

The false myth of the rise in Italian self-citations, and the impressively positive effect of bibliometric evaluations on the increase of the impact of Italian research

Pietro D'Antuono^{a,b}, Michele Ciavarella^{a,*}

^a*Department of Mechanics, Mathematics & Management, Polytechnic University of Bari, Via Edoardo Orabona, 4, 70126 Bari BA*

^b*Leonardo Helicopters Division Cascina Costa, Fatigue Office, Via Giovanni Agusta, 520 Samarate, Varese, IT 21017*

Abstract

It has recently been claimed by Baccini and coauthors that due to ANVUR's bibliometric evaluations of individuals, departments, and universities, in Italy there has been a surge in self-citations in the last ten years, thus increasing the “inwardness” of Italian research more than has happened abroad. We have studied the database of Ioannidis et al. published on 12 August 2019 of the one hundred thousand most “highly cited” scientists, including about two thousand Italians, and we found that the problem of self-citations in relation to this scientific elite is not significant in Italy, while perhaps observing a small deviation in the low scores in the rankings. The effect indicated by Baccini et al. consequently, does not seem worrying for the scientific elite (we quantified it in 2% of the total of scientists of the “best” one hundred thousand), and is probably largely concentrated in the further less cited scientists. Evaluation agencies like ANVUR should probably exclude self-citations in future evaluations, for the noise introduced by the young researchers. The overall state of health of the Italian research system and the positive effect of the ANVUR assessments are demonstrated by the number of Italian researchers in the top one hundred thousand, which has increased by comparing the “career” databased of 22 years, with that of the “young” researchers in the “2017” database. Italy, looking at the elite researchers, not only is not the most indulgent in self-citations, but has shown the best improvements, proving that the introduction of ANVUR had a positive effect. Indeed, all countries apart from Italy have suffered a decline, even substantial (−20% on a national Japan scale), of the number of researchers present in the 2017 data sets compared to career data. Italy instead shows a +0.2% on a global basis and an impressive +11.53% on a national basis.

*Corresponding author.

Indirizzo E-mail: mciaava@poliba.it (M. Ciavarella).

1 Introduction

The recent article by Baccini et al. [1] has received large coverage in the general Italian and foreign press (see Stella in the *Corriere Della Sera*¹), after suggesting a “boom” of self-citations of Italian scientists and therefore a deleterious effect of the evaluations of the Italian National Agency for Evaluation of the University System and Research (ANVUR)² which has been set in Italy in 2009, for evaluating individuals, departments and universities with bibliometric parameters. This largely critical article had a greater effect on the press than positive results in the news, such as the one that indicates that in 2012 Italian publications have surpassed those in the United States in terms of weighted citation impact, or that Italy has reached second place after the United Kingdom among the G8 countries, despite the level of public spending on research and development is still (1.3%) well below the European average (1.9%) [2, 3, 4]. We therefore tried to read data starting from a different perspective of elite scientists, stemming from an even more recent article, the ranking of Ioannidis et al. [5] on PLoS Biology³, which recognizes that citation metrics are widely but sometimes improperly used. Ioannidis and co-authors have published data on over 100,000 leading scientists⁴ who provide standardized information on citations, eliminating self-citations and even “citation farms”, thus overcoming all the problems Baccini et al. raised, although limited to elite scientists, and not the full 7.5 million scientists in the Scopus database. As Prof. Ioannidis says in a recent *Nature* coverage of Baccini’s work, Baccini’s effect is not extreme, and “could be due to chance. If the increase were real, it could be attributable to a minority of researchers”⁵. This statement motivated the present study and the question we asked ourselves is: “Is it a minority or the tail of less-rated researchers (either because they are poorly performing or because they are young newcomers) trying to beat the system to rig the bibliometric evaluations through the few self-citations that they can manufacture?”. Baccini et al. observe that the “inwardness” (a parameter defined by them as an index of self-reference) has undergone an increase in slope since 2009, ten years ago. Is this true also for elite scientists in Ioannidis et al database? And it this true for other countries?

This note is therefore structured as follows:

- (i) the concepts of citation ratio and ranking ratio that are used in this work are defined,
- (ii) (therefore an analysis is made of the citation ratio distribution and the ranking ratio of the data within the single selected sample countries;
- (iii) the same analysis is then extended to the single set of countries under exam, in order to obtain the real trend of the position fluctuation in the ranking relative to the weight of the self-citations.
- (iv) the study is specialized to the “Italy in the world” case, making it possible to draw the final conclusions.

Ioannidis’ data are the result of a very complete analysis of the data present on Scopus, even more than the data used by Baccini on SciVal.com, as Ioannidis’ career database dates back to the 1960s, providing a long-term performance of the single researchers, in order to evaluate their position in the Italian or world rankings, and the change of this position with and without self-citations. The data are separated in two databases: one for for “career” spanning a 22 years, where there are scientists of greater standing and already

¹ https://www.corriere.it/cronache/19_settembre_11/i-professori-si-citano-soli-cosi-si-gonfia-ricerca-c471954a-d4cf-11e9-8dcf-5bb1c565a76e.shtml (in Italian)

² <https://www.roars.it/online/citarsi-addosso-ascesa-scientifica-dellitalia-no-solo-doping-per-inseguire-i-criteri-anvur/> (in Italian)

³ <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000384> (in Italian)

⁴ The word scientists here is used improperly to refer to all the components of the lists made by Ioannidis et al.

⁵ Richard Van Noorden, Italy’s rise in research impact pinned on ‘citation doping’, Citation of Italian-authored papers by Italian researchers rose after the introduction of metrics-based thresholds for promotions. *Nature NEWS* 13 SEPTEMBER 2019

well-established career, and the “2017” classification, among which it is interesting to highlight both the “young promises” and the “elderly” who continue giving their contribution to the Literature.

2 Data analysis for the year 2017 Ioannidis database

As already mentioned the data used are part of a large database collected by Ioannidis et al. [2] and published in the authoritative journal PLoS Biology. In this study, the data for the single year (2017) were considered, whose file contains the data of the 106,368 most cited scientists in the world. We have considered six sample nations, or a base of 64,368 scientists, equivalent to about 60.5% of the total. The six countries examined are the United States (USA – 42,455 scientists), Great Britain (GBR – 9,467 scientists), Germany (DEU – 5,225 scientists), Japan (JPN – 2,674 scientists), Italy (ITA – 2,303 scientists) and France (FRA – 2,193 scientists). Of the large database provided by Ioannidis and co-authors, the number of citations per scientist with and without self-citations was considered in this note. The citation ratio q_c is therefore simply defined as the ratio between self-citations and citations by others, while the ranking ratio q_r is defined as the ratio of the difference between position in the ranking with and without self-citations and the position including self-citations, viz.:

$$\begin{aligned} q_c &= \text{autocites} / (\text{cites} - \text{autocites}) & (a) \\ q_r &= [\text{pos}(\text{cites}) - \text{pos}(\text{cites} - \text{autocites})] / \text{pos}(\text{cites}) & (b) \end{aligned} \quad (1)$$

Therefore, a scientist particularly accustomed to self-referencing could have a q_c even greater than one, while the ranking ratio is a number that, if negative, indicates a deterioration in the ranking after excluding self-citations (vice versa if it is positive).

2.1 Analysis on a national basis

2.1.1 Citation ratio

In the analysis for individual countries, we have considered the excursion in the national classification including or excluding self-citations, as well as analyzing the distribution of the citation ratio for each nation. Collecting data for each of the nations in the form of a histogram, we observed that the national distributions of q_c remain similar (see Table 1).

Table 1: trend of the citation ratio at national level for the 2017 database. The initials WRD identify the “world” as a union of the nations in question. Figures truncated to $q_c \geq 1$.

2017	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_c=0$	8	36	33	22	224	1087	1410
$0 < q_c \leq 0.1$	504	1343	937	663	4037	20772	28256
$0.1 < q_c \leq 0.2$	852	1883	991	891	3203	13655	21475
$0.2 < q_c \leq 0.3$	466	1055	397	367	1219	4171	7675
$0.3 < q_c \leq 0.4$	217	436	156	121	417	1483	2830
$0.4 < q_c \leq 0.5$	111	214	71	55	142	563	1156
$0.5 < q_c \leq 0.6$	66	107	34	27	76	290	600
$0.6 < q_c \leq 0.7$	28	58	12	17	72	157	344
$0.7 < q_c \leq 0.8$	23	33	11	11	34	108	220
$0.8 < q_c \leq 0.9$	15	21	15	6	12	48	117
$0.9 < q_c \leq 1$	14	13	2	5	5	31	70

From Table 1 we note that q_c reaches the peak value typically between 0.1 and 0.2, or between 0 and 0.1 for GBR and USA. That is, most probably an ITA, DEU, JPN and FRA scientist tends to cite himself once every

seven citations received, while a GBR or US scientist will tend to be more likely to be self-cited once every around twenty citations received. Based on the trend of the data, we assumed that a two-parameter Weibull distribution could fit the trend of the citation ratios. Therefore, we calculated the parameters of the Weibull curves by applying the maximum likelihood estimate (MLE) from which we could trace Figure 1, which clearly shows that, apart from USA and GBR, ITA data are perfectly in agreement with FRA, JPN, and DEU. The global data (WRD) is obviously heavily influenced by the USA and GBR scientists, which alone make up about 80% of the statistical sample. Surprisingly, the DEU (Germany) data appear to show the highest mode, albeit being very similar compared to the others.

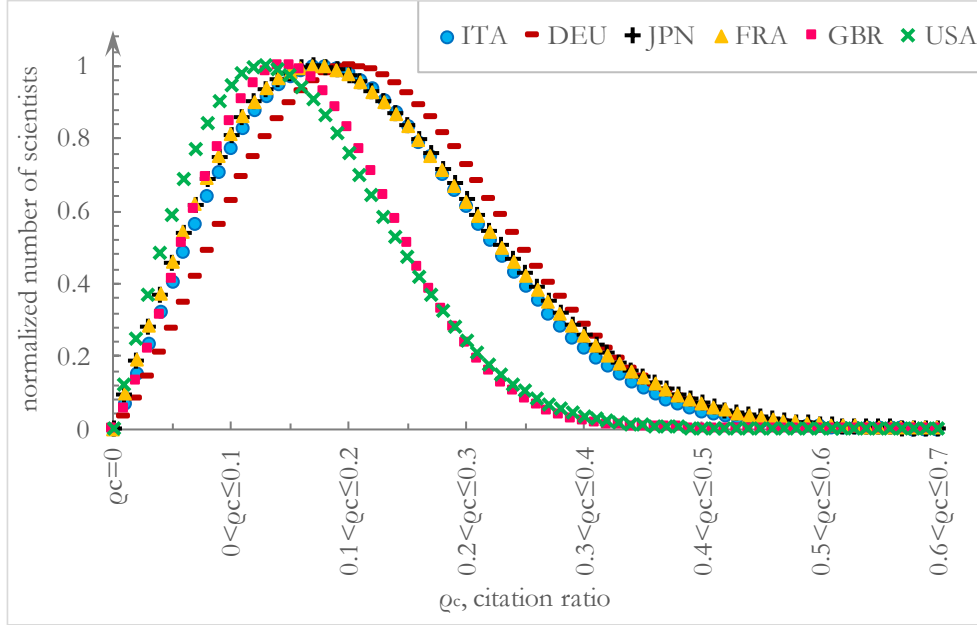


Figure 1: Distribution of the country-by-country citation ratio using Weibull probability density functions

It is noteworthy that the data in question do not, however, consider the “very smart” (in the negative sense), i.e. those who have a citation ratio greater than one, indicating that they cite themselves more than others cite them. This case has been singularly studied and the result would seem to show a greater tendency in the ITA block to self-citation within the investigated sample, as shown in Table 2, where both the absolute number and the percentage of the analyzed national sample is shown. Therefore, in this type of analysis, a certain “Baccini effect” would seem to be found, however the conditional is compulsory since we do not know with certainty whether this data is statistically significant, being only a mere 2% of the total. In addition to this, Table 2 also shows the average position in absolute and relative terms of the “smart” scientists, highlighting the theory we put forward: typically, those who tend to self-cite are in a low position at the national level. This position fluctuates, in a normalized ranking on a percentage basis, between 73 and 92 cents with the Italians ranking around halfway with 84%.

Table 2: Number of scientists with a citation ratio greater than one and a percentage of the national total

	ITA (2303)	DEU (5225)	JPN (2674)	FRA (2193)	GBR (9467)	USA (42455)
$q_c \geq 1$	48	26	15	8	26	90
$q_c \geq 1 / \text{Tot}$	2.08%	0.50%	0.56%	0.36%	0.27%	0.21%
average[$\text{pos}(q_c \geq 1)$]	1944	4093	2460	2011	6941	34255
average[$\text{pos}(q_c \geq 1) / \text{Tot}$]	84%	78%	92%	92%	73%	81%

2.1.2 Ranking ratio

As for the analysis of the ranking ratio q_r , we have produced a graph for each of the six countries studied, highlighting the samples with $q_r < 0$, i.e. in gray. those who without self-citations go down into the global ranking, and into black the samples with $q_r > 0$. The results are shown in Figure 2. The ITA data in gray do not show significant differences to the total, in fact, all the figures tend to show a belly at $q_r > 0$, and to widen quite evenly if not for the fact that they appear slightly more scattered.

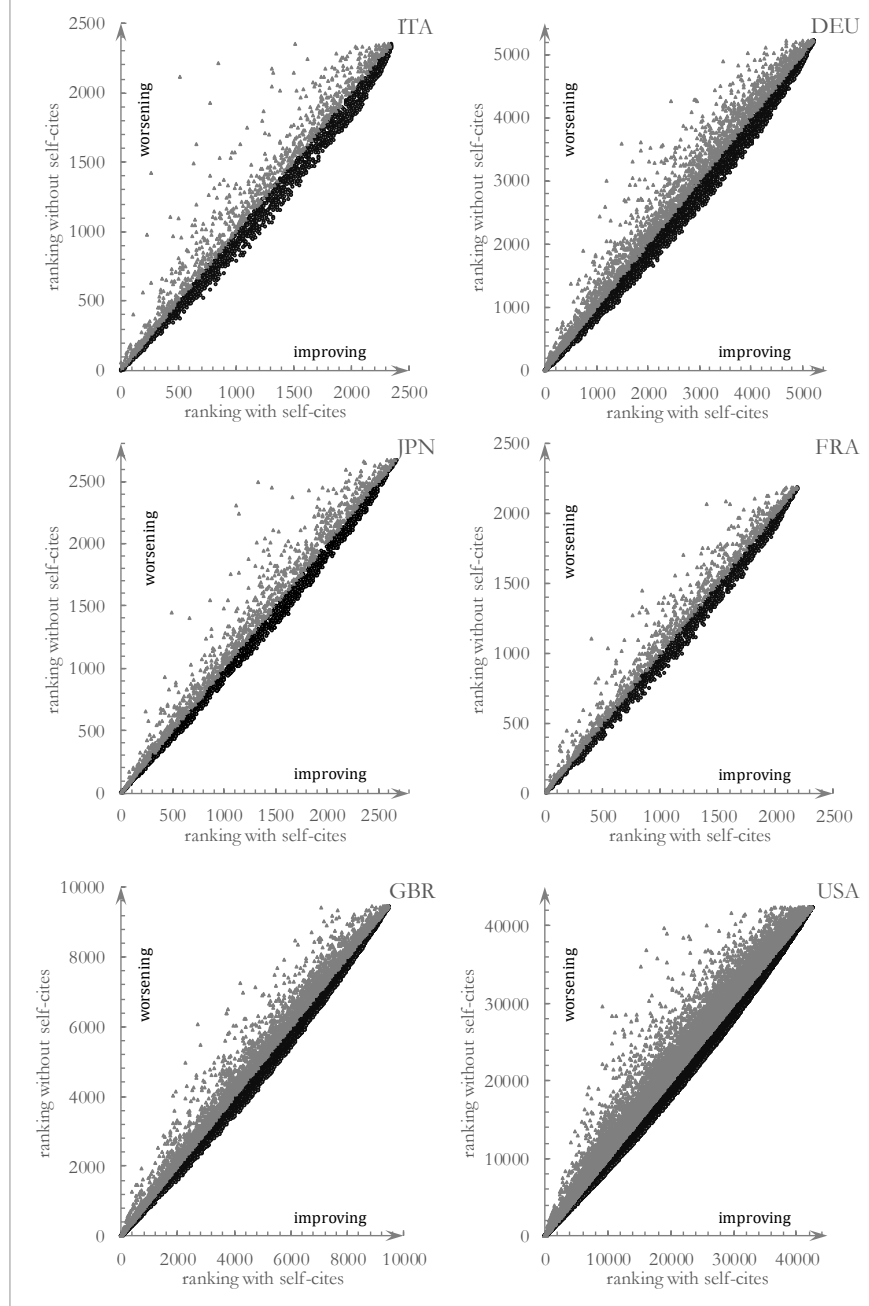


Figure 2: analysis of the trend of the ranking on a national basis. Note how the data with $q_r < 0$ (gray) corresponding to a drop in the ranking are more scattered than the data with $q_r > 0$ (black) identifying a rise in the ranking

We, therefore, decided to quantify this scatter in q_r by studying its distribution both for each nation and for the entire WRD sample. The distributions with the numeric data are shown in Table 3, from which we note that

the ratio between the number of samples in $-0.20 \leq q_r \leq 0$ and in $-0.40 \leq q_r \leq -0.2$ for ITA and DEU (Germany) is the highest of all the nations (and not Italy!). This indicates a higher worsening on average once the self-citations have been removed.

Table 3: trend of the ranking ratio on a national basis. The code WRD identifies the “world” intended as union of the nations under exam. All the distributions are centered around $0 \leq q_r \leq 0.2$ and show a comparable shape, confirming what observed in Figure 2

2017	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_r \leq -1$	27	30	7	7	28	120	233
$-1 < q_r \leq -0.8$	7	17	2	3	22	45	97
$-0.8 < q_r \leq -0.6$	17	25	17	10	38	105	202
$-0.6 < q_r \leq -0.4$	37	68	30	23	66	360	636
$-0.4 < q_r \leq -0.2$	95	238	109	66	312	1259	2149
$-0.2 < q_r \leq 0$	711	1739	911	787	3175	14486	21505
$0 < q_r \leq 0.2$	1395	2980	1578	1269	5755	25985	39254
$0.2 < q_r \leq 0.4$	59	128	20	28	71	95	291
$0.4 < q_r \leq 0.6$	5	0	0	0	0	0	0
$0.6 < q_r \leq 0.8$	0	0	0	0	0	0	0
$0.8 < q_r \leq 1$	0	0	0	0	0	0	0

The qualitative indication deriving from the histograms in Table 3 is shown quantitatively in Table 4. What appears is once again a certain “Baccini effect” in the “average of worsening” line which shows how, tending to exclude self-citations, the position of people who worsen in the ranking drops on average by 20%, compared to around 10% of other nations. This effect is attenuated by weighing the figure for the ratio between the number of samples that fell in the standings compared to the total, falling to 7.5% compared to around 4.5% of foreign nations. However, it should be highlighted that the number of Italians improving on the national ranking excluding self-citations is significantly higher than those worsening, showing the highest ratio of the entire statistical base (1.65 vs. ~1.55 on average). To understand if this data is significant, we compared it with the corresponding result for Italy based on WRD.

Table 4: Analysis of q_r on a national basis

Analysis of exclusion of self-citations	ITA (2303)	DEU (5225)	JPN (2674)	FRA (2193)	GBR (9467)	USA (42455)
risen in ranking	1459	3108	1598	1297	5826	26080
dropped in ranking	884	2100	1056	869	3610	16351
ratio(risen/dropped)	1.65	1.48	1.51	1.49	1.61	1.6
neutral	10	17	20	27	31	24
average(risen)	7.80%	7.10%	6.40%	6.10%	5.40%	5.10%
average(dropped)	-19.50%	-13.20%	-11.70%	-11.10%	-10.40%	-9.70%
weighted average(risen)	5.00%	4.20%	3.80%	3.60%	3.30%	3.10%
weighted average(dropped)	-7.50%	-5.30%	-4.60%	-4.40%	-4.00%	-3.70%

2.2 Specialization to ITA data on the WRD basis

To repeat and generalize the analysis on a “global” basis, we followed a procedure like that adopted on the

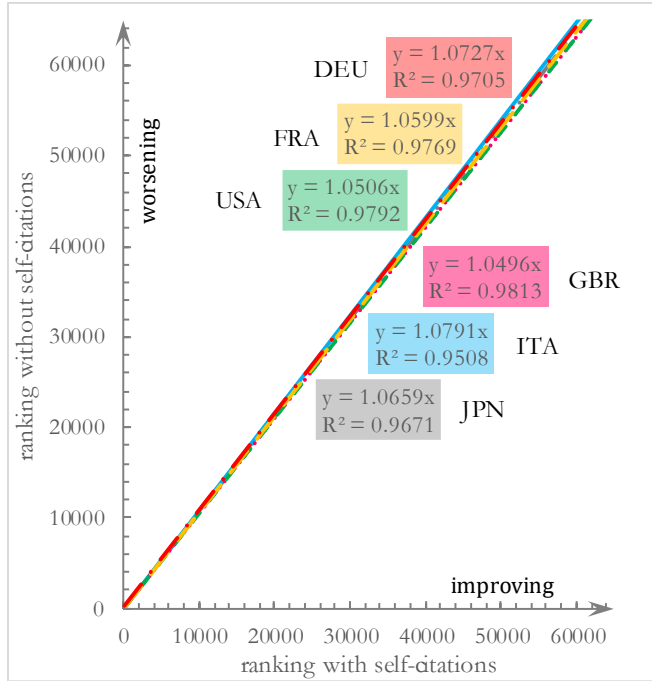


Figure 3: analysis of the ranking on the global basis for the only scientists having $q_r < 0$, corresponding to a drop in the rankings. All the regression lines are close to each other, with ITA showing a slightly more pronounced average worsening.

national basis by combining the scientists of the individual nations into the WRD statistical base. The derived table of global rankings for Italian scientists has been called ITW. The results obtained do not differ significantly from the results on the ITA basis, and this was predictable, given the similarity between the various nations and the distribution of citation and ranking ratios, as shown in Table 1, Figure 1 and Table 3. In Figure 3 we report the trend line showing the correlation between the ranking of researchers with $q_r < 0$ with and without self-citing, highlighting a very slight, almost imperceptible by eye, difference between the various countries. Italy indeed shows the worst datum, both in terms of slope and of data dispersion (lowest R^2), but we are talking about a difference between 1.079 and 1.072, or a mere 0.55%. Finally, we observe in Figure 4 that the data on a national scale are also representative of data on a global scale since the trend of the normalized position in terms of percentages with varying q_c for the ITA base and the ITW base are practically overlapping.

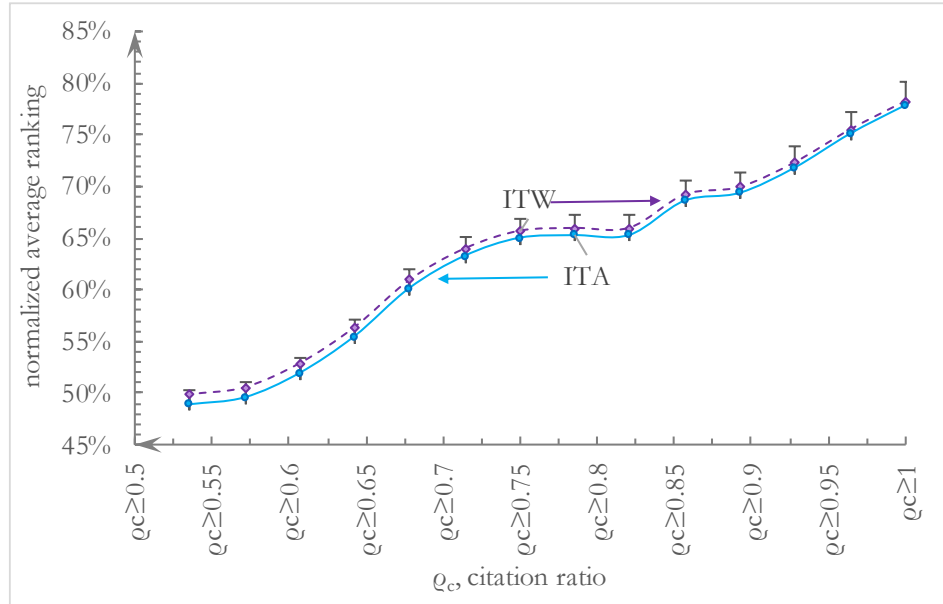


Figure 4: comparison between the normalized average (on a percent scale) ranking ITA and ITW with reducing the citation ratio threshold. The two curves are always within the 3% scatter error bands.

3 Discussion

3.1 The career data

To substantiate the findings and to confirm that there is not a trend in time of increasing the self-citations amongst the best performing scientists of Italy, a comparative analysis with the “career” database from Ioannidis et al. is required. If the q_c distribution tends to higher values in time, then the “Baccini effect” is demonstrated, otherwise, it is confuted. The same procedure described for the “2017” data has been adopted here for the career data, which presumably indicate a backward in time analysis, and the plots shown in Figure 5, and Table 5 and Table 6 have been obtained. As it is clear from the figure, the index of self-citations q_c does not show any alarming increase in ITA in time; conversely, the curve seems to shrink around the mode. Furthermore, surprisingly there has been a sudden increase in self-citations in FRA, and maybe this effect should be further investigated, if we consider also that the WRD 2017 for FRA sees 9% loss on the national base with respect to the WRD career database.

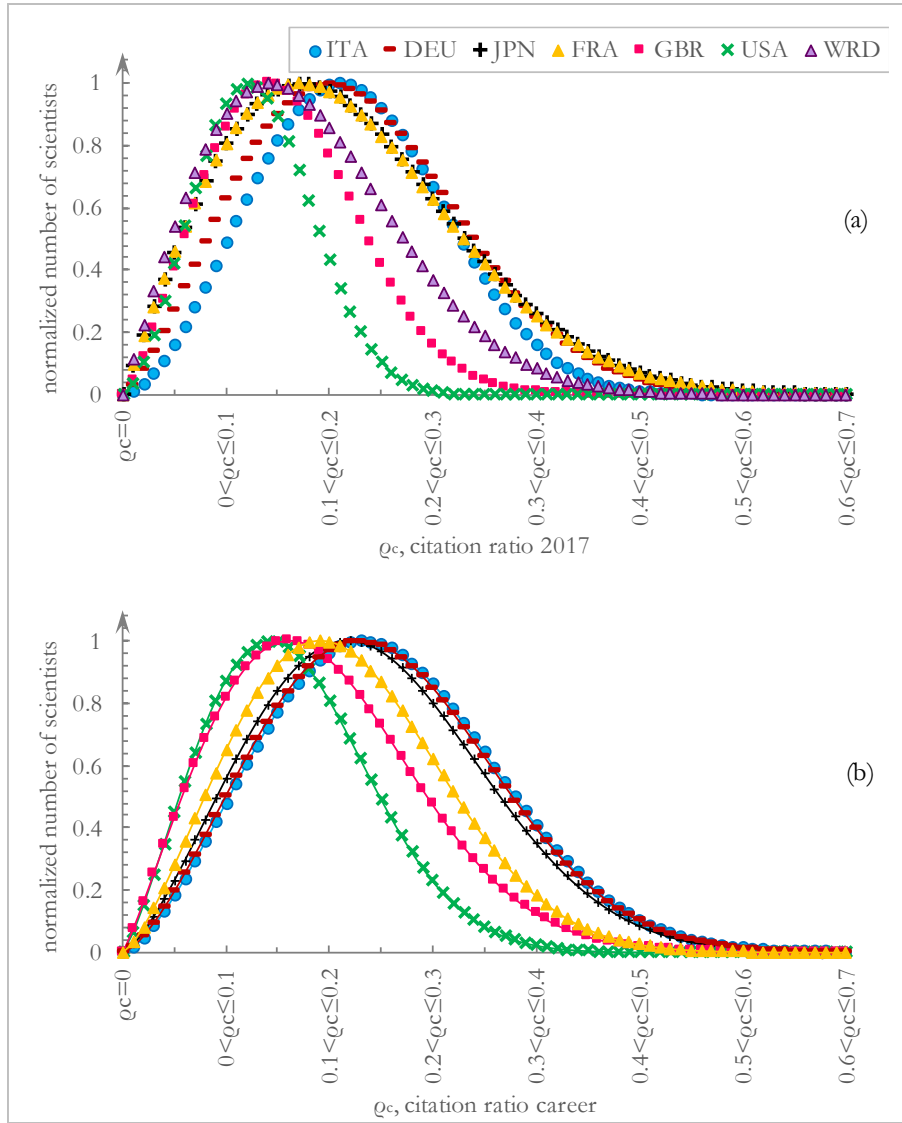


Figure 5: The distribution of the citation ratio, as regards ITA, DEU, and JPN, looks very similar in the 2017 (a) and the career (b) in terms of mode, while USA and GBR tend to self-cite more in the career than in 2017 data. Conversely, the most substantial increase in self-citations is observed in FRA

Table 5: trend of the citation ratio at national level for the career database. Figures are not truncated to $q_c \geq 1$, since no-one of the top 100,000 scientists in its entire career has recurred to self-citations more times than he has been cited.

Career	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_c=0$	0	2	1	0	5	32	40
$0 < q_c \leq 0.1$	 224	 681	 469	 437	 2879	 17604	 22294
$0.1 < q_c \leq 0.2$	 871	 2172	 1517	 1171	 4392	 19507	 29630
$0.2 < q_c \leq 0.3$	 543	 1460	 789	 502	 1673	 5872	 10839
$0.3 < q_c \leq 0.4$	 214	 576	 308	 162	 487	 1752	 3499
$0.4 < q_c \leq 0.5$	 92	 265	 134	 62	 168	 654	 1375
$0.5 < q_c \leq 0.6$	 51	 136	 68	 38	 76	 302	 671
$0.6 < q_c \leq 0.7$	 24	 77	 38	 19	 44	 134	 336
$0.7 < q_c \leq 0.8$	 22	 43	 25	 9	 20	 61	 180
$0.8 < q_c \leq 0.9$	 7	 18	 11	 9	 8	 33	 86
$0.9 < q_c \leq 1$	 2	 14	 3	 0	 10	 18	 47

Table 6: trend of the ranking ratio on a national basis for the career database. Again, all the distributions are centered around $0 \leq q_r \leq 0.2$ as in the 2017 database

Career	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_r \leq -1$	 13	 19	 5	 8	 23	 41	 110
$-1 < q_r \leq 0.8$	 6	 10	 5	 7	 15	 44	 84
$-0.8 < q_r \leq -0.6$	 6	 26	 10	 8	 17	 101	 184
$-0.6 < q_r \leq -0.4$	 18	 58	 38	 14	 54	 228	 429
$-0.4 < q_r \leq -0.2$	 73	 238	 127	 82	 286	 1164	 2131
$-0.2 < q_r \leq 0$	 745	 1924	 1230	 878	 3661	 17435	 25408
$0 < q_r \leq 0.2$	 1174	 3108	 1939	 1401	 5676	 26919	 40523
$0.2 < q_r \leq 0.4$	 28	 75	 27	 17	 48	 81	 243
$0.4 < q_r \leq 0.6$	 1	 1	 0	 0	 0	 0	 0
$0.6 < q_r \leq 0.8$	 0	 0	 0	 0	 0	 0	 0
$0.8 < q_r \leq 1$	 0	 0	 0	 0	 0	 0	 0

In light of the demonstrated fact that the effect of self-citations on highly cited scientists is negligible and with the aim of highlighting the overall state of health of the Italian research system and the probably positive effect of the ANVUR assessments, we show how the ITA presence has increased in the ranking of Ioannidis et al. comparing the quantity of scientists in the 2017 database, with respect to the “career” database, both in terms of absolute and percent increments. The emerging fact is that all the countries apart from Italy have suffered a decline, even substantial (−20% on a national JPN scale) of the number of samples present in the 2017 data sets compared to career data. Italy instead shows a +0.2% on a global basis and an impressive +11.53% on a national basis.

Table 7: Relative and absolute analysis of the presences by nation in the Career and 2017 databases

Presences in the rankings		ITA	DEU	JPN	FRA	GBR	USA
Absolute presence	Career	↓ 2065	↑ 5459	↑ 3382	↑ 2415	↑ 9780	↑ 46013
	2017	↑ 2303	↓ 5225	↓ 2674	↓ 2193	↓ 9647	↓ 42455
Percentual presence	Career	↓ 1.97%	↑ 5.20%	↑ 3.22%	↑ 2.30%	↑ 9.31%	↑ 43.81%
	2017	↑ 2.17%	↓ 4.91%	↓ 2.51%	↓ 2.06%	↓ 9.07%	↓ 39.91%
	National Δ	11.53%	-4.29%	-20.93%	-9.19%	-1.36%	-7.73%
	Global Δ	0.20%	-0.29%	-0.71%	-0.24%	-0.24%	-3.90%

4 Conclusions

The effect claimed by Baccini et al. that due to ANVUR’s bibliometric evaluations of individuals, departments, and universities, in Italy there has been a surge in self-citations in the last ten years, has been denied regarding the data of the hundred thousand most "highly cited" scientists in the world, including about two thousand Italians, being extremely limited. We have found that only those who are in a low position at the national level have a slight tendency to self-cite a lot. While we obviously advise evaluation agencies such as ANVUR to exclude self-citation in future assessments, particularly of young researchers or low rank researchers (which is very simple to do in all databases), we provide a set of original conclusions. Indeed, by comparing the “career” database of Ioannidis et al, with that of the young scientists “2017” database, we have found a much more important fact: all the countries we studied apart from Italy have suffered a decline, even substantial (−20% on a national JPN scale) of the number of scientists. Italy instead shows a +0.2% on a global basis and an impressive +11.53% on a national basis.

References

- [1] A. Baccini, G. De Nicolao, and E. Petrovich, ‘Citation gaming induced by bibliometric evaluation: A country-level comparative analysis’, *PLoS ONE*, vol. 14, no. 9, Sep. 2019.
- [2] Nature, Seven days: 6–12 December 2013, <https://www.nature.com/news/seven-days-6-12-december-2013-1.14335>
- [3] BIS, U. K. International Comparative Performance of the UK Research Base–2016. 2016, <https://www.elsevier.com/research-intelligence/research-initiatives/beis2016>
- [4] Nature, Editorial, ‘Memo to Italy’s president: your researchers need you’, 27 August 2019, <https://www.nature.com/articles/d41586-019-02560-1>
- [5] J. P. Ioannidis, J. Baas, R. Klavans, and K. W. Boyack, ‘A standardized citation metrics author database annotated for scientific field’, *PLoS biology*, vol. 17, no. 8, p. e3000384, 2019.

Il falso mito dell'aumento delle auto-citazioni, e l'impressionantemente positivo effetto delle valutazioni bibliometriche sull'aumento dell'impatto della ricerca italiana

Pietro D'Antuono^{a,b}, Michele Ciavarella^{a,*}

^aDepartment of Mechanics, Mathematics & Management, Polytechnic University of Bari, Via Edoardo Orabona, 4, 70126 Bari BA

^bLeonardo Helicopters Division Cascina Costa, Fatigue Office, Via Giovanni Agusta, 520 Samarate, Varese, IT 21017

Sommario

È stato recentemente sostenuto da Baccini e coautori che a causa delle valutazioni bibliometriche di ANVUR dei singoli, dei dipartimenti, e delle Università, in Italia si sia assistito ad una impennata delle autocitazioni negli ultimi dieci anni, aumentando così la “*inwardness*” (autoreferenzialità) della ricerca italiana più di quanto non sia accaduto all'estero. Abbiamo studiato in questo lavoro il database di Ioannidis et al. pubblicato il 12 agosto 2019 dei centomila scienziati più “altamente citati”, incluso circa duemila italiani, ed abbiamo constatato che il problema delle autocitazioni relativamente a questa élite scientifica non è significativo in Italia, osservando comunque una piccola deviazione negli “ultimi in classifica”. L'effetto indicato da Baccini et al. dunque, non appare preoccupante per la élite scientifica (lo abbiamo quantificato nel 2 % del totale di scienziati dei centomila “migliori”), ed è probabilmente in larga misura concentrato nelle fasce ulteriormente meno citate. In altre parole, tipicamente chi tende ad autocitarsi molto si trova in una posizione bassa a livello nazionale. Consigliamo certamente ad ANVUR di utilizzare nelle prossime valutazioni l'esclusione delle autocitazioni, peraltro molto semplice da eseguire in tutti i database (potrebbe essere lievemente più problematico riconoscere le “*citation farms*”, come comunque fatto da Ioannidis et al.). Ciononostante, lo stato di salute complessivo del sistema della ricerca italiana e l'effetto positivo delle valutazioni ANVUR sono dimostrati dal numero di ricercatori italiani nella top centomila, che è aumentato paragonando il database di Ioannidis per la “carriera” di 22 anni con quello dei ricercatori “giovani” nel database 2017. Pertanto, il caso dell'Italia non è il più eclatante in termini di autocitazioni (la Germania sembra più prona a ciò), e viceversa si è mostrato quello con i più marcati miglioramenti, mostrando che l'introduzione di ANVUR ha avuto un effetto globalmente positivo. Tutti gli stati a parte l'Italia hanno subito un declino, anche sostanziale (Giappone -20% su base nazionale), del numero di ricercatori presenti nel set di dati 2017 rispetto ai dati carriera. In controtendenza l'Italia che mostra un +0.2% su base globale ed un impressionante +11.53% su base nazionale.

*Corresponding author.

Indirizzo E-mail: mciaava@poliba.it (M. Ciavarella).

1 Introduzione

Il recente articolo di Baccini et al. [1] di cui hanno parlato molto la stampa italiana e straniera (cfr. Stella sul Corriere della Sera⁶), suggerendo un “boom” delle autocitazioni e quindi un effetto deleterio delle valutazioni di ANVUR⁷, introdotta in Italia nel 2009 per la valutazione di individui, dipartimenti ed università con parametri bibliometrici. Questo aspramente critico articolo ha avuto maggiore impatto sui media di quanto non ne abbiano avuto i risultati positivi della ricerca italiana, come quello che indica che nel 2012 le pubblicazioni italiane hanno sorpassato quelle negli Stati Uniti in termini di impatto citazionale pesato, o che l’Italia ha raggiunto il secondo posto dopo Regno Unito tra i Paesi G8. Pertanto, nonostante il livello di spesa pubblica in ricerca e sviluppo sia ancora (1.3 %) ben al di sotto della media europea (1.9 %), la ricerca italiana sembra continuare a migliorare [2, 3, 4]. Abbiamo quindi provato a dare una nostra lettura dei dati partendo dalla classifica di Ioannidis et al. [5] su PLoS Biology⁸ che riconosce come le metriche di citazione siano ampiamente utilizzate e riutilizzate in modo improprio. Ioannidis e coautori hanno pubblicato i dati riguardanti oltre 100,000 scienziati⁹ di spicco che forniscono informazioni standardizzate su citazioni, eliminando le autocitazioni e anche le “*citation farms*”, e superando quindi tutti i problemi di cui parlano Baccini e Stella. Come dice in un recente coverage su Nature del lavoro di Baccini, Ioannidis dichiara che i dati di Baccini non sono estremi, e potrebbero essere dovuti al caso. Se anche l’aumento fosse reale, potrebbe essere attribuibile ad una minoranza di ricercatori¹⁰. Questa affermazione ha motivato il presente studio e la domanda che ci siamo posti: “Si tratta di una minoranza, o della coda di ricercatori meno quotati (o perché poco performanti o perché giovani esordienti) che tentano di giocare delle carte avendo indici così bassi da poter essere “truccati” anche dalle pochissime autocitazioni che essi stessi possono fabbricare?”. Baccini et al. osservano che la *inwardness* (parametro da essi definito come indice di autoreferenzialità) ha subito un incremento di pendenza a partire dal 2009, dieci anni orsono. Ciò detto, è possibile che i parametri bibliometrici stabiliti da ANVUR in una certa percentuale possano aver influito sull’eventuale uso/abuso delle autocitazioni, sicché non è nostro scopo screditare la ricerca di Baccini e coautori, anzi, piuttosto questa nota è finalizzata al cercare altre cause che abbiano influito sul dimostrato aumento di autoreferenzialità che però non vadano ad intaccare la globalmente indiscussa qualità della ricerca *made in Italy*. La presente nota è quindi così strutturata: (i) vengono definiti i concetti di rapporto di citazione e rapporto di ranking che sono usati in questo lavoro, (ii) dunque viene effettuata una analisi della distribuzione del rapporto di citazione e del rapporto di ranking dei dati all’interno dei singoli paesi campione scelti; (iii) la stessa analisi viene poi estesa all’insieme unico delle nazioni in esame, in modo da ottenere l’andamento reale dell’oscillazione di posizione in classifica relativa al peso delle autocitazioni. (iv) Lo studio viene infine specializzato al caso “Italia nel mondo” permettendo di trarre delle ultime conclusioni. I dati di Ioannidis sono risultati di una analisi molto completa dei dati presenti su Scopus, anche più dei dati usati da Baccini su SciVal.com, in quanto risale indietro agli anni ’60 fornendo una performance a lungo termine del ricercatore singolo, in modo da poter valutare la sua posizione nel ranking italiano o mondiale, e il cambiamento di questa posizione con e senza autocitazioni. I dati si distinguono in classifica per “career”, dove ovviamente

⁶https://www.corriere.it/cronache/19_settembre_11/i-professori-si-citano-soli-così-si-gonfia-ricerca-c471954a-d4cf-11e9-8dcf-5bb1c565a76e.shtml

⁷ <https://www.roars.it/online/citarsi-addosso-ascesa-scientifica-dellitalia-no-solo-doping-per-inseguire-i-criteri-anvur/>

⁸ <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000384>

⁹ Il termine scienziati nella nota viene utilizzato in maniera impropria per riferirsi a tutti i componenti delle liste di Ioannidis e coautori.

¹⁰ Richard Van Noorden, Italy’s rise in research impact pinned on ‘citation doping’, Citation of Italian-authored papers by Italian researchers rose after the introduction of metrics-based thresholds for promotions. Nature NEWS 13 SEPTEMBER 2019

risultano scienziati di maggiore peso e carriera già ben avviata, e i classifica “2017”, tra cui è interessante segnalare sia le “giovani promesse” che gli “anziani” che continuano a produrre bene.

2 Analisi dei dati anno dal database Ioannidis 2017

Come già accennato, i dati utilizzati fanno parte di un vasto database collezionato da Ioannidis et al. [2] e pubblicato sulla autorevole PloS Biology. Per chi volesse dilettersi, le tabelle complete da noi utilizzate estratte da Ioannidis e coautori sono disponibili al link seguente: [22](#). Nel presente studio sono stati considerati i dati relativi al singolo anno (2017), il cui file contiene i dati dei 106,368 scienziati più citati al mondo. Abbiamo preso in considerazione sei nazioni campione, ovvero una base di 64,368 scienziati, equivalente a circa il 60.5 % del totale. Le sei nazioni prese in esame sono Stati Uniti (USA – 42455 scienziati), Gran Bretagna (GBR – 9467 scienziati), Germania (DEU – 5225 scienziati), Giappone (JPN – 2674 scienziati), Italia (ITA – 2303 scienziati) e Francia (FRA – 2193 scienziati). Del grande database fornito da Ioannidis e coautori, in questa nota sono stati considerati il numero di citazioni per scienziato con e senza autocitazioni. Il rapporto di citazione ρ_c è quindi definito banalmente come il rapporto tra le autocitazioni e le citazioni da parte di altri, mentre il rapporto di ranking ρ_r è definito come il rapporto della differenza tra posizione in classifica con e senza autocitazioni e la posizione incluse le autocitazioni, ovvero:

$$\begin{aligned} \rho_c &= \text{autocites}/(\text{cites} - \text{autocites}) & (a) \\ \rho_r &= [\text{pos}(\text{cites}) - \text{pos}(\text{cites} - \text{autocites})]/\text{pos}(\text{cites}) & (b) \end{aligned}$$

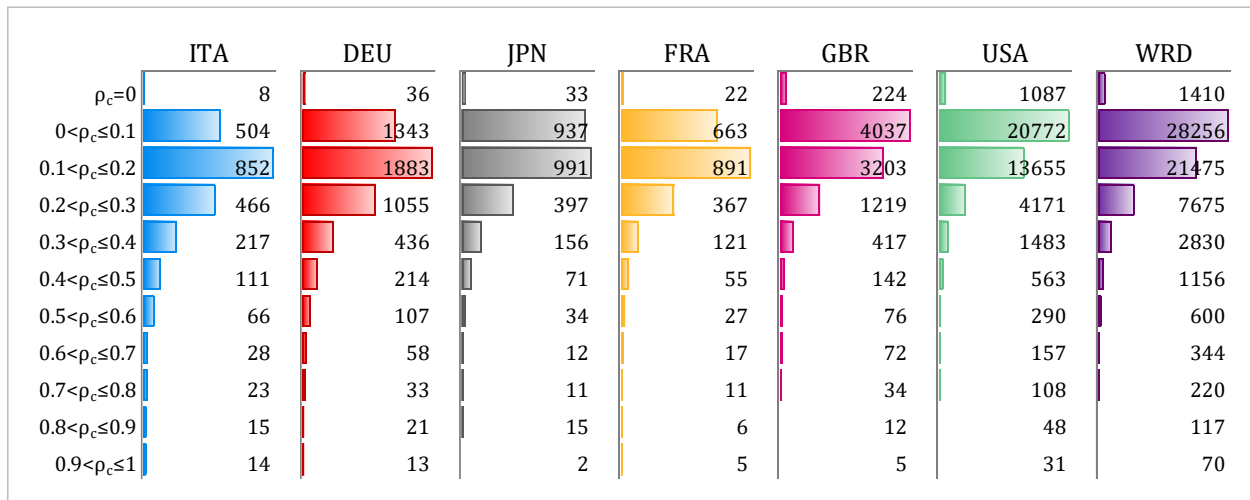
Va da sé che uno scienziato particolarmente avvezzo ad autocitarsi potrebbe avere un ρ_c anche maggiore di uno, mentre il rapporto di ranking è un numero che, se è negativo, indica un peggioramento in classifica dopo aver escluso le autocitazioni (viceversa se è positivo).

2.1 Analisi su base nazionale

2.1.1 Rapporto di citazione

Nell’analisi per singole nazioni abbiamo considerato l’escursione nella classifica nazionale includendo o escludendo le autocitazioni, oltre ad analizzare la distribuzione del rapporto di citazione per ogni nazione. Raccogliendo i dati per ognuna delle nazioni in forma di istogramma, abbiamo osservato che le distribuzioni nazionali di ρ_c si mantengono simili

Tabella 1: andamento del rapporto di citazione livello nazionale. La sigla WRD identifica il “mondo” inteso come unione delle nazioni in esame. Figure troncate per $\rho_c \geq 1$.



Dalla Tabella 1 notiamo come q_c raggiunga il valore di picco tipicamente tra 0.1 e 0.2, oppure tra 0 e 0.1 per GBR e USA. Ovvero, con maggiore probabilità uno scienziato ITA, DEU, JPN e FRA tenderà ad autocitare ogni sette citazioni ricevute, mentre uno scienziato GBR o USA tenderà ad autocitarsi con maggiore probabilità una volta ogni circa venti citazioni ricevute. In base all'andamento dei dati, abbiamo supposto che una distribuzione di Weibull a due parametri potesse fittare l'andamento dei rapporti di citazione. Pertanto, abbiamo calcolato i parametri delle curve di Weibull applicando il maximum likelihood estimate (MLE) da cui abbiamo potuto tracciare la Figura 1, che mostra chiaramente come, a parte USA e GBR, i dati ITA siano perfettamente in accordo con FRA, JPN e DEU. Il dato globale (WRD) è ovviamente influenzato pesantemente da USA e GBR che da sole costituiscono circa l'80 % del campione statistico. E, sorprendentemente, è proprio il dato DEU quello che appare mostrare la moda più elevata, sebbene di pochissimo rispetto agli altri.

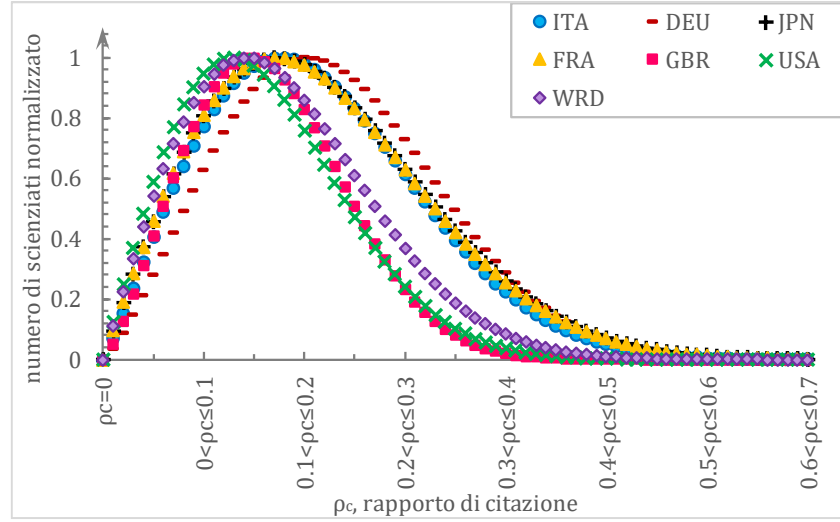


Figura 1: Distribuzione del rapporto di citazione paese per paese tramite funzione di densità di probabilità di Weibull

È importante far osservare che i dati in esame non tengono però conto dei “molto furbi”, ovvero coloro i quali hanno un rapporto di citazione maggiore di 1, indicatore di chi cita più volte se stesso di quanto non venga citato da altri. Questo caso è stato singolarmente studiato ed il risultato sembrerebbe mostrare una maggiore tendenza ITA ad autocitarsi all’interno del campione indagato, come mostrato in Numero di scienziati con rapporto di citazione maggiore di uno e percentuale rispetto al totale nazionale, dove sia il numero assoluto che la percentuale sul campione nazionale analizzato sono mostrati. Pertanto, in questo tipo di analisi, parrebbe riscontrabile un certo “effetto Baccini”, tuttavia il condizionale è d’obbligo in quanto non sappiamo dire con certezza se questo dato sia statisticamente rilevante, essendo il 2 % del totale. Oltre a ciò, Tabella 2 mostra anche la posizione media in termini assoluti e relativi degli scienziati “furbi”, evidenziando la teoria da noi avanzata: tipicamente chi tende ad autocitarsi molto si trova in una posizione bassa a livello nazionale. Tale posizione oscilla, in una classifica normalizzata su base percentuale, tra i 73 ed i 92 centesimi con gli italiani che si collocano circa a metà con un 84 %.

Tabella 2: Numero di scienziati con rapporto di citazione maggiore di uno e percentuale rispetto al totale nazionale

	ITA (2303)	DEU (5225)	JPN (2674)	FRA (2193)	GBR (9467)	USA (42455)
$q_c \geq 1$	48	26	15	8	26	90
$q_c \geq 1 / \text{Tot}$	2.08%	0.50%	0.56%	0.36%	0.27%	0.21%
media[pos($q_c \geq 1$)]	1944	4093	2460	2011	6941	34255
media[pos($q_c \geq 1$)]/Tot	84%	78%	92%	92%	73%	81%

2.1.2 Rapporto di ranking

Quanto all'analisi del rapporto di ranking q_r , abbiamo creato un grafico per ognuna delle sei nazioni studiate evidenziando in grigio i campioni aventi $q_r < 0$, i.e. coloro i quali senza le autocitazioni scendono nella graduatoria globale, ed in nero i campioni con $q_r > 0$. I risultati sono mostrati in Figura 2. I dati ITA in grigio non mostrano significative differenze rispetto al totale, infatti tutte le figure tendono a mostrare una pancia per quanto riguarda $q_r > 0$, ed ad allargarsi abbastanza uniformemente se non per il fatto che appaiono leggermente più scatterati.

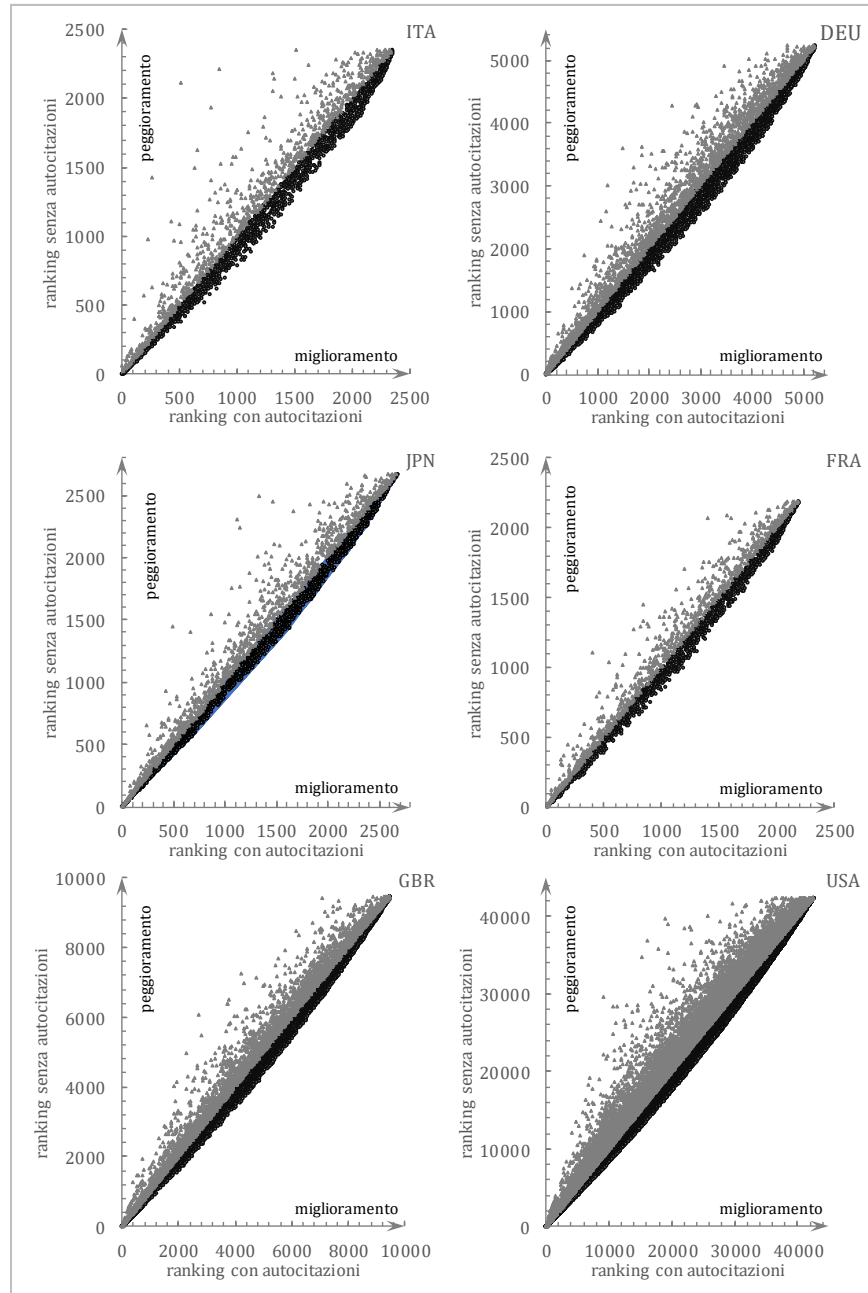


Figura 2: analisi dell'andamento della posizione in classifica su base nazionale. Si nota come i dati con $q_r < 0$ (grigio), corrispondenti ad un peggioramento in classifica, siano più dispersi di quelli con $q_r > 0$ (nero) che invece identificano un miglioramento in classifica

Abbiamo dunque deciso di quantificare questo scatter nel q_r studiandone la distribuzione sia per ogni nazione che per l'intero campione WRD. Le distribuzioni con i dati sono mostrate in Tabella 3, da cui si nota che il rapporto tra il numero di campioni in $-0.20 \leq q_r \leq 0$ e in $-0.40 \leq q_r \leq -0.2$ per ITA e DEU è il più alto tra tutte le nazioni. Ciò indica un peggioramento in media più elevato una volta rimosse le autocitazioni.

Tabella 3: andamento del rapporto di ranking livello nazionale. La sigla WRD identifica il “mondo” inteso come unione delle nazioni in esame. Tutte le distribuzioni sono centrate in $0 \leq q_r \leq 0.2$ e presentano forma paragonabile, confermando quanto mostrato in Figura 2

	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_r \leq -1$	27	30	7	7	28	120	233
$-1 < q_r \leq -0.8$	7	17	2	3	22	45	97
$-0.8 < q_r \leq -0.6$	17	25	17	10	38	105	202
$-0.6 < q_r \leq -0.4$	37	68	30	23	66	360	636
$-0.4 < q_r \leq -0.2$	95	238	109	66	312	1259	2149
$-0.2 < q_r \leq 0$	711	1739	911	787	3175	14486	21505
$0 < q_r \leq 0.2$	1395	2980	1578	1269	5755	25985	39254
$0.2 < q_r \leq 0.4$	59	128	20	28	71	95	291
$0.4 < q_r \leq 0.6$	5	0	0	0	0	0	0
$0.6 < q_r \leq 0.8$	0	0	0	0	0	0	0
$0.8 < q_r \leq 1$	0	0	0	0	0	0	0

L'indicazione qualitativa derivante dagli istogrammi in Tabella 3 è mostrata quantitativamente in Tabella 4. Ciò che appare è nuovamente un certo effetto Baccini nella riga “media dei peggioramenti” che mostra come tendenzialmente, escludendo le autocitazioni, la posizione delle persone che peggiorano in classifica scende mediamente del 20 %, rispetto ad un 10 % circa delle altre nazioni. Tale effetto viene attenuato pesando il dato per il rapporto tra il numero dei campioni che sono scesi in classifica rispetto al totale, scendendo ad un 7.5 % rispetto ad un 4.5 % circa delle nazioni estere. Va però evidenziato che il numero di italiani che migliorano in graduatoria nazionale escludendo le autocitazioni è sensibilmente più alto di coloro che peggiorano, mostrando il rapporto più alto dell'intera base statistica (1.65 vs. ~1.55 medio). Per capire se questo dato sia significativo, l'abbiamo confrontato con il corrispettivo risultato per l'Italia in base WRD.

Tabella 4: Analisi del q_r su base nazionale. Notare, sia il rapporto tra migliorati e peggiorati in classifica una volta escluse le autocitazioni che in Italia è il più alto tra i paesi considerati.

Analisi del rapporto di ranking	ITA (2303)	DEU (5225)	JPN (2674)	FRA (2193)	GBR (9467)	USA (42455)
migliorati in classifica ($q_r > 0$)	1459	3108	1598	1297	5826	26080
peggioramenti in classifica ($q_r < 0$)	884	2100	1056	869	3610	16351
rapporto(migliorati/peggiorati)	1.65	1.48	1.51	1.49	1.61	1.6
neutro ($q_r = 0$)	10	17	20	27	31	24
media dei miglioramenti	7.80%	7.10%	6.40%	6.10%	5.40%	5.10%
media dei peggioramenti	-19.50%	-13.20%	-11.70%	-11.10%	-10.40%	-9.70%
media pesata dei miglioramenti	5.00%	4.20%	3.80%	3.60%	3.30%	3.10%
media pesata dei peggioramenti	-7.50%	-5.30%	-4.60%	-4.40%	-4.00%	-3.70%

2.2 Analisi dell'Italia su base globale

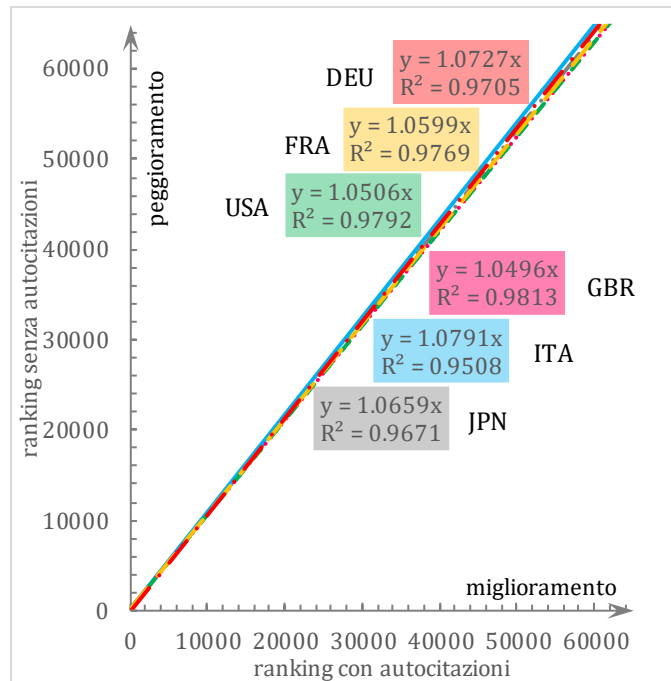


Figura 3: analisi dell'andamento della posizione in classifica su base globale dei soli dati con $\rho_r < 0$, corrispondenti ad un peggioramento in classifica. Si nota che siano tutti vicini, ma l'Italia ha un peggioramento leggermente più accentuato, ma solo dello 0.5%.

Per le analisi su base “globale” abbiamo seguito un procedimento simile a quello adottato su base nazionale unendo i campioni delle singole nazioni in un'unica base statistica WRD. La nuova tabella di posizioni a livello globale per gli scienziati italiani è stata denominata ITW. I risultati ottenuti non si discostano significativamente dai risultati su base nazionale, e questo era prevedibile, vista somiglianza tra le varie nazioni e della distribuzione dei rapporti di citazione e di ranking, come evidenziato in Tabella 1, Figura 1 e Tabella 3. In Figura 3 riportiamo linea di tendenza che mostra la correlazione tra il ranking dei ricercatori con $\rho_r < 0$ con e senza autocitazioni, evidenziando una leggerissima differenza tra i vari Paesi, quasi impercettibile. È pur vero che l'Italia mostra il dato peggiore, anche in termini di dispersione dei dati, ma parliamo di una differenza tra 1.079 e 1.072 ossia dello 0.55%. Infine, osserviamo in Figura 4 che i dati su scala nazionale sono rappresentativi anche dei dati su scala globale, in quanto l'andamento della posizione normalizzata in base cento al variare di ρ_c per la base ITA e per la base ITW sono praticamente sovrapponibili.

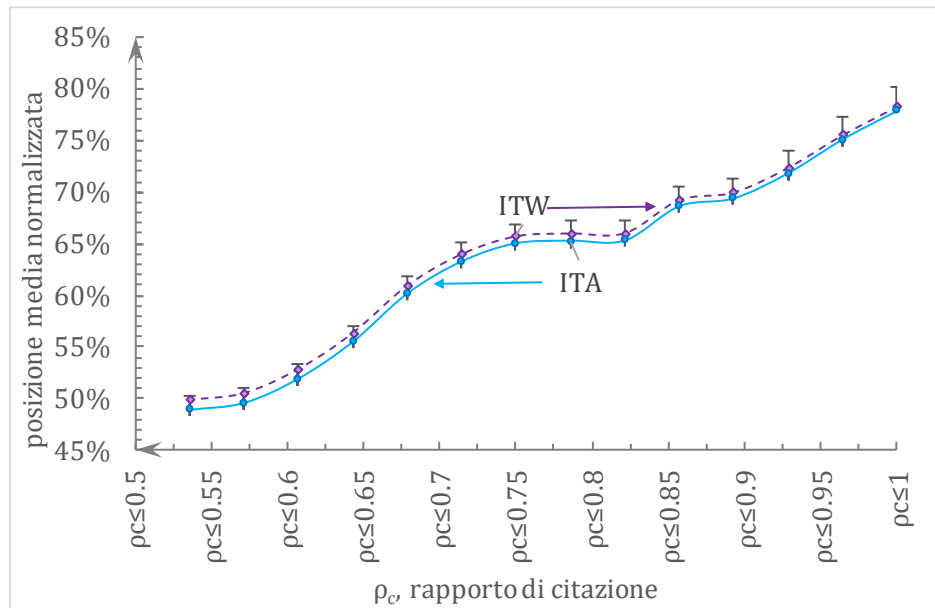


Figura 4: analisi dell'andamento della posizione in classifica su base nazionale. Si nota come i dati con $\rho_r < 0$ (grigio), corrispondenti ad un peggioramento. Le due curve si trovano sempre all'interno delle error bands del 3%

3 Discussione

3.1 I dati carriera

Con lo scopo di sostanziare quanto trovato e confermare che non c'è tendenza nel tempo ad aumentare la autoreferenzialità tra gli scienziati più performanti di Italia, è necessaria anche un'analisi comparativa database “carriera” di Ioannidis e cuautori. Se la distribuzione di q_c tende a crescere “nel tempo”, quello che abbiamo indicato come “effetto Baccini” è confermato, altrimenti è confutato. Pertanto, la stessa procedura usata per i dati “2017” è stata adottata per i dati carriera che presumibilmente contiene una certa isteresi della *inwardness*, e i risultati sono mostrati in Figura 5, e Tabella 6 e 7. Come appare evidente dalle figure, l'indice di autocitazione q_c non mostra nessun incremento allarmante in ITA nel tempo; al contrario, la curva sembra stringersi attorno alla moda. Inoltre, sorprendentemente c'è stato un marcato aumento di autoreferenzialità in FRA, e probabilmente questo dato richiederebbe ulteriore analisi, se a ciò aggiungiamo anche che il WRD 2017 per FRA vede una perdita del 9% su base nazionale rispetto ai dati carriera.

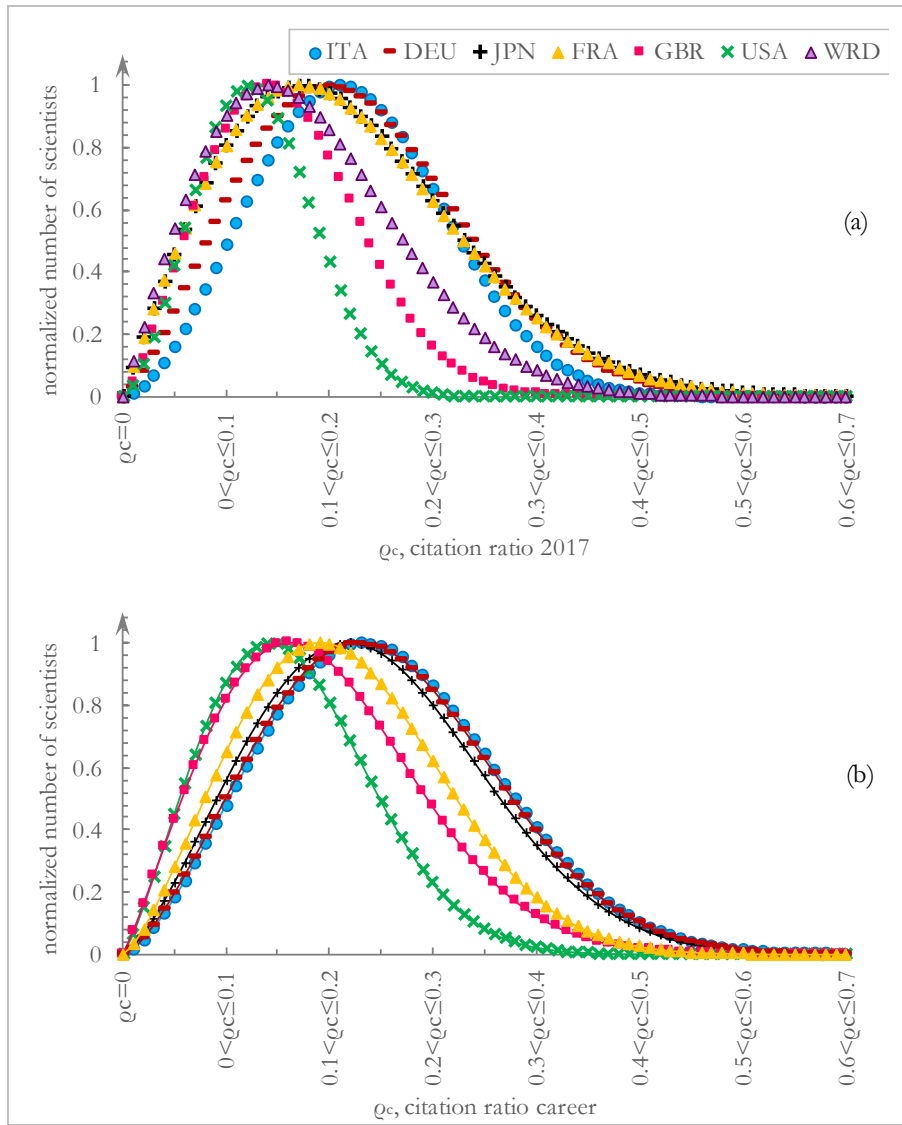


Figura 5: La distribuzione di q_c per ITA, DEU e JPN rimane pressoché invariata tra 2017 (a) e carriera (b), mentre USA e GBR tendono ad autocitarsi più in carriera che nei dati 2017. Al contrario, l'aumento più marcato di autoreferenzialità si registra in FRA

Tabella 5: andamento del rapporto di citazione a livello nazionale per il database carriera. Le figure non sono truncate per $q_c \geq 1$, dato che (ovviamente) nessuno dei top 100,000 scienziati al mondo cita se stesso più di quanto non venga citato

Career	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_c=0$	0	2	1	0	5	32	40
$0 < q_c \leq 0.1$	224	681	469	437	2879	17604	22294
$0.1 < q_c \leq 0.2$	871	2172	1517	1171	4392	19507	29630
$0.2 < q_c \leq 0.3$	543	1460	789	502	1673	5872	10839
$0.3 < q_c \leq 0.4$	214	576	308	162	487	1752	3499
$0.4 < q_c \leq 0.5$	92	265	134	62	168	654	1375
$0.5 < q_c \leq 0.6$	51	136	68	38	76	302	671
$0.6 < q_c \leq 0.7$	24	77	38	19	44	134	336
$0.7 < q_c \leq 0.8$	22	43	25	9	20	61	180
$0.8 < q_c \leq 0.9$	7	18	11	9	8	33	86
$0.9 < q_c \leq 1$	2	14	3	0	10	18	47

Tabella 6: andamento del rapporto di ranking su base nazionale per il database carriera. Nuovamente, tutte le distribuzioni sono centrate attorno a $0 \leq q_r \leq 0.2$ come per il database 2017

Career	ITA	DEU	JPN	FRA	GBR	USA	WRD
$q_r \leq -1$	13	19	5	8	23	41	110
$-1 < q_r \leq 0.8$	6	10	5	7	15	44	84
$-0.8 < q_r \leq -0.6$	6	26	10	8	17	101	184
$-0.6 < q_r \leq -0.4$	18	58	38	14	54	228	429
$-0.4 < q_r \leq -0.2$	73	238	127	82	286	1164	2131
$-0.2 < q_r \leq 0$	745	1924	1230	878	3661	17435	25408
$0 < q_r \leq 0.2$	1174	3108	1939	1401	5676	26919	40523
$0.2 < q_r \leq 0.4$	28	75	27	17	48	81	243
$0.4 < q_r \leq 0.6$	1	1	0	0	0	0	0
$0.6 < q_r \leq 0.8$	0	0	0	0	0	0	0
$0.8 < q_r \leq 1$	0	0	0	0	0	0	0

Alla luce del fatto che l'effetto delle autocitazioni sugli scienziati altamente citati è trascurabile e con lo scopo di rimarcare lo stato di salute complessivo del sistema della ricerca italiana e l'effetto probabilmente positivo delle valutazioni ANVUR, mostriamo come è aumentata la presenza nazionale ITA nella classifica di Ioannidis e coautori confrontando il quantitativo di presenze sia in termini di incrementi assoluti. Il dato importante che salta all'occhio è che tutte le nazioni a parte l'Italia hanno subito una flessione, anche sostanziale (−20 % su scala nazionale JPN) del numero di campioni presenti nei set di dati 2017 rispetto ai dati carriera. L'Italia invece mostra un +0.2 % su base globale ed un impressionante +11.53 % su base nazionale.

Tabella 7: analisi relativa ed assoluta delle presenze per nazione all'interno del database carriera e 2017.

Presenza in classifica		ITA	DEU	JPN	FRA	GBR	USA
presenza assoluta	dati carriera	↓ 2065	↑ 5459	↑ 3382	↑ 2415	↑ 9780	↑ 46013
	dati 2017	↑ 2303	↓ 5225	↓ 2674	↓ 2193	↓ 9647	↓ 42455
Presenza percentuale	dati carriera	↓ 1.97%	↑ 5.20%	↑ 3.22%	↑ 2.30%	↑ 9.31%	↑ 43.81%
	dati 2017	↑ 2.17%	↓ 4.91%	↓ 2.51%	↓ 2.06%	↓ 9.07%	↓ 39.91%
	Δ nazionale	11.53%	-4.29%	-20.93%	-9.19%	-1.36%	-7.73%
	Δ globale	0.20%	-0.29%	-0.71%	-0.24%	-0.24%	-3.90%

4 Conclusioni

L'effetto sostenuto da Baccini et al. che a causa delle valutazioni bibliometriche di ANVUR dei singoli, dei dipartimenti, e delle Università, in Italia si sia assistito ad una impennata delle autocitazioni negli ultimi 10 anni, è stato smentito relativamente ai dati dei centomila scienziati più “altamente citati”, incluso circa duemila italiani, essendo estremamente limitato. Infatti, abbiamo riscontrato che solo chi si trova in una posizione bassa a livello nazionale ha una leggera tendenza ad autocitarsi molto. Consigliamo pertanto ad ANVUR di utilizzare nelle prossime valutazioni l'esclusione delle autocitazioni, peraltro molto semplice da eseguire in tutti i database. In ogni caso, paragonando i dati “del database carriera” di Ioannidis et al con quello relativo agli “scienziati giovani” (database 2017), abbiamo riscontrato un dato molto rilevante: tutte le nazioni esaminate a parte l'Italia hanno sofferto un declino, anche sostanziale (−20% su base nazionale JPN) del numero di scienziati presenti nel data set 2017 paragonato al dato carriera. L'Italia invece mostra un +0.2% su base globale ed un impressionante +11.53% su base nazionale.

Riferimenti

- [1] A. Baccini, G. De Nicolao, and E. Petrovich, ‘Citation gaming induced by bibliometric evaluation: A country-level comparative analysis’, *PLoS ONE*, vol. 14, no. 9, Sep. 2019.
- [2] Nature, Seven days: 6–12 December 2013, <https://www.nature.com/news/seven-days-6-12-december-2013-1.14335>
- [3] BIS, U. K. International Comparative Performance of the UK Research Base–2016. 2016, <https://www.elsevier.com/research-intelligence/research-initiatives/beis2016>
- [4] Nature, Editorial, ‘Memo to Italy’s president: your researchers need you’, 27 August 2019, <https://www.nature.com/articles/d41586-019-02560-1>
- [5] J. P. Ioannidis, J. Baas, R. Klavans, and K. W. Boyack, ‘A standardized citation metrics author database annotated for scientific field’, *PLoS biology*, vol. 17, no. 8, p. e3000384, 2019.