio Squazzoni (done, under review)
 - ▶ Brief introduction
 - ▶ Method & Data
 - ▶ Results
 - ▶ Conclusions
- ▶ **2nd paper:** “Evaluating the effects of ANVUR policies on research productivity of Italian sociologists”; by: Aliakbar Akbaritabar; Flaminio Squazzoni (80% done)
 - ▶ Brief Introduction
 - ▶ Method & Data
 - ▶ Results
 - ▶ Conclusions (preliminary)

Outline of presentation 2/2

- ▶ **3rd paper:** “Bibliometric vs. peer review. Simulating VQR rankings of Italian sociology through quantitative indicators”; by: Aliakbar Akbaritabar; Flaminio Squazzoni (10% done)
 - ▶ Main questions
 - ▶ Method & Data
 - ▶ What we are going to do
- ▶ **4th paper:** “Gendered and ethnicized patterns of publication in top sociology journals”; by: Aliakbar Akbaritabar; Flaminio Squazzoni (40% done)
 - ▶ Main questions
 - ▶ Method & Data
 - ▶ What we are going to do

1st paper: “The conundrum of research productivity. A study on sociologists in Italy”

- ▶ **How is the situation of our community ?**
- ▶ **Is it a matter of embeddedness?**

Introduction

- ▶ Research work **doesn't** happen in a social and institutional vacuum
- ▶ Certain institutional and structural factors might influence scientist's work, including research priorities, publication strategies and collaboration (Shapin 2009)
- ▶ Understanding institutional embeddedness of scientist work is complicated due to the multi-layered nature of institutions.
- ▶ why Italy?
 - ▶ First, in 2006, the Italian government established an independent evaluation agency, i.e., **ANVUR**, with the aim to assess the performance of all research institutes; empowering the importance of productivity and quantitative indicators
 - ▶ Secondly, ANVUR was also involved in **developing common standards for the national habilitation** of all new associate and full professors, which **linked promotion and resources to research productivity**

Introduction

- ▶ Why sociologists?
 - ▶ First, 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
 - ▶ Secondly, research on scientists' productivity in Italy have looked mainly at the case of **hard sciences** (e.g., Abramo, D'Angelo, and Rosati 2016) stating that research in humanities and social sciences **cannot be examined quantitatively** (e.g., Abramo and D'Angelo 2014; Bellotti, Kronegger, and Guadalupi 2016; Nederhof 2006)
- ▶ Our aim:
 - ▶ fill this gap by providing a **quantitative analysis of research productivity of Italian sociologists**

Method & Data 1/2

- ▶ Scopus database
 - ▶ Title, authors' names, source title, year of publication, number of citations received, permanent link to the publication, authors' affiliations, abstract, keywords, references.
 - ▶ **Type of documents produced by each scientist** (i.e., "Article", "Review", "Book Chapter", "Book", "Erratum", "Editorial", "Note", "Conference Paper", "Article in Press", "Letter", "Short Survey"),
 - ▶ **Publication language** (i.e., "Italian", "English", "French", "Dutch", "Spanish", "Slovak", "German", "Portuguese", "Croatian", "Hungarian" were the languages Italian sociologists have been publishing).
- ▶ Our indexes:
 - ▶ **Productivity index**, $Pr = \frac{P}{t}$

Method & Data 1/2

- ▶ **FSS** (Fractional Scientific Strength) (by Abramo, D'Angelo, and Di Costa 2011; Abramo and D'Angelo 2011); a production function by microeconomic approach. (a) resources used by scientists, e.g., salary and time, and (b) scientific outputs, e.g., publications, and c) citations

$$FSS_R = \frac{1}{t} \sum_{i=1}^N \frac{c_i}{\bar{c}} f_i$$

- ▶ t = is the time a researcher has been publishing papers, *Sep2016 – p.date*
- ▶ N is the number of publications a sociologist have;
- ▶ i each paper of this author
- ▶ c_i is the number of citations of publication i
- ▶ \bar{c} is the average number of citations each paper have received
 - ▶ compared to other papers published in the same year
 - ▶ on each sub-community (SPS07 – SPS12 in MIUR categorization)
 - ▶ all sociologists in our sample
- ▶ f_i is inverse of number of authors of this paper

Method & Data 1/2

- ▶ **Internationalisation index**, number of authors with non-Italian affiliations a_f on the total number of authors of each paper a .

$$IntScore = \frac{1}{N} \sum_{i=1}^N \frac{a_f}{a}$$

- ▶ Co-authorship composition

$$AverageCoauthors = \frac{1}{N} \sum_{i=1}^N a_i$$

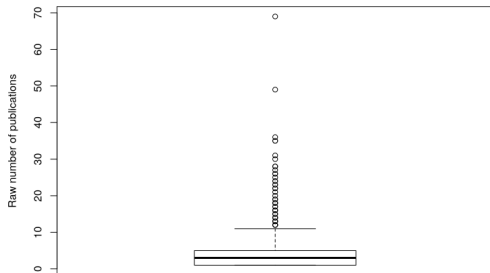
- ▶ **Co-author variety index**, which measured the degree of similarity or variety of co-authors in all publications for each scientist by inverse log-weighted similarity (Csardi and Nepusz 2006)

Hierarchical linear modeling

- ▶ Examining the importance of institutional embeddedness
- ▶ Each scientist is nested in different clusters
- ▶ Three clustering levels:
 - ▶ department; promotion and careers are eventually decided locally
 - ▶ scientific disciplinary sector
 - ▶ university, important in Italy after the last reform of the national university system in 2010; governance for incentivising and measuring scientist productivity, e.g., by establishing rewards for research excellence and allocating internal resources on productivity indicators.
- ▶ Following (Snijders and Bosker 1999; Faraway 2005; Zuur et al. 2009), we used hierarchical linear models
 - ▶ for example, an assistant professor of political sociology (i.e., this is the scientific disciplinary sector) who was enrolled in a department of social sciences in a given university could have different intercept (starting point) and slope (growth rate) in a regression model compared to a full professor of economic sociology in the same university and department.

Results

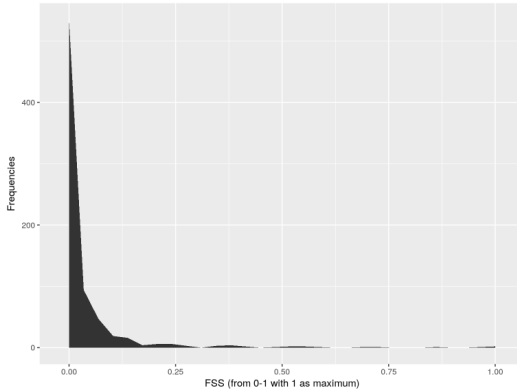
- Highly skewed distribution of publications; in line with (Nygaard 2015; Ramsden 1994; Coile 1977; Ellwein, Khachab, and Waldman 1989).



Box-plot distribution (black line is the median)

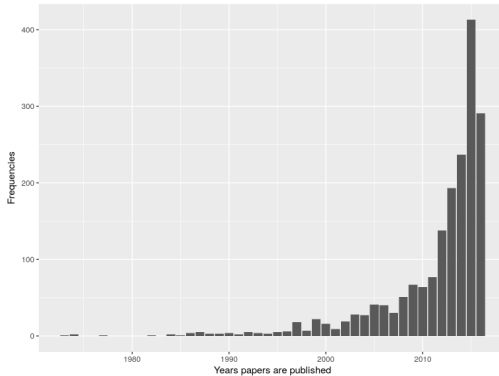
Results

- Skewed distribution of citations (e.g., Bornmann 2010): there were sociologists having up to nine records without a single citation.



Results

- Delayed recognition of scientific contributions (e.g., Garfield 1980 as cited in Abramo and D'Angelo (2011a))
- Before 2000 without any citation!



Results

Table 1: Comparative table of Multi level regression models

	Research productivity as dependent variable	
	FSS	Total Publications
	(1)	(2)
internationalisation	0.062*** (0.023)	0.099*** (0.025)
Coauthors Similarity	0.042 (0.048)	0.460*** (0.050)
Gender (male)	0.032*** (0.010)	0.033*** (0.010)
Associate professor	−0.011 (0.011)	0.015 (0.012)
Full professor	−0.025** (0.013)	0.035** (0.014)
Postdoc	0.081*** (0.016)	0.102*** (0.018)
Constant	0.036*** (0.011)	0.015 (0.012)
Observations	456	473
Log Likelihood	389.517	362.026
Akaike Inf. Crit.	−757.034	−702.052
Bayesian Inf. Crit.	−711.686	−656.302

Note:

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

Table 2: Comparative table of Macro level regression models

	Research productivity as dependent variable	
	FSS	Total Publications
	(1)	(2)
internationalisation	0.247*** (0.021)	0.189*** (0.022)
Coauthors Similarity	0.093 (0.070)	0.631*** (0.076)
Gender (male)	0.005 (0.004)	0.005 (0.004)
Associate professor	-0.008* (0.004)	-0.001 (0.005)
Full professor	-0.006 (0.005)	0.004 (0.005)
Postdoc	0.008 (0.007)	0.020*** (0.007)
Constant	0.016*** (0.005)	0.023*** (0.005)
Observations	477	477
R ²	0.251	0.267
Adjusted R ²	0.242	0.257
Residual Std. Error (df = 470)	0.040	0.043
F Statistic (df = 6; 470)	26.307***	28.470***

Note:

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

So what ?

- ▶ If individual, institutional and contextual factors could explain differences in research productivity.
- ▶ Internationalization is key to increase productivity. (e.g. F. Narin, Stevens, and Whitlow 1991; Francis Narin and Whitlow 1991; Khor and Yu 2016)
- ▶ Collaborating with colleagues abroad; being more exposed to international standards of research, have higher probability to publish in prestigious journals and be recognized by an international audience
- ▶ Relatively good performance of younger scholars compared to older generations - Scientific disciplinary sectors have no influence on scientist productivity
- ▶ These sectors seem less relevant to define research standards. They are a political and administrative “ontology” rather than an institutional scaffold that establishes research standards and promote excellence and innovation.
- ▶ This could bring us to re-discuss the importance of disciplinary sectors-dominated hiring and in National research assessments

Our limitations

- ▶ Coverage of Scopus
- ▶ Could not compensate by collecting data on Google Scholar, as only a few sociologists in Italy have activated a Google Scholar profile (19.7 %) While time effects were controlled on citations, our analysis did not consider important temporal aspects, such as the changing academic status of sociologists over their career, which might have relevant implications on productivity and important cumulative effects.
- ▶ Improve the analysis on embeddedness factors by looking at multi-level network effects, e.g., understanding whether collaboration networks and institutional embeddedness self-reinforce each other in generating effects on productivity.

2nd paper: “Evaluating the effects of ANVUR policies on research productivity of Italian sociologists”

- ▶ **Can policies reorient publication patterns against endogenous forces?**

Repeated measurements Anova analysis with nested structure

- ▶ All publications of authors in **5** full years **before** and **after** ANVUR (2006-2010; 2011-2015)
- ▶ Examine the importance of institutional embeddedness
- ▶ Compare individuals embedded in different institutional levels before and after ANVUR
 - ▶ Academic level (status)
 - ▶ Scientific disciplinary sector
- ▶ Following (Snijders and Bosker 1999; Faraway 2005; Zuur et al. 2009), we used repeated measurements ANOVA models with nested structure
- ▶ An example (**within** and **between** group measurements)
 - ▶ Assistant professors are significantly more productive, measured by FSS and number of papers, compared to associate or full professors?
 - ▶ How assistant professors' research productivity have changed after ANVUR compared to before?

Results 1/4

	Contrasting within and between group measurements				
	FSS as dependent variable				
	(1)	(2)	(3)	(4)	(5)
SPS/08	0.009 (0.009)				
SPS/09	0.034 ^{***} (0.012)				
SPS/10	0.020 (0.016)				
SPS/11	0.029 (0.020)				
SPS/12	-0.005 (0.020)				
Associate professors		0.026 ^{***} (0.009)			
Full professors		0.051 ^{***} (0.010)			
Level changed from 2010			0.002 (0.008)		
Medium FSS at 2010				0.010 (0.008)	
High FSS at 2010				0.265 ^{***} (0.010)	
Medium n.o papers at 2010					0.111 ^{***} (0.016)
High n.o papers at 2010					0.391 ^{***} (0.032)
FSS after ANVUR	-0.007 (0.007)	0.002 (0.006)	-0.015 ^{***} (0.006)	0.006 (0.004)	-0.002 (0.011)
SPS/08 vs FSS af. ANVUR	-0.006 (0.010)				
SPS/09 vs FSS af. ANVUR	-0.028 ^{***} (0.013)				
SPS/10 vs FSS af. ANVUR	-0.017 (0.018)				
SPS/11 vs FSS af. ANVUR	-0.030 (0.022)				
SPS/12 vs FSS af. ANVUR	0.001 (0.022)				
Associate prof. vs FSS af. ANVUR		-0.019 [†] (0.010)			
Full prof. vs FSS af. ANVUR		-0.049 ^{***} (0.011)			
Level changed vs FSS af. ANVUR			0.001 (0.009)		
Medium FSS at 2010 vs FSS af. ANVUR				-0.002 (0.009)	
High FSS at 2010 vs FSS af. ANVUR				-0.199 ^{***} (0.012)	
Medium n.o papers at 2010 vs FSS af. ANVUR					-0.080 ^{***} (0.018)
High n.o papers at 2010 vs FSS af. ANVUR					-0.268 ^{***} (0.035)
Constant	0.018 ^{***} (0.006)	0.009 (0.006)	0.028 ^{***} (0.005)	0.000 (0.004)	0.008 (0.010)

Results 2/4

- ▶ SPS/09 higher FSS
- ▶ SPS/09 decreased FSS after ANVUR
- ▶ Associate & Full professors higher FSS
- ▶ Associate & Full professors decreased FSS after ANVUR
- ▶ High FSS at 2010, higher FSS at 2016, compared to ppl with low FSS at 2010
- ▶ High FSS at 2010 decreased FSS after ANVUR
- ▶ Medium & High number of papers published at 2010, higher FSS at 2016, compared to ppl with low number of papers published at 2010
- ▶ Both have decreased FSS after ANVUR

Results 3/4

	Contrasting within and between group measurements				
	N.o papers as dependent variable				
	(1)	(2)	(3)	(4)	(5)
SPS/08	-0.021 (0.014)				
SPS/09	0.009 (0.017)				
SPS/10	-0.008 (0.025)				
SPS/11	0.037 (0.032)				
SPS/12	-0.012 (0.025)				
Associate professors		0.053 ^{***} (0.014)			
Full professors		0.051 ^{***} (0.014)			
Level changed from 2010			-0.011 (0.012)		
Medium FSS at 2010				0.005 (0.012)	
High FSS at 2010				0.160 ^{***} (0.014)	
Medium n.o papers at 2010					0.120 ^{***} (0.012)
High n.o papers at 2010					0.418 ^{***} (0.023)
N.o papers after ANVUR	-0.004 (0.010)	0.023 ^{**} (0.010)	-0.003 (0.008)	0.007 (0.010)	0.055 ^{***} (0.008)
SPS/08 vs papers af. ANVUR	0.027 [*] (0.014)				
SPS/09 vs papers af. ANVUR	0.010 (0.017)				
SPS/10 vs papers af. ANVUR	0.009 (0.026)				
SPS/11 vs papers af. ANVUR	-0.049 (0.032)				
SPS/12 vs papers af. ANVUR	0.005 (0.027)				
Associate prof. vs papers af. ANVUR		-0.032 ^{**} (0.014)			
Full prof. vs papers af. ANVUR		-0.018 (0.014)			
Level changed vs papers af. ANVUR			0.020 [*] (0.012)		
Medium FSS at 2010 vs papers af. ANVUR				0.039 ^{***} (0.013)	
High FSS at 2010 vs papers af. ANVUR				-0.007 (0.015)	
Medium n.o papers at 2010 vs papers af. ANVUR					-0.044 ^{***} (0.013)
High n.o papers at 2010 vs papers af. ANVUR					-0.196 ^{***} (0.024)
Constant	0.054 ^{***} (0.010)	0.018 [*] (0.010)	0.054 ^{***} (0.007)	0.026 ^{***} (0.009)	0.000 (0.007)

Results 4/4

- ▶ SPS/08 increased n.o. papers after ANVUR
- ▶ Associate & Full professors higher n.o. papers
- ▶ Associate professors decreased n.o. papers after ANVUR
- ▶ ppl who have changed academic level from 2010 till 2016 publish higher after ANVUR
- ▶ High FSS at 2010, higher n.o. papers at 2016, compared to ppl with low FSS at 2010
- ▶ Medium FSS at 2010 increased n.o. papers after ANVUR
- ▶ Medium & High number of papers published at 2010, higher n.o. papers at 2016, compared to ppl with low number of papers published at 2010
- ▶ Both have decreased n.o. papers after ANVUR

So what ?

- ▶ We **DO NOT** see a uniform trend of increase in research productivity (both FSS and n.o. papers) after ANVUR to conclude that this trend has been under the effect of National policy change
- ▶ Instead, we see some sub-group and sub-category level similarity and differences which seem to be mainly due to personal and individual level differences, that can be further looked into by analyzing individual coauthorship networks structure and evolution

3rd paper: “Bibliometric vs. peer review. Simulating VQR rankings of Italian sociology through quantitative indicators”

- ▶ **Is the “anti-quantitative indicators” belief of peer review supporters in research assessment only a mantra?**

What we are going to do next?

- ▶ Peer review evaluation (selecting three of your already accepted and published papers to be peer reviewed by unknown number of anonymous reviewers, 6 for research based positions)
- ▶ Time period of publications: 2004-2010
- ▶ They use bibliometric indexes too, but among Social Sciences only for Economics and Statistics (area 13) and in a small part of Psychology (area 11)
- ▶ Bibliometric evaluation of research productivity
 - ▶ possibility of looking at all the publications
 - ▶ time and cost effective
 - ▶ a good complementary approach to peer review
- ▶ We have extracted subset of our data on publications from beginning of 2011 till end of 2014
- ▶ Evaluating the research productivity by both FSS and number of papers published (taking citations each paper has received, as index of impact, into account)
- ▶ Next we need to write the analysis scripts, and report

4th paper: “Gendered and ethnicized patterns of publication in top sociology journals”

- ▶ **Can we find gendered and ethnicized publication patterns in top sociology journals?**
- ▶ **Are these “biases” due to social mechanisms of scientific excellence?**
- ▶ A collaboration with Michele Lamont, Harvard university
- ▶ Replicating some of the main results of a research from MIT **see here**

Main questions of paper 1/2

- ▶ Is publications in AJS and ASR, which are sociology's top ranked journals biased and limited to a group of scientists?
- ▶ Is there a gender bias in publishing in AJS and ASR ?
- ▶ Is there a USA non-USA bias between where scientists are affiliated ?
- ▶ Is there an university and university ranking bias ?
- ▶ Is there an ethnicity bias? (White, Black, Hispanic, Asian, other)
- ▶ How each of the above biases (in case they exist) has evolved through the time ?
- ▶ Is there a bias toward star scientists to publish more ? (rich gets richer ? or not? -> through preferential attachment assessment in scale free network approach)

Main questions of paper 2/2

- ▶ Can we find cohesiveness between riches working more with rich people, or a core periphery structure of riches surrounded by poorer (in terms of citations as well as total number of publications in the sample)
- ▶ Is there a community and sub-groups bias ? (disconnected sub-groups who sometimes work together, but they are more cohesive within than between themselves ? -> through small worlds assessment)
- ▶ Can we observe an overall cohesive structure (which is the case in interdisciplinary sciences, with overlapping collaborations between different parts of community)
- ▶ Can we find other structurally deterministic properties that can help us define who is working with whom ? (through blockmodelling and looking at network of collaborations considering authors attributes, like gender, university affiliations, country, city etc)

Preliminary Results 1/4

- ▶ Data has been gathered on January 20th 2017
- ▶ First issue until 2017
- ▶ Total of 4127
 - ▶ American Journal of Sociology, 1153 papers (27.94 %)
 - ▶ American Sociological Review, 1440 papers (34.89 %)
 - ▶ Annual Review of Sociology, 520 papers (12.60 %)
 - ▶ Social Networks 1014 papers (24.57 %)
- ▶ Social networks as a specific (and rather more recent than others) to be compared, and ARS to checkout patterns in AJS and ASR
- ▶ Based on the current (latest) affiliations recorded in Scopus, there are 3019 (66.24 % of total) authors affiliated in USA, and 1539 (33.76 % of total) scientists affiliated out of USA

Preliminary Results 2/4

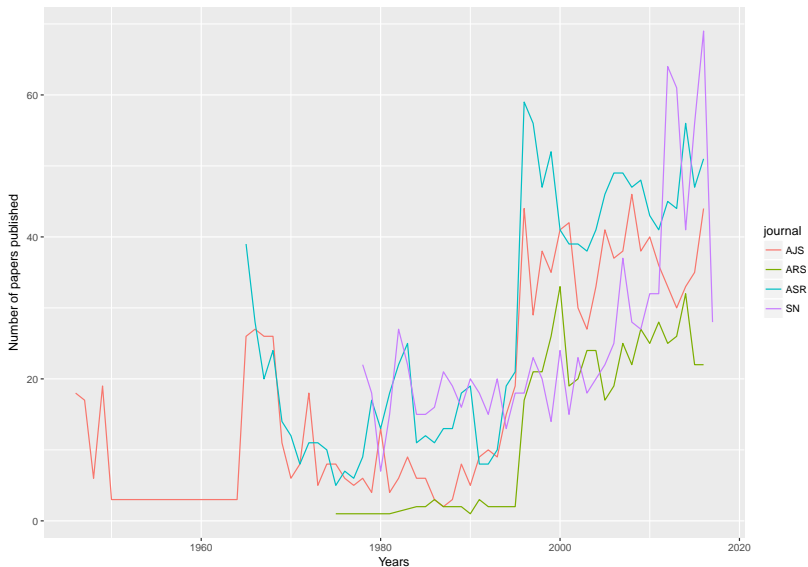
US cities	Freq
New York	185
Ann Arbor	102
Los Angeles	95
Palo Alto	89
Cambridge	88

And below we see the five cities out of US with highest number of authors in this sample:

non-US ities	Freq
London	56
Toronto	51
Amsterdam	44
Oxford	31
Utrecht	31

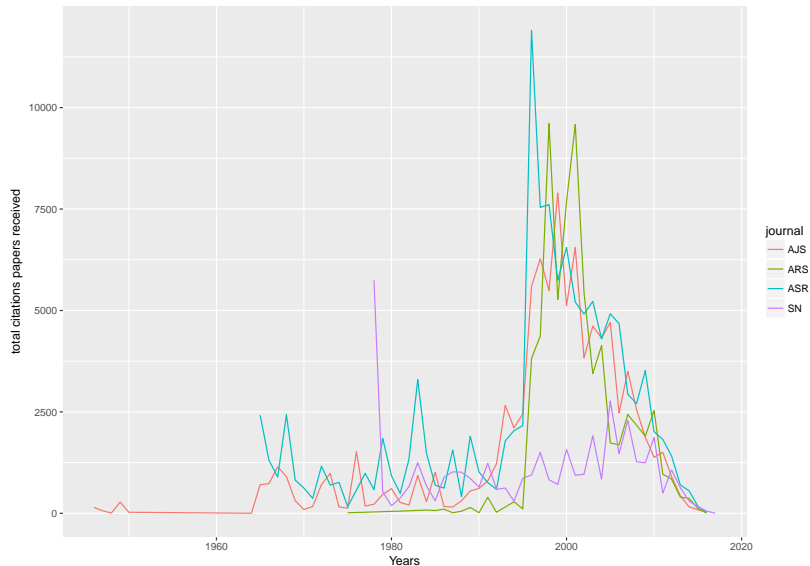
Preliminary Results 3/4

Comparative plot of number of papers published in each journal throughout the time



Preliminary Results 4/4

Number of citations those papers have received



What we are going to do next?

- ▶ Taking out **gender** of the authors based on their first name
- ▶ Take out their **ethnicity** based on the available information including, surname, first name, country, city and gender
- ▶ Cleaning data on university names and affiliations
- ▶ Add the latest rankings of universities to the picture intending to evaluate the probable ivy league universities bias in publications

Thanks a lot for your attention

References

- Abramo, Giovanni, and Ciriaco Andrea D'Angelo. 2011. "National-Scale Research Performance Assessment at the Individual Level." *Scientometrics* 86 (2): 347–64. doi:10.1007/s11192-010-0297-2.
- . 2014. "How Do You Define and Measure Research Productivity?" *Scientometrics* 101 (2): 1129–44. doi:10.1007/s11192-014-1269-8.
- Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Flavia Di Costa. 2011. "Research Productivity: Are Higher Academic Ranks More Productive Than Lower Ones?" *Scientometrics* 88 (3). Akadémiai Kiadó, co-published with Springer Science+ Business Media BV, Formerly Kluwer Academic Publishers BV: 915–28.
- Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Francesco Rosati. 2016. "The North–south Divide in the Italian Higher Education System." *Scientometrics* 109 (3): 2093–2117. doi:10.1007/s11192-016-2141-9.
- Bellotti, Elisa, Luka Kronegger, and Luigi Guadalupi. 2016. "The