A walk on the wild side: an investigation into the quantity and quality of 'predatory' publications in Italian academia

Manuel Bagues [*] Mauro Sylos-Labini [†] Natalia Zinovyeva [‡]

December 31, 2016

Abstract

In recent years the academic world has experienced a mushrooming of journals that falsely pretend to provide peer review. We study the quantity and quality of publications in dubious journals using information from the CVs of 46,000 researchers seeking promotion in Italian academia. We find that about 5% of these researchers have published in journals included in the blacklist of 'potential, possible, or probable predatory journals' elaborated by the scholarly librarian Jeffrey Beall. To better understand the quality of these publications and the motivations of authors, we collected bibliometric information and we conducted a survey among one thousand researchers (response rate=54%). According to respondents, at least a third of these journals either did not offer regular peer review or engaged in some type of irregular editorial practice. The proportion of journals with reported malpractices is similar among journals from Beall's list that are indexed in Scopus. On the other hand, we also find evidence suggesting that some journals identified by Beall may be legitimate. Overall, our results indicate that the use of white and black lists in research evaluations needs to be complemented with expert evaluations.

Keywords: predatory journals, Italian academia

JEL Classification: I23, J45.

[*]Aalto University and IZA, Helsinki, Finland; email: manuel.bagues@aalto.fi
[†]University of Pisa, Italy; email: mauro.syloslabini@unipi.it
[‡]Aalto University, Helsinki, Finland; email: natalia.zinovyeva@aalto.fi

1 Introduction

The referee system occupies a central place in the reward structure in science. Its limitations and inefficiencies notwithstanding, it provides an institutional basis for the reliability and accumulation of knowledge (Zuckerman and Merton, 1971; Dasgupta and David 1994; Csiszar, 2016). Lately, its functioning has been threatened by the emergence of a new industry of journals that falsely claim to offer peer review (Butler, 2013). The extent of this phenomenon has been studied by the scholarly librarian Jeffrey Beall, who catalogues journals and publishers that allegedly engage in fraudulent practices in a list of 'potential, possible, or probable predatory journals'. Beall argues that these journals accept submitted papers without conducting any proper peer-review process and they engage in a number of deceptive and dishonest practices in order to provide the appearance of being a legitimate academic journal. The reliability of Beall's list was largely confirmed by an experiment conducted in 2013 (Bohannon, 2013). The author submitted a fake medical paper with easily detectable flaws to around 300 open-access journals, including a hundred journals from Beall's list. While these journals claimed officially to have a referee process, 84% of them automatically accepted the paper without questioning its content or providing any feedback.

The number of predatory journals seems to be growing exponentially. Shen and Björk (2015) estimate that the number of articles published in journals included in Beall's list grew from 53,000 articles in 2010 to 420,000 in 2014. The rise and ultimate success of 'predatory' publishers rests on a combination of (at least) two factors. First, the open-access model, where journals are funded using publication fees, coupled with ICT technologies makes viable and relatively cheap for dishonest publishers to run counterfeit scholarly journals for profits. Second, the pressure for academic recognition, sometimes called "publish or perish", may push some naive or unethical scientists to publish their research in dodgy journals, specially in institutional contexts that lack an efficient system of academic evaluation. When most of the emphasis is placed on quantity rather than quality, some researchers may exploit the existence of predatory journals to fatten their CVs. To make matters worse, as we show below in more

detail, some predatory journals have managed to be included in indexes such as Scopus, which are used in many countries as *de facto* whitelists. Academic evaluations that automatically give credit to publications in journals included in these lists may give unfair advantages to predatory authors.

In this paper, we provide the first analysis on the quantity and quality of predatory publications at the country level using information from one the world's biggest producer of high-quality research, Italy.[1] Our database includes the CVs of more than 46,000 researchers seeking promotion in Italian academia. The sample accounts for around 61% of assistant professors and 60% of associate professors in the country.[2] According to our findings, about 5% of these researchers have published at least once in a journal covered by Beall's list. Predatory publications are relatively more frequent in departments with lower research quality, in universities located in Southern Italy, and in engineering, economics and business. We also observe that novel researchers are significantly more likely to engage in predatory publishing.

To obtain a better understanding of the quality of journals included in Beall's list where Italian researchers have published, we conducted a survey where we asked a random sample of approximately 1,000 authors about their experiences with the journals, including whether they received referee reports or whether they have some concerns about the integrity of the journals. Around 54% of researchers replied providing information on the editorial functioning of 268 journals. According to our respondents, around one third of these journals did not provide referee reports during the revision process or had an editorial behavior that cast doubts on their integrity. Given the possibility of misreporting, we interpret this figure as a lower bound of the extend of fraud within this set of journals. Results are very similar for the subsample of 74 journals that are indexed in Scopus.

We have also collected bibliometric information from Google Scholar. The academic

[1] According to the Scimago Country Ranking, Italy is the 7th country with the largest h-index (see http://www.scimagojr.com/countryrank.php?order=h&ord=desc. Similarly, King (2004) argues that Italy belongs to the scientific 'premier league'.

[2] Source: Our own calculations using information from the Italian Ministry of Education on the identity of all assistant (*Ricercatori*) and associate professors (*Associati*) in Italy on December 31 2012.

impact of these journals tends to be low. Only 38% of journals have published in the previous 5 years at least 5 articles that received at least 5 citations, and about a third of articles published by Italian authors have received no citations. Yet, there are also a few highly cited articles.

The proliferation of predatory journals in Italy may be a worrying symptom of the pitfalls of performance-based research funding systems that rely on objective measures (Butler, 2003; Hicks, 2012; Bertocchi et al., 2015). An excessive emphasis on the quantity of publications at the expense of quality may provide incentives to publish in questionable journals. Our results also cast doubts on the mechanic use of white- and black- lists in research evaluations, suggesting the need to complement journal lists with expert evaluations.

2 Background information

2.1 Beall's list

Spurred by the proliferation of questionable journals, the academic librarian Jeffrey Beall created in 2010 a 'black' list of suspicious outlets. This list is intended to help scientists and scientific evaluators to identify journals that claim to be academic but do not provide the editorial and publishing services associated with legitimate journals. The list consists of two distinct subsets: the 'list of publishers' and the 'list of stand alone journals'. In both cases, inclusion rests on a set of criteria related to the identity and behavior of the editors and staff, the business management of the publisher, and the overall integrity policy of the publisher.[3] For instance, many of these journals do not conduct any peer-review, they publicize fake impact factors and editorial boards, or falsely claim a nonexistent association with an academic institution or geographic location. In 2012, the black list included 143 standalone journals and about 269 publishers; in 2014, it included 468 standalone journals and 667 publishers, and by the end of 2016 there were 1260 standalone journals and 1155 publishers.

[3]The list can be accessed at https://scholarlyoa.com. The complete description of criteria can be found on line at https://scholarlyoa.files.wordpress.com/2015/01/criteria-2015.pdf.

Beall's list has been subject to some controversy. While in the experiment conducted by Bohannon (2013) a fake article was accepted by approximately 84% of the journals in Beall's list that were targeted, the remaining 16% did reject it, suggesting that maybe some of these journals do not actually qualify as predatory. As Beall himself acknowledges, in some instances publishers and journals may change their policies for the better or for the worse and, as a result, the status of each outlet may not always be up to date. Moreover, sometimes it might be difficult to distinguish between a poorly managed journal and an illegitimate one.[4]

2.2 Citation indexes

A common (and sometimes controversial) way to evaluate research production relies on the use of citation indexes, which allow to rank journals according to some bibliometric indicator. The most used ones are those provided by Scopus and Web of Science. These lists include journals that allegedly satisfy some minimum requirements. For instance, Scopus, the most comprehensive of the two, claims to include only journals that (i) provide-peer review, (ii) publish issues on a regular basis, (iii) satisfy a minimum level of relevancy and readability for an international audience, and (iv) have an ethics and malpractice statement (Rew 2015).[5] Some evaluation systems consider the inclusion in Scopus and Web of Science as a mark for quality and use the set of indexed journals as a *de facto* whitelist.[6]

One potential advantage of using citation indexes is they may be less prone to conflict of interests and they are less expensive than expert evaluations (Régibeau and Rockett, 2006; Bertocchi et al., 2015). On the flip side, they may provide an inaccurate measure of quality and they may encourage agents to game the incentive system.[7] As

[4]For instance, publishers such as MDPI and Hindawi have been removed from the list after a successful appeal.

[5]Scopus has a broader coverage of the scientific literature than Web of Science. The selection criteria of Web of Science and Scopus are explained in more detail respectively at http://wokinfo.com/essays/journal-selection-process/ and https://www.elsevier.com/solutions/scopus/content/content-policy-and-selection

[6]For instance, this is the case in Italy (National Quality Assessment and National Scientific Qualification), the Russian Federation (The National Excellence Initiatives and the Russian Science Foundation) and Spain (Acreditación).

[7]Moed (2006) and Vinkler (2010) offer systematic overviews of bibliometric indicators for research

Holmstrom and Milgrom (1991) point out, when the 'true output' cannot be easily measured, the reliance on objective performance measures may generate distortions. If authors anticipate the metrics that will be used to evaluate them, a problem of goal displacement may arise whereas scoring high on performance measures becomes a goal in itself, rather than a means of measuring whether a certain performance level has been attained (Wouters et al., 2015). For example, Butler (2003) finds that, in Australia, a greater reliance on journals indexed in ISI (today known as Web of Science) was followed by an increase of quantity but not quality of research. Similarly, Moosa (2016) shows that the adoption of a bucket classification system encourages researchers to submit their research to the lowest-quality journal within each bucket. Moreover, some of the journals included in these indexes, particularly in Scopus, may have limited scientific value and, allegedly, they may not even be legitimate academic outlets (Sterligov and Savina, 2016).

More recently, some agencies are also using the information provided by Google Scholar (Bertocchi et al., 2015). In this database, citations are computed and updated automatically and reflect all documents accessible to Google search robots that are formatted in ways that make it possible for their indexing algorithms to identify their bibliographic data or references. Hence, compared to other citations databases, it is the most comprehensive (and the less selective) one.

2.3 Scientific evaluations in Italy

Until recently, scientific evaluations in Italy did not rely explicitly on journal lists or bibliometric data.[8] Two major innovations have changed the evaluation landscape. The first concerns universities. Since 2011, universities are periodically evaluated on the quality of their research production, and the outcome of this assessment has an impact on their funding. This evaluation combines peer-review of journal articles with

evaluation. More generally, Gibbons (1998) provides an overview of the economic literature on objective performance measures.

[8]Some authors have also argued that publications played a relatively small role in promotion decisions. For instance, Perotti (2002) shows that in economics the number of referred publication is positively correlated with promotion, but the correlation is low and not statistically significant.

the use of bibliometric indicators (Rebora and Turri, 2015).[9] The second innovation concerns the tenure process. Since 2012 all promotions in Italian universities are decided within a two-stage evaluation system.[10] In the first stage, candidates to associate and full professorships are required to qualify in a national-level evaluation known as the National Scientific Qualification (NSQ) (*Abilitazione Scientifica Nazionale*). Evaluations are conducted separately by 184 committees, one for each scientific field defined by the Ministry of Education. Qualified candidates can participate in the second stage, which is managed locally by each university.

The first edition of the NSQ took place between 2012 and 2014.[11] At the beginning of the process an evaluation agency appointed by the Ministry collected information on the scientific production of all candidates and provided it to the evaluation committees. In scientific areas, this information is based on the number of articles published in journals indexed by Scopus or Web of Science. In Social Sciences and Humanities, it relies, among other sources, on a list of journals compiled by local experts. Committees were suggested to take this information into consideration.[12]

3 Empirical analysis

We examine the publication record of Italian researchers using information from a large sample of authors. We identify the number of publications covered by Web of Science and Scopus as in 2012, and also the number of publications included in the blacklist created by Beall.[13] Then, to better understand the quality of publications in Beall's list journals, we report the results from a survey conducted among a random sample

[9] In 2003, the first edition of this evaluation, known as Triennial Research Assessment (VTR), relied only on peer-review.

[10] Law number 240/2010, also known as "Gelmini reform" after the name of the minister of Education. Bagues, Sylos-Labini and Zinovyeva (2016, 2017) provide detailed information about this process.

[11] Official documents regulating the process are available at `http://abilitazione.miur.it/public/index.php?lang=eng`, retrieved on February 2016. A detailed description of the system can be also found in Bagues, Sylos-Labini, Zinovyeva (2017).

[12] While in the first edition of the NSQ the bibliometric information was just a nudge, in the second edition celebrated in 2016 it became a strictly necessary condition to be evaluated.

[13] We collected this information in January 2015. The list combines journals listed in Beall's list of 'potential, possible, or probable predatory standalone journals' and journals published by publishers included in Beall's list of 'potential, possible, or probable predatory publishers'.

of authors, complemented with bibliometric information on their publications.

3.1 Publication record of participants in the NSQ

We collected the publication records of all Italian researchers who participated in the 1st wave of the NSQ.[14] The dataset includes information on 46,244 researchers. This accounts for around 61% of assistant professors and 60% of associate professors in Italy.[15]

Applicants' CVs provide a rich amount of demographic and academic information. As shown in Table 1, 40% of researchers are women and, on average, researchers have 16 years of experience measured as time elapsed since their first publication. Most researchers hold a permanent position in an Italian university, either as assistant professors (*ricercatori*) (64%) or as associate professors (*professori associati*) (36%). About a third of candidates have a fixed-term labor contract or they are not affiliated to an Italian university. Approximately half of the researchers working in Italy are based in the North, one quarter are based in the central Italy, and another quarter is based in the South.[16]

The CVs also provide information on researchers' publications between 2002 and 2012. During this 10-year period the average applicant has published 45 items. Out of these 45 items, 24 (53%) are journal articles, 8 (18%) are conference proceedings, 8 (18%) are books and books chapters, and 5 (11%) correspond to other types of publications such as an abstract in a conference, a database, a translation, a comment on a court sentence. The type of publications varies significantly across different disciplines. In sciences and medical disciplines, journal articles are the main type of academic communication. In engineering, conference proceedings are more popular than journal articles. In economics and business and, even to a larger extent, in social sciences and

[14]We downloaded in January 2014 the CVs of all applicants from the official page of the *National Scientific Qualification* (http://abilitazione.miur.it/public/index.php?lang=eng.

[15]Source: Our own calculations using information from the Italian Ministry of Education on the identity of all assistant (*ricercatori*) and associate professors (*associati*) in Italy on December 31 2012.

[16]Southern regions refer to Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia; central regions include Lazio, Marche, Toscana and Umbria; and northern regions are Emilia Romagna, Piemonte, Lombardia, Valle d'Aosta, Veneto, Trentino-Alto Adige, Friuli-Venezia Giulia and Liguria

humanities, many academics also write books or chapters of books.

Within the set of articles published in journals, the large majority (75%) were published in outlets indexed by Scopus or Web of Science. The proportion is significantly lower in social sciences and humanities, probably reflecting that in these areas only 10% of articles are written in English, compared to 60% in economics and business and around 90% in Science, Technology, Engineering, Math, and Medicine (STEM&M).[17]

To proxy for the quality of articles we consider two different measures. In STEM&M fields, we ranked journals in each field by their Articles Influence Score (AIS)[18] and using this information we calculated the number of articles that each author has published in the top quartile (Q1) journals. In economics and business as well as in social sciences and humanities, we measure the quality of journals using the list of high-impact journals prepared by the Italian evaluation agency (so-called 'A-journals').

In STEM&M fields about 45% of articles are in top 25% of journals in the Web of Knowledge. In economics and business, about 20% of articles are A-journals. About 30% of articles in the fields in social sciences and humanities are in A-journals.

We also examine how many articles are published in journals included in Beall's list.[19] Out of the 1.8M articles published by our sample of researchers, 5,798 were published in a predatory journal. Overall, 0.5% of articles of an average researcher are published in Beall's list journals (see Table 1, column 1).[20] If we examine the information at the author level, there are 2,225 researchers (about 5% of the entire population) in the sample with at least one article published in a journal of the Beall's list (Table 1, column 1).

Since Beall typically monitors English-language journals, it is informative to calculate the proportion of English-language articles by Italian authors (79% of the total)

[17] We identified the language of each article based on the language used in the title.

[18] Article Influence Score is similar to the journal 5-year Impact Factor, but (i) it weights citations by the quality of the citing journal and by the inverse of the number of references in citing journal and (ii) it excludes self-citations. It is available only for journals indexed in Web of Science.

[19] We collected information on Beall's list journals in January 2015.

[20] In order to identify how many articles researchers have published in journals included in Beall's list of 'potential, possible, or probable predatory journals', we match the information provided by researchers on the name and the ISSN code of journals where they have published with the ISSN code of journals included in Beall's list.

that have been published in Beall's list journals. In this case the share of predatory articles is slightly larger, around 0.9% of all articles published in English.

The propensity to publish in predatory journals differs substantially across fields. In sciences and in medical sciences, only 0.4% of articles in English were published in a predatory journal. The figure is larger in social sciences and humanities (0.7%) and in Engineering (1.6%), and it is the highest in economics and business, where 4.1% of all articles in English were published in a journal classified by Beall as predatory.

Figure 1 also shows how the proportion of publications in predatory journals has evolved over time. Starting in 2010 we observe a sharp increase in the proportion of articles published in Beall's list journals, particularly in economics and business. In the last year of our sample, 2012, over 5% of all articles written in English in economics and business were published in journals included in Beall's list.

In order to describe the profile of a typical candidate who publishes in Beall's list journals, we estimate the following equation:

$$B_{i,e} = \beta_0 + \mathbf{X_{i,e}}\beta_\mathbf{1} + \mu_e + \epsilon_{i,e}, \tag{1}$$

where $B_{i,e}$ is an indicator for candidates who have published in Beall's list journals, $\mathbf{X_{i,e}}$ is a set of individual characteristics and μ_e are discipline times promotion category fixed effects.

Authors with predatory publications are more likely to have a relatively higher publication count, but at the same time with fewer high impact publications (Table 2). They also tend to be less experienced. Women are less likely to have publications in Beall's list than men, but this difference is not statistically significant. Predatory publications are also more common among candidates who are based in Italian universities than among candidates based in foreign universities or candidates with non-academic jobs. Publishing in Beall's list journals is particularly common in departments with relatively poorer research quality and departments located in the South of Italy.[21]

[21]For the subset of researchers based in Italian universities, we measure the research quality of their departments using the score obtained by each department in a national assessment of research quality that was conducted in 2011, based on publications by faculty members between 2004 and

3.2 Journals in Beall's list: bibliometric information

There are 2,225 researchers in our sample who have published at least one article in a journal included in Beall's list. Overall, they have published in 599 different journals, which represents slightly below 10% of the total number of journals included in Beall's list and having an ISSN code (see Table 3, columns 1 and 2). As expected, the sample is positively selected: it includes around half of all journals from Beall's list that have also been indexed by Scopus or Web of Science. Nonetheless, these journals do not tend to be highly cited. In 2012, where the NSQ took place, 132 of these journals were indexed in Scopus and 10 in Web of Science, but only three of them are in the top quartile according to Scopus and none of them is in the top quartile of most cited journals within Web of Science.[22] In Social Sciences and Humanities, we examine the overlap with the list of journals elaborated by the Italian academic authorities for the NSQ. Overall, there are 107 journals that were included in the NSQ white list, although only one journal is included in the list of A-journals. In order to obtain a broader coverage, we have also collected in Fall 2016 information at the journal level from Google Scholar for a (random) sample of these journals. Only 38% of journals in the sample satisfy the criterion for being indexed in Google Scholar: they have an h-index of five or more based on the articles published in the previous five years and comply with simple formatting rules.[23]

We have also used Google Scholar to collect information about the number of citations received by each article.[24] As shown in Figure 2, most of these articles have not attracted much attention from the literature. 23% of articles have not received a single citation (not even from the authors) and the median article has received only 3 cita-

2010. The assessment was organized by ANVUR and it was carried out by independent experts who reviewed a selected number of research products. The resulting score varies between zero (low quality) and one (high quality). According to this metrics, the average researcher is based in a department with score 0.6 (standard deviation is 0.2). More detailed information is available at http://www.anvur.org/rapporto/.

[22] In the case of Web of Science we consider the ranking according to the Article Influence Score; for Scopus we consider the Scimago Journal Rank (SJR).

[23] Google scholar provides information on journals' h5. This index indicates the largest number h such that h articles published in the previous five years have at least h citations each.

[24] This information was collected in September-October 2016. It only covers the (random) subset of articles that was selected for the survey.

tions.[25] If we exclude self-citations, the share of articles that has not received a single citation raise to one third. At the same time, we also observe that there is a group of articles which has received a non-trivial number of citations. The 10% most cited articles in our sample has received at least 20 citations and one article had received 399 citations.

3.3 Journals in Beall's list: survey information

The journals in our sample tend to have a low scientific impact. However, this does not necessarily imply that their editorial practices are fraudulent. In fact, as mentioned earlier, around 16% of journals from Beall's list that participated in Bohannon's (2013) experiment rejected the fake paper sent by the researcher. The proportion of journals that, despite being part of Beall's list, actually follow a legitimate editorial process might be even larger within the sample of journals in our study. In order to obtain information about their editorial practices, we conducted a survey among a random sample of researchers who had published in one of these journals. To select the sample we used a randomized design with stratification by university and field. Our overall population includes 2,225 authors, who were based in 1558 different departments. We randomly assigned half of these departments to participate in the survey and, as a result, we surveyed a sample of 1,088 researchers who were based in 779 different departments. As expected, the characteristics of the target population are statistically similar to the characteristics of the survey sample (see Table 4, columns 1, 2 and 4).

We contacted authors by email between February 2016 and May 2016 and we asked them to complete online a survey. The translated versions of our contact email and questionnaire are reproduced in Appendix A. In our contact email, we point out that one of the articles of the recipient has been published in a journal included in Beall's list, and we brief the recipient about the nature of this list.[26] In our message we

[25] The titles of some of these articles provide some additional hints about their quality. In some cases the grammar is not correct (e.g. "Income Don't Influence Health"), or the paper deals with questions of presumably limited academic interest (e.g. "Influence of Parmigiano Reggiano Diet on Male Sexual Behavior in Rats: Behavioral and Neurochemical Study").

[26] In the case of authors who had more than one publication in a predatory journal, for the purpose of our survey we selected randomly one single publication.

also indicate that Beall's list may have some shortcomings, that our survey aimed at shedding some light on its reliability, and that we would be grateful if the recipient could provide some feedback on the editorial functioning of the journal where her article was published. To minimize the possibility that our message might be perceived as offensive, we also state clearly that it is not our intention to question the integrity or the scientific quality of the recipient. If a given author did not answer, we sent her/him up to three reminders.

Naturally, the interpretation of the survey results is subject to a number of important concerns. A first threat stems from the potential existence of a non-response bias. From a theoretical perspective, the direction of this bias is not clear. Researchers who were aware of the fraudulent nature of the journal where they published may feel too embarrassed or guilty about their past behavior to participate in the survey, generating a 'positive' selection in the sample of respondents. On the other hand, researchers who have a higher opportunity cost of their time may be less likely to reply. If these researchers are less likely to publish in truly predatory journal this might lead to a problem of negative selection.

Another potential source of concern is the existence of a recall bias. We ask researchers information about events that happened several years earlier. Some of them may fail to remember accurately their experience. More worrying is the possibility of a reluctance bias. Researchers may be reluctant to express negative opinions on a journal that accepted one of their papers. There might be also a problem of cognitive dissonance: subjects may report opinions that legitimize their past behaviors and past opinions.

In general, we expect these biases to lead to the underreporting of malpractices and, in this respect, our results should probably be interpreted as a lower bound of the actual degree of fraud. Keeping in mind these potential caveats, below we analyze the results of the survey.

3.3.1 Response rate

Out of the 1,088 authors contacted, 584 completed the survey (response rate = 54%). Due to coauthorships, some respondents provided information about the same article. Overall, respondents provide information on 549 different articles (around 59% of the initial set of articles) and 268 different journals (around 70% of the set of journals initially identified).

The set of authors who participated in the survey is (somehow surprisingly) similar to the target population in terms of the individual characteristics and also of the journals where they have published (see Table 4, columns 2, 3, and 5). The sample of respondents only differs significantly from the target sample in terms of their probability of holding a permanent position in an Italian university (73% in the target sample vs. 77% in the sample of respondents). A potential explanation for this divergence may be related to the validity of the email addresses that we used to contact authors.[27]

3.3.2 Survey answers

Below we describe the main results of the survey. A more detailed summary can be found in Table B1.

One third of respondents (answer that they) learned about the journal from one of their colleagues, 27% reacted to an e-mail sent by the journal, and in 12% of the cases the submission was linked to the participation in a conference. Only 16% of respondents chose the journal because they had previously read some of the articles published there. Some respondents also provide information about publication fees. Around 38% of respondents remembered having paid a fee, which on average amounted to 467 USD.

Authors were asked several questions about their experience with the editorial process. About 8% of respondents who recall the details of the process admit that they did not receive any referee reports and the share more than doubles (22%) when we

[27] We searched contact details based on the information provided in researchers' applications in 2012 for the NSQ evaluation. It is possible that researchers with a permanent position were more likely to use the same email address in 2016 than researchers with fixed-term positions.

consider also those who reveal having received only trivial comments related to the editing (see Table 5, column 1). The figure raises to 26% when we also consider respondents who point out that, during the revision process, they noticed something that made them distrust the integrity of the journal.

Given that some respondents may underreport fraudulent practices but it is unlikely that they overreport them,[28] we construct a variable defined at the journal level that takes value one if at least one author has reported a case of malpractice. According to this metrics, the proportion of fraudulent journals would be equal to 36% (Table 5, column 2).[29]

Around 30 respondents left a comment in the survey describing in more detail why they consider that the editorial practices of the journal were suspicious. As one respondent points out, some journals provide referee reports but nonetheless behave in a predatory way:

> *I was invited to join the editorial board of the journal, and this is why I did not pay to get published. Subsequently, I was asked to serve as referee but I realized that my comments did not have any impact: the papers were published without any improvement. This journal, as many others, do not have a real editor, but a graphical technician who deals with both referees and authors. I then wrote to the editor to resign but nobody even bother to reply.*

Another type of inadequate behavior, revealed by three respondents, is that some journals falsely claimed to be indexed in Scopus or in Web of Science and this information misled them to submit their papers. For instance, one respondent argues that:

> *On their webpage they reported the impact factor of the journal. Unfortunately, I trusted them and did not check it out. Only after they charged me*

[28] Even if errors due to recall bias are always possible, it is unlikely that someone mistakenly remembers (or does not pay attention to) an episode that raises doubts on the quality of her research work

[29] The list of journals, for which at least one author has reported a case of malpractice, is provided in Table B2 in the Appendix.

the publication fees, I realized that the journal was not indexed neither in ISI nor in Scopus.

An additional type of irregularity refers to pricing policies. One author reports that, shortly after the official publication fee was paid, she received a new request for two additional fees: the first for compulsory copy editing and the second for an optional 'fast track' publication required to see her/his article published within a month. Another researcher reports that, as a requirement for publication, she was requested ex-post to attend a conference and pay an extra cost. Interestingly, three authors had the opposite experience. They refused to pay (an unexpected) publication fee but, nonetheless, the journal eventually published their papers. As one of them wrote:

I was asked to pay only after my paper was accepted. I firmly refused, but they published my article anyway.

Another reported malpractice is the publication of articles without the explicit consent of the author. Two respondents argue that their articles were accepted without their knowledge following their attendance to a conference. Similarly, an additional suspicious practice, reported by three authors, is receiving an acceptance decision at an unusually fast pace: either automatic or in less than 48 hours from the submission. Of course, in none of these cases the authors received referee reports.

Some respondents complain that they were instructed by the editor to include among the references at least three articles published in the same journal, irrespectively of their actual content. This practice, known as *coercive citation*, has been also widely documented among more 'standard' academic journals (Wilhite and Fong, 2012).

We also asked researchers what is, according to their experience, the average value given in scientific evaluations to articles published in the corresponding journal. Around 24% of the respondents believe that articles published in the journal tend to receive a high or average valuation in scientific evaluations, 39% think that the value is low and 25% declare that it receives no (or negative) value. Interestingly, when researchers are asked about their own opinion of the journal, around 29% of them think that the

journal should receive more credit than it typically gets, while only 7% believe that it is overvalued.

A number of respondents also expressed their opinion about the quality of the journal in their open comments. Among the 226 authors who sent us comments, around 10% (22 respondents) provided us with positive feedbacks either praising the quality of the revision process or highlighting the academic stand of the editor in charge. For instance, one of the respondents declares that:

> *My experience with [JOURNAL TITLE] was very positive. I had the impression of a very careful and rigorous revision process, comparable to other Journals of the same scientific field. I remember we had two very competent Reviewers who addressed pertinent issues in the paper and helped us to improve our article. To me, this is a "trustable" Journal.*

Similarly, another author argues that:

> *The referees did a very good job. The paper has improved substantially after their comments and suggestions.*

Two authors even sent us the copies of the referee reports they received as a way to demonstrate the quality of the editorial process.

On the other hand, some respondents voice very negative opinions about the journal where they published their work. Two authors put it quite bluntly:

> *I think that the journal should be shut down.*

> *The editor in chief is a crook.*

And another one expresses regret about his experience:

> *In 2011 I participated to a conference they organized. They run several journals and they offered me to publish on a fast track in one of them. (...) I needed a publication for the National Scientific Qualification and I accepted to publish in this journal. Today, I regret that decision.*

In the last question of the survey, we asked researchers if they were familiar with Beall's list. Most of them were unaware of the existence of this 'black' list: only 10% had ever heard about it.

3.4 White- and blacklists: 'It don't matter if you're black or white'

Some countries, including Italy, use citations indexes such as Scopus as a de-facto whitelist. In fact, some respondents argued that the main motivation to submit their paper to that particular journal was the fact that it is included in a whitelist.[30]

However, according to our calculations, almost 300 journals indexed by Scopus have been also classified by Beall as predatory (see Table 3). Some of the open comments that we have received from authors suggest that some of these journals do not qualify as legitimate academic outlets:

> *It is not a publication I am proud of. Indeed, I am a bit ashamed. Let me add that all the journals published by [NAME OF THE PUBLISHER] are just trash. I cannot understand how can they be indexed by Scopus and, thus, count for the Research Quality Assessment.*

We use the evidence provided by the survey, which includes information on 74 Scopus journals, to estimate in a more systematic way the reliability of the editorial practices of Scopus journals that have been included in Beall's list. For the sake of comparison, we also analyze separately the subsample of journals covered by Google Scholar. As shown in Table 5, according to survey respondents at least 27% of Scopus journals exhibit some editorial irregularity. This figure is statistically similar to the share among journals indexed by Google Scholar (29%) and journals that are not indexed neither by Scopus nor by Google.

[30] 21 feedbacks of our respondents (around 10% of total feedbacks) are related to the fact that the journal under scrutiny is indexed in Scopus or Web of Science. Additionally, 13 feedbacks mention at least one of the two Italian national evaluations (VQR or NSQ). Interestingly, in 5 cases respondents acknowledge that the publication was useful for the evaluation. According to one of them, who now deeply regrets her decision, submission was driven by the closing timeline of the evaluation.

We also analyze whether journal ranks in Scopus and Google Scholar are associated with the frequency of fraudulent practices reported by the respondents. We estimate the following equation:

$$P_j = \beta_0 + \mathbf{X_j}\beta_\mathbf{1} + \epsilon_j, \tag{2}$$

where P_j is an indicator for the journals that, according to the information collected in the survey, exhibit predatory practices and $\mathbf{X_j}$ is a measure of the number of citations received by journal j. In particular, we use the Scimago journal rank for journals included in Scopus and the Google Scholar h-index for journals indexed in Google Scholar. As shown in Table 6, within the set of indexed journals there is a significant negative correlation between the journal rank and the probability that the journal was reported as fraudulent.

In sum, within the set of journals included in Beall's list, being indexed by Scopus or Google Scholar does not seem to be associated to better editorial practices. However, for the subset of indexed journals, we do observe a significant correlation between the number of citations received by the journals and the probability that these journals engage in dubious editorial practices.

4 Conclusion

In this paper, we use data from the CVs of a large of sample of Italian researchers in all disciplines to examine the quantity and the quality of publications in so-called 'predatory' journals.

The amount of publications in Beall's list journals is relatively small, but not negligible. In a sample of around 1.8 million publications, we identify approximately 6,000 publications in journals that have been included by Beall in his blacklist. The extent of these publications has been growing in recent years and, among all fields, it is particularly relevant in Economics and Business. In the last year of our sample, 2012, approximately 5% of all articles by Italian economists and management scholars in

English-language journals were published in one of these journals.

In order to understand better whether these journals are legitimate academic journals or not, we conducted a survey among a sample of around 1,100 researchers. Approximately 54% of these researchers replied to the survey providing information on their experience. Our analysis of these replies, combined with bibliometric information, provides four main findings.

First, a significant amount of journals in our sample can be indeed classified as predatory journals. In particular, according to the survey information about a third of journals in the sample incurred in some type of fraudulent editorial practice. Not surprisingly, most journals have a very low scientific impact. According to Google Scholar, only 38% of these journals have an h-index of five or more based on the articles published in the previous five years.

Second, our findings also provide some qualitative information about the motivation of authors who publish in predatory journals. Some authors argue that they were fooled by the journals, for instance due to false claims about their impact factor, and they hope that more research is done to uncover dishonest journals. As one of the authors puts it:

> *I am happy that someone is trying to shed some light on these journals. However, I hope that this is not going to harm those who in good faith devoted their time and money (with the shortage of research fund it is mostly out of the pocket).*

On the other hand, some authors confess that their main motivation was linked to upcoming evaluations where publications in these journals are given credit despite their lack of rigor. Unfortunately, we are unable to determine based on our evidence the share of each group of authors.

The third general result that emerges from our investigation is that not all journals included in the Beall's list are 'dodgy'. According to survey respondents over half of the journals offer meaningful peer review and did not incur in any unusual editorial practice that might be considered predatory. This is consistent with the fact that some

of the articles receive a significant number of citations. For example, among the top cited articles of our sample, there is research cited in articles published in top-journals like Science, PNAS, Nature review cancer, and The Lancet. But even among less cited articles, according the feedback of our respondents, some were peer reviewed for the Italian Research Quality Assessment (VQR) and received the highest mark (excellent).

Fourth, we have detected a significant number of journals indexed by Scopus that exhibit predatory practices. This is particularly worrying given than in many countries, including Italy, publications included in this index sometimes automatically receive credit in scientific evaluations.

Overall, our study casts doubts on the mechanical use of white and black lists in evaluation processes. For instance, the presence within Scopus of some journals which are likely to accept for publication any manuscript that vaguely resembles an academic paper in exchange for a payment strongly advises against the automatic use of this index as a whitelist. Similarly, our findings suggest that not every journal included in Beall's list should be considered as predatory. In general, a more nuanced approach to evaluations may be needed, whereas lists are combined with expert evaluations.

Our survey results also suggest that feedback from authors may be used to detect predatory journals. While surveys may be less precise than experiments that involve sending fake papers to journals (e.g. Bohannon 2013), they allow focussing on a more relevant sample of journals and they also have the advantage of avoiding the ethical concerns associated with imposing a cost on potentially legitimate editorial boards and referees without their consent.

References

Bagues, M., Sylos-Labini, M., Zinovyeva, N., 2017. Does the gender composition of scientific committees matter? American Economic Review, Forthcoming.

Bagues, M., Sylos-Labini, M., Zinovyeva, N., 2016. Connections in Scientific Committees and Applicants' Self-selection: Evidence from a Natural Randomized

Experiment. IZA Discussion Paper No. 9594.

Beall, J., 2012. Predatory publishers are corrupting open access. Nature, 489 (7415), 179-179.

Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C.A., Peracchi, F., 2015. Bibliometric evaluation vs. informed peer review: Evidence from Italy. Research Policy, 44(2), 451-466.

Bertrand, M., Mullainathan, S., 2001. Do People Mean What They Say? Implications for Subjective Survey Data. American Economic Review, 91(2), 67-72.

Bohannon, J., 2013. Who's afraid of peer review? Science, 342, 60-65.

Butler, D., 2013. The dark side of publishing. Nature, 495, 433-435.

Butler, L., 2003. Explaining Australia's increased share of ISI publications: the effects of a funding formula based on publication counts. Research Policy, 32(1), 143-155.

Clark, J., Smith, R., 2015. Firm action needed on predatory journals. BMJ, 350. doi:http://dx.doi.org/10.1136/bmj.h210.

Csiszar, A., 2016. Peer review: Troubled from the start. Nature, 532, 306-308.

Dasgupta, P., David, P. A., 1994. Toward a New Economics of Science. Research Policy, 23, 487-521.

Garfield, E., 1972. Citation Analysis as a Tool in Journal Evaluation Science, 178, 471-479.

Gibbons, R., 1998. Incentives in Organization. Journal of Economic Perspectives, 12(4), 115-132.

Hicks, D., 2012. Performance based university research funding systems. Research Policy, 41(2), 251-261.

Holmstr om, B., Milgrom, P., 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. Journal of Law, Economics, and Organization, 7,24-52.

King, D.A., 2004 The scientific impact of nations. Nature, 430, 311-316.

Moed, H.F., 2005. Citation Analysis in Research Evaluation. Springer.

Moosa, I.A., 2016. A critique of the bucket classification of journals: The ABDC list as an example. Economic Record, 92(298), 448-463.

Perotti, R., 2002. The Italian University System: rules versus incentives. ISAE conference Monitoring Italy.

Rebora, G., Turri, R., 2013. The UK and Italian research assessment exercises face to face. Research Policy, 42(9), 1657-1666.

Régibeau, K., Rocke, E., 2016. Research assessment and recognized excellence: simple bibliometrics for more efficient academic research evaluations. Economic Policy, 31 (88), 607-609.

Rew, D., 2015. An Introduction to the Scopus Content Selection and Advisory Board (CSAB). https://www.elsevier.com/__data/assets/pdf_file/0004/95116/general_introduction_csab.pdf

Seglen, P.O., 1997. Why the impact factor of journals should not be used for evaluating research. BMJ., 14(7079), 498-502.

Shen, C., B. Björk, 2015. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Medicine, 13:230.

Sterligov, I., Savina, T., (2016). Riding with the Metric Tide: 'Predatory' Journals in Scopus. Higher Education in Russia and Beyond, 1(7), 9-12.

Stern, D.I., 2013. Uncertainty Measures for Economics Journal Impact Factors. Journal of Economic Literature, 51(1), 173-89.

Vickler, P., 2010. The Evaluation of Research by Scientometric Indicators. Oxford, Chandos Publishing.

Wilhite, A., Fong, E.A., 2012. Coercive Citation in Academic Publishing. Science, 335 (6068), 542-543.

Wouters, P., Thelwall, M., Kousha, K., Waltman, L., de Rijcke, S., Rushforth, A., Franssen, T., 2015. The Metric Tide: Literature Review. Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management. HEFCE.

Xia J., 2015. Predatory journals and their article publishing charges. Learned Publishing, 28(1), 69-74.

Xia J., Harmon, J.L., Connolly K.G., Donnelly, R.M., Anderson, M.R., Howard H.A., 2015. Who publishes in 'predatory' journals? Journal of the Association for Information Science and Technology, 66(7), July 2015, 1406-1417.

Zuckerman, H., Merton, R.K., 1971. Patterns of evaluation in science: institutionalization, structure and function of the referee system. Minerva, 9, 66-100.

Table 1: Descriptive statistics – Candidates in NSQ

	All		Sciences		Medical Sciences		Engineering		Econ. & Business		Soc.Sc. & Hum.	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics:												
Female	0.4	0.49	0.4	0.49	0.37	0.48	0.22	0.42	0.41	0.49	0.46	0.5
Experience	16	8	16	7	18	8	13	6	12	6	15	8
Permanent position in an Italian university:	0.6	0.49	0.57	0.5	0.59	0.49	0.7	0.46	0.75	0.43	0.57	0.5
- Assistant professor	0.64	0.48	0.66	0.48	0.64	0.48	0.62	0.49	0.61	0.49	0.64	0.48
- Associate professor	0.36	0.48	0.34	0.48	0.36	0.48	0.38	0.49	0.39	0.49	0.36	0.48
University location:												
- North	0.45	0.5	0.45	0.5	0.46	0.5	0.46	0.5	0.48	0.5	0.42	0.49
- Center	0.26	0.44	0.27	0.44	0.28	0.45	0.23	0.42	0.24	0.43	0.27	0.44
- South	0.29	0.45	0.28	0.45	0.27	0.44	0.31	0.46	0.28	0.45	0.31	0.46
Department research score	0.59	0.21	0.7	0.18	0.55	0.23	0.71	0.18	0.35	0.21	0.56	0.16
Publications	45	41	50	44	60	52	58	39	26	21	31	26
Journal articles	24	30	35	38	38	34	21	17	10	11	10	13
- ISI/Scopus articles	18	29	32	36	33	33	17	16	5	7	1	3
- Articles in English	19	29	33	35	33	33	19	17	6	8	1	3
- Q1-journal articles (STEM&M)	-		17	23	15	19	7	9	-		-	
- A-journal articles (SS&H)	-		-		-		-		2	3	3	5
Conference proceedings	8	15	7	14	7	15	30	25	4	7	3	5
Chapters	6	9	2	4	4	6	4	6	7	8	11	12
Books	2	3	0.5	1.6	0.8	2.1	0.9	2.1	2	3	4	4
Other publications	5	15	5	12	10	25	3	8	2	6	3	8
Publications in Beall's list												
Number of Beall's list articles	0.08	0.52	0.07	0.39	0.09	0.48	0.22	1.03	0.21	0.94	0.01	0.015
At least one Beall's list article	0.05	0.21	0.05	0.22	0.06	0.24	0.11	0.31	0.11	0.32	0.01	0.08
% of Beall's list articles in journal articles	0.5	3.5	0.3	1.7	0.3	1.9	1.4	6.4	2.2	8.5	0.1	1.6
% of Beall's list articles in journal articles in English	0.9	6.0	0.3	2.3	0.5	3.3	1.6	7.0	4.1	14.2	0.7	6.9
Observations	46,244		11,953		10,712		4,607		3,256		15,716	

Notes: University location is identified for all applicants with a permanent or temporal position in an Italian university. Department research score is from the 2011 department assessment by ANVUR (*Valutazione della Qualitá della Ricerca*). Q1-journals are journals in the first quartile in the corresponding field in ISI Web of Knowledge in terms of the Article Influence Score. A-journals are high-impact journals in the fields of Economics, Business, Social Sciences and Humanities as defined by ANVUR expert committee. Experience is defined as the number of years since the first publication. Publication data refer to publications between 2002 and 2012 listed in applicants' CVs.

Table 2: Who publishes in Beall's list?

	1	2	3
	All	**Disciplinary group:**	
		STEM&Med	**Soc.Sc.&Hum.**
Female	−0.001	−0.003	0.001
	(0.002)	(0.003)	(0.002)
Experience	−0.004***	−0.004***	−0.004***
	(0.001)	(0.002)	(0.001)
Position (benchmark - non-tenured university position):			
- Assistant Professor, tenured	0.003	0.002	0.002
	(0.003)	(0.005)	(0.003)
- Associate Professor, tenured	0.004	0.002	0.004
	(0.005)	(0.008)	(0.006)
- Abroad or non-university position	−0.021***	−0.027***	−0.020***
	(0.006)	(0.009)	(0.006)
University location:			
- Central Italy	0.006**	0.007	0.004
	(0.003)	(0.004)	(0.003)
- Southern Italy	0.029***	0.039***	0.014***
	(0.003)	(0.004)	(0.003)
University ranking	−0.036***	−0.040***	−0.042***
	(0.007)	(0.010)	(0.008)
Total number of publications	0.018***	0.022***	0.011***
	(0.001)	(0.001)	(0.001)
Proportion of Q1/A-journal articles	−0.011***	−0.014***	−0.005***
	(0.001)	(0.002)	(0.001)
Mean dependent variable	0.048	0.065	0.024
Evaluation panel FE	Yes	Yes	Yes
Adjusted R-Squared	0.062	0.044	0.099
Observations	46,244	27,272	18,972

Note: OLS estimates. Dependent variable is an indicator for authors who have publications in Beall's list journals. All productivity indicators in the prediction model exclude publications in Beall's list. Productivity indicators and experience are normalized to have zero mean and unit standard deviation for all applicants in a given field and category.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Sample selection: Beall's list journals

	1	2	3	4
	All journals (non-missing ISSN)	**Journals where Italians published**	**Journals in the targeted sample**	**Journals with survey info**
Number of journals	7210	599	379	268
Indexed in:				
Scopus	284	131	88	74
- Q1 SJR	6	3	3	3
Web of Science	14	10	6	5
- Q1 AIS	0	0	0	0
NSQ list	273	213	128	102
- A-journal	2	2	2	2
Google Scholar	n.a.	n.a.	143	112
- GS h-index (mean)	n.a.	n.a.	12	12

Notes: Column 1 includes information on all journals from Beall's list with non-missing ISSN codes. *Q1 SJR* stands for the top-quartile journals in Scopus according to Scimago Journal Rank. *Q1 AIS* indicates top-quartile journals is Web of Science according to the Article Influence Score. n.a. - information is not available (not collected).

Table 4: Sample selection: researchers

	1	2	3		4	5
	Authors with Beall's list articles	**Targeted**	**Replied**		**p-value for the t-test of means differences: targeted/not**	**replied/not**
Female	0.36	0.37	0.35		0.487	0.213
Experience	15	16	16		0.163	0.238
Disciplinary group:						
- Sciences	0.27	0.27	0.29		0.780	0.152
- Engineering	0.22	0.22	0.21		0.606	0.788
- Medicine	0.31	0.31	0.28		0.737	0.026
- Economics and Business	0.16	0.16	0.17		0.658	0.235
- Social Sciences and Humanities	0.04	0.04	0.05		0.622	0.749
Permanent position in an Italian university:	0.74	0.73	0.77		0.865	0.003
- Assistant professor	0.62	0.59	0.57		0.019	0.226
- Associate professor	0.38	0.41	0.43		0.019	0.226
University location:						
- North	0.37	0.35	0.36		0.152	0.615
- Center	0.23	0.24	0.25		0.544	0.150
- South	0.41	0.41	0.39		0.376	0.084
All publications	62	63	62		0.775	0.857
Journal articles	32	33	32	0.920	0.324	
Q1 or A-journal articles	9.6	9.4	9.2	0.376	0.714	
Bealls list articles	1.6	1.6	1.6	0.214	0.579	
Characteristics of the Beall's list article inquired about in the survey:						
Year of publication	-	2010	2010		-	0.783
Citations in Google Scholar	-	8.9	9.3		-	0.469
Journal is indexed in Google Scholar	-	0.52	0.54		-	0.109
Journal Google Scholar H-index	-	17	16		-	0.721
Journal is indexed in Scopus	-	0.43	0.45		-	0.201
Journal Scopus SJR score	-	0.55	0.55		-	0.698
Journal is indexed in WoS	-	0.13	0.14		-	0.628
Journal WoS AIS score	-	0.87	0.85		-	0.316
Observations	2225	1088	584			

Table 5: Survey responses on Beall's list journals

	1	2
	% of answers	**% with at least one answer**
	All journals	
No referee report	8	12
No or superficial referee report	22	30
No or superficial referee report, or something odd	26	36
	Not indexed in Scopus and Google Scholar	
No referee report	8	12
No or superficial referee report	22	30
No or superficial referee report, or something odd	26	36
	Indexed in Scopus	
No referee report	9	19
No or superficial referee report	21	33
No or superficial referee report, or something odd	27	40
	Indexed in Google Scholar	
No referee report	8	15
No or superficial referee report	25	38
No or superficial referee report, or something odd	29	44

Notes: Information is on 242 journals, for which at least one respondent could recall whether the publication involved a referee report.

Table 6: What predicts bad practices?

	1	2	3
Sample:	**All journals**	**Indexed in Scopus**	**Indexed in Google Scholar**
Indexed in Scopus	-0.010		
	(0.044)		
Indexed in Google Scholar	0.016		
	(0.044)		
Scopus Journal Rank		-0.119*	
		(0.060)	
Google Scholar h-index			-0.011***
			(0.003)
Constant	0.267***	0.332***	0.452***
	(0.037)	(0.047)	(0.056)
Adjusted R-sq	-0.008	0.042	0.101
N	242	67	108

Notes: OLS estimates. Dependent variable is the proportion of survey answers for a given journal indicating the existence of fraudulent practices. Information is on 242 journals, for which at least one respondent could recall whether the publication involved a referee report. We apply analytic weights for the number of received responses per journal.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Beall's list articles (%)

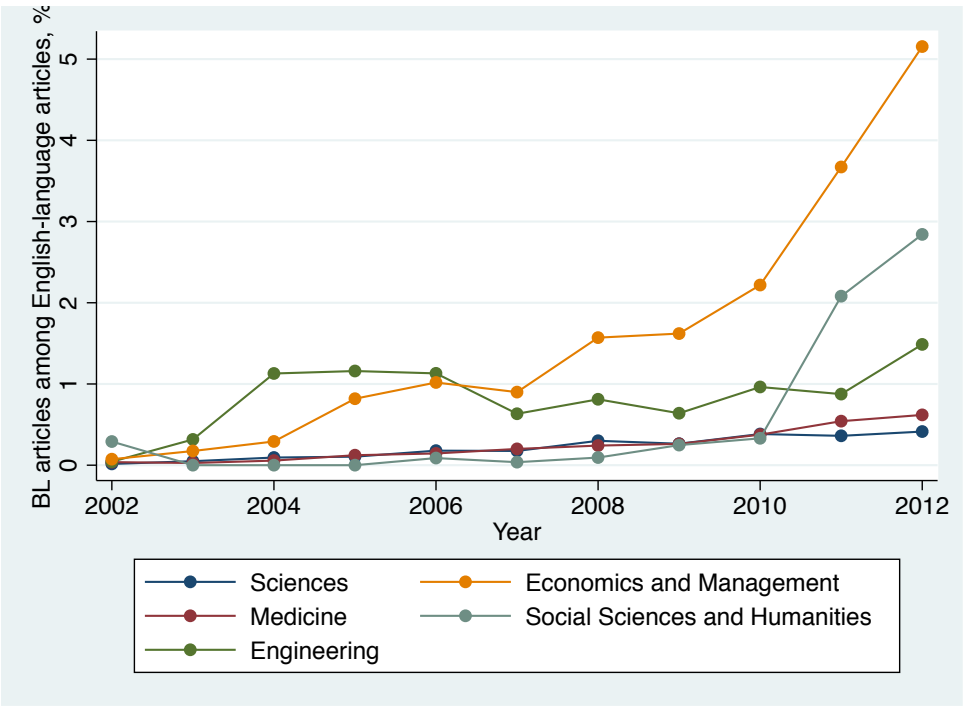
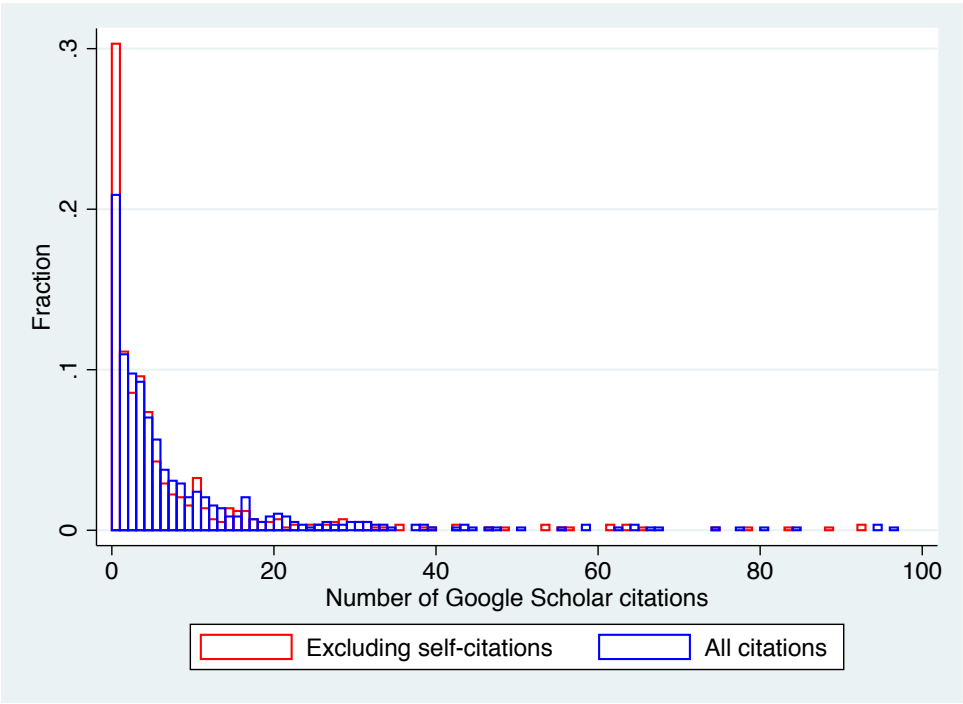


Figure 2: Citations of articles in Beall's list journals



Note: The figure does not include 3 observations with more than 100 citations.

Appendix A. Data.

The data on the participants in Italian evaluations was available at the website of the Italian Ministry of Higher Education and Research during the evaluation process. We extracted all the individual characteristics that we use in the analysis from these CVs. Information on tenured researchers' affiliation and the length of tenure was obtained from the Consortium of Italian universities (CINECA). Affiliation of nontenured researchers is from the most recent publication of the CV.

Candidates were required to report the ISSN code of their journal publications in their submitted CVs. This facilitates the process of merging information in candidates' CVs with various journal lists. In order to identify publications in Beall's list journals, we first collected titles and the ISSN codes of the whole set of journals of each publisher that appeared in Beall's list in January 2015. We also collected the ISSN codes of journals included in the list of standalone predatory journals. We then matched the whole set of articles with candidates' CVs lists using the ISSN code of the journals. To avoid mistakes, we ignored Beall's list journals with no ISSN codes (about 40%). We also exclude from the list the journals run by MDPI, a publisher that in 2015 was removed from the Beall's list after its successful appeal.

We also collected information on bibliometric information from Scopus and Web of Science. First, we check whether journals in candidates' CVs were indexed in Scopus and Web of Science at the moment when candidates submitted their applications in 2012. Second, we consider journals citations ranks based on Scimago Journal Rank (2011) and Article Influence Score (2012). For a subsample of journals we consider whether they are indexed in Google Scholar as in 2016 and their Google Scholar h-index.

Appendix B. Survey email and questionnaire

In this section, we report the original and translated versions of email and questionnaire used in our survey.

ORIGINAL VERSION

Email *Oggetto*: Un suo articolo scientifico pubblicato

Gentile [NOME COGNOME],

la contattiamo in relazione al suo articolo "[TITOLO DELL'ARTICOLO]" pubblicato sulla rivista: [TITOLO DELLA RIVISTA].

La rivista in questione (o la sua casa editrice) è stata inserita dal professor Jeffrey Beall (Università del Colorado) in una lista di editori che, a suo parere, "potenzialmente, possibilmente o probabilmente" non rispettano gli standard scientifici internazionali. Per esempio, secondo la ricerca del professor Beall, alcune riviste pubblicate dalle case editrici della lista, accettano articoli senza un processo di revisione, oppure riportano sulle loro pagine web informazioni non corrette circa il loro impact factor, gli editor o l'indirizzo della loro sede. Può trovare maggiori informazioni sulla lista e sui criteri utilizzati a questo link:
[HYPERLINK ALLA BEALL'S LIST]

La lista del professor Beall ha creato numerose controversie nel mondo accademico e alcune riviste e case editrici inizialmente sospettate di comportamenti non in linea con gli standard scientifici hanno invece dimostrato la loro assoluta correttezza. In un progetto condotto in collaborazione presso l'Università di Pisa e l'Università Aalto (Helsinki, Finlandia), vorremmo approfondire in che misura la lista è affidabile, individuare le sue criticità e, possibilmente, migliorarla. La sua esperienza con la rivista in questione è per noi molto importante.

Le saremmo quindi molto grati se volesse dedicare solo 2 minuti del suo tempo per rispondere a 7 brevi domande relative alla sua esperienza con la rivista [TITOLO DELLA RIVISTA]. Puó farlo semplicemente attraverso questo link:
[HYPERLINK AL QUESTIONARIO]

Ci teniamo a chiarire che, ovviamente, le domande del questionario non hanno come obiettivo mettere in discussione la serietà del suo lavoro né quella di altri articoli pubblicati sulla stessa rivista. L'unico interesse è comprendere il funzionamento editoriale della rivista.

Le risposte saranno usate in forma anonima e solo per scopi di ricerca. Una volta completato il nostro studio, sarà nostra cura, anche per ringraziarla per la sua collaborazione, inviarle un breve report con informazioni relative a tutte le riviste che abbiamo analizzato.

Non esiti a contattarci per qualsiasi informazione aggiuntiva o suggerimento circa il nostro studio.

Cordialmente,

Manuel Bagues, Università di Aalto, Helsinki (email: manuel.bagues@aalto.fi)
Mauro Sylos Labini, Università di Pisa (email: mauro.syloslabini@unipi.it)
Natalia Zinovyeva, Università di Aalto, Helsinki (email: natalia.zinovyeva@aalto.fi)

Questionario

1. Come ha conosciuto l'esistenza della rivista in questione?
    - Attraverso la segnalazione di un collega
    - Con un email/invito della rivista
    - Partecipando ad una conferenza
    - Leggendo articoli pubblicati dalla rivista
    - Altro/Non ricordo
2. Dopo aver inviato il suo articolo, ha ricevuto referee report?
    - No
    - Sì, ho ricevuto referee report superficiali solo relativi alla forma
    - Sì, ho ricevuto referee report anche relativi ai contenuti
    - Non ricordo
3. La pubblicazione del suo articolo è stata a pagamento?
    - Sì, era evidente nelle regole della rivista
    - Sì, l'ho scoperto dopo che l'articolo è stato accettato

- No, non c'è stato alcun costo
- Non ricordo

Se ha pagato per la pubblicazione, può indicarci qual era il prezzo (in $ USA)?

4. Nel processo di revisione e accettazione dell?articolo, c'è stato qualcosa che le ha fatto dubitare dell'integrità/professionalità della rivista?

   - No
   - Sì
   - Non ricordo

   Se ha risposto sì alla domanda precedente, può descrivere brevemente cosa?

5. In base alla sua esperienza professionale, che valore hanno accademicamente le pubblicazioni su questa rivista (per esempio per la Valutazione della Qualità della Ricerca (VQR), l'Abilitazione Scientifica Nazionale (ASN), i fondi di ricerca)?

   - Alto
   - Medio
   - Basso
   - Nullo
   - Negativo

6. In base alla sua opinione, che valore dovrebbero avere accademicamente le pubblicazioni su questa rivista (per esempio per la Valutazione della Qualità della Ricerca (VQR), l'Abilitazione Scientifica Nazionale (ASN), i fondi di ricerca)?

   - Alto
   - Medio

- Basso
- Nullo
- Negativo

7. Prima di ricevere la nostra email, conosceva la lista di Beall?

   - Sì
   - No

   Nel caso volesse aggiungere commenti, suggerimenti o riflessioni, può utilizzare lo spazio sottostante:

TRANSLATED VERSION

Email *Subject*: One of your published articles

Dear [NAME SURNAME],

We contact you regarding your article "[ARTICLE TITLE]" published in [JOURNAL TITLE].

The aforementioned article (or its publisher) had been included by prof. Jeffrey Beall (University of Colorado) in a list of "potentially, possibly, or probably" do not respect international scientific standards. To give you an example, according to prof. Beall, some of the included journals accept articles without a proper refereeing system, or include in their web pages inaccurate information concerning their impact factor, the composition of their editorial board, or their precise location. You can find more information on the list and Beall criteria at this link:
[HYPERLINK TO BEALL'S LIST]

"Beall's list" created a few controversies and some publishers initially included have been removed after a successful appeal. In an on-going research project, our main goal

is to shed some light on the reliability of the list and possibly to improve it. Your experience with the aforementioned journal is very important for our research.

Hence, we shall be pleased if you could devote just 2 minutes of your time to answer 7 short questions concerned with your experience with [JOURNAL TITLE]:
[HYPERLINK TO QUESTIONNAIRE]

We wish to make clear that our goal is not to question the seriousness of your scientific work or of other articles published in the same aforementioned journal. Our only goal is to understand its editorial functioning.

Your answers will be anonymized and employed only for research purposes. After the end of our project, also to thank you for your collaboration, we shall send you a report on the main results.

Yours sincerely,

Manuel Bagues, Aalto University, Helsinki (email: manuel.bagues@aalto.fi)
Mauro Sylos Labini, University of Pisa (email: mauro.syloslabini@unipi.it)
Natalia Zinovyeva, Aalto University, Helsinki (email: natalia.zinovyeva@aalto.fi)

Questionnaire

1. How did you first hear about the aforementioned journal?

    - I received suggestion from a colleague
    - An email/invitation to submit a paper
    - Reading its published articles
    - Attending a conference
    - Other/I do not remember

2. After submitting you article, did you receive referee reports?

    - No, I did not
    - Yes, I did receive trivial referee report only addressing the editing
    - Yes, I did and they also addressed the paper's contents

- I do not remember

3. Did you pay any fee to publish your article?

   - No, I did not
   - Yes, I did and it was clear from the journal guidelines
   - Yes, I did and I found out after the article was accepted for publication
   - I do not remember

   If you did, could you please indicate how much you paid? (in US dollar)

4. During the process of revision/acceptance of your paper, did you notice anything that made you distrust the integrity/professionalism of the journal?

   - No, I did not
   - Yes, I did
   - I do not remember

   In case you did, could you briefly tell us what?

5. According to your past experience, what is the academic value of articles published on this journal?

   - High
   - Average
   - Low
   - None
   - Negative

6. According to your opinion, what should be the academic value of articles published on this journal?

- High
- Average
- Low
- None
- Negative

7. Have you ever heard about "Beall list" before getting the invitation to participate to this survey?

   - Yes
   - No

   In case you would like to add any comment or suggestion you can use the box below

Table B1: Survey responses

	1	2	3
	All	**In Scopus/WoS:**	
		No	**Yes**
How did you first hear about this journal?			
A suggestion from a colleague	33	28	39
An email/invitation to submit a paper	27	30	23
Reading its published articles	16	15	17
Attending a conference	12	13	11
Other / Do not remember	13	15	10
After submitting you article, did you receive referee reports?			
Yes, I did and they also addressed paper's contents	64	65	62
Yes, I did and they only addressed the editing	12	13	10
No, I did not	7	7	8
I do not remember	18	15	20
Did you pay any fee?			
No, I did not	44	45	42
Yes, I did and it was clear from the journal guidelines	30	31	28
Yes, I did and I found out after the article was accepted for publication	8	10	6
I do not remember	18	14	24
Did you find anything odd in the process of revision/acceptance of the paper?			
No, I did not	78	77	80
Yes, I did	10	10	11
I do not remember	12	14	9
According to your experience, what's the academic value of publications in this journal (in formal national research evaluations)?			
High	4	4	3
Average	20	15	26
Low	39	38	41
None	24	30	16
Negative	1	0.3	1.9
I do not know	13	14	12
In your opinion, what should be the academic value of publications in this journal (in formal national research evaluations)?			
High	5	5	5
Average	35	33	37
Low	35	36	34
None	10	11	9
Negative	1.2	0.3	2.3
I do not know	14	15	13
Have you ever heard about Beall's list before?			
No	90	88	93
Yes	10	12	7

Notes: The table reports percentages of non-missing answers to each question.

Table B2: Journals with at least one reported bad practice

Journal title	Replies	Any bad practice, %	Scopus 2012	WoS 2012	NSQ 2012	Google Scholar
APPLIED MATHEMATICS	3	100	0	0	1	0
CHINA-USA BUSINESS REVIEW	2	100	0	0	1	1
ENVIRONMENTAL ECONOMICS	2	100	0	0	1	0
INTERNATIONAL JOURNAL OF CONTEMPORARY MATHEMATICAL SCIENCES	2	100	0	0	1	0
THE OPEN FUELS & ENERGY SCIENCE JOURNAL	2	100	1	0	0	0
THE OPEN PROTEOMICS JOURNAL	2	100	1	0	0	0
ACTA INFORMATICA MEDICA	1	100	1	0	0	1
AFRICAN JOURNAL OF AGRICULTURAL RESEARCH	1	100	1	0	1	1
AMERICAN JOURNAL OF FOOD TECHNOLOGY	1	100	1	0	1	0
ATMOSPHERIC AND CLIMATE SCIENCE	1	100	0	0	0	1
CREATIVE EDUCATION	1	100	0	0	0	1
ENERGY AND ENVIRONMENT RESEARCH	1	100	0	0	0	1
EUROPEAN JOURNAL OF MANAGEMENT (EJM)	1	100	0	0	1	0
EUROPEAN JOURNAL OF SCIENTIFIC RESEARCH	1	100	0	0	1	0
GLOBAL ECONOMY AND FINANCE JOURNAL	1	100	0	0	1	0
INTERNATIONAL JOURNAL OF APPLIED PSYCHOLOGY	1	100	0	0	0	0
INTERNATIONAL JOURNAL OF COMPUTER SCIENCE ISSUES	1	100	0	0	0	1
INTERNATIONAL JOURNAL OF HUMANITIES AND SOCIAL SCIENCE	1	100	0	0	1	1
INTERNATIONAL JOURNAL OF MATERIAL SCIENCE	1	100	0	0	0	0
INTERNATIONAL JOURNAL OF MATHEMATICAL ANALYSIS	1	100	1	0	1	0
INTERNATIONAL JOURNAL ON ADVANCES IN SOFTWARE	1	100	0	0	0	0
JOURNAL OF ANTIVIRALS AND ANTIRETROVIRALS	1	100	1	0	0	0
JOURNAL OF AQUACULTURE RESEARCH & DEVELOPMENT	1	100	0	0	0	1
JOURNAL OF CLINICAL MEDICINE RESEARCH	1	100	0	0	0	1
JOURNAL OF FOOD RESEARCH	1	100	0	0	0	1
JOURNAL OF INTERNATIONAL ENVIRONMENTAL APPLICATION & SCIENCE	1	100	0	0	0	1
JOURNAL OF MATHEMATICAL SCIENCES: ADVANCES AND APPLICATIONS	1	100	0	0	1	0
JOURNAL OF NANOMEDICINE & BIOTHERAPEUTIC DISCOVERY	1	100	0	0	0	0
JOURNAL OF PHARMACY AND NUTRITION SCIENCES	1	100	0	0	0	1
JOURNAL OF TRAUMA & TREATMENT	1	100	0	0	0	0

Continued on next page

Table B2 – Journals with at least one reported bad practice (continued)

Journal title	Replies	Any bad practice, %	Scopus 2012	WoS 2012	NSQ 2012	Google Scholar
NATURAL SCIENCE	1	100	0	0	0	1
OPEN CONFERENCE PROCEEDINGS JOURNAL	1	100	0	0	0	0
RESEARCH IN APPLIED ECONOMICS	1	100	0	0	1	1
SURGICAL SCIENCE	1	100	0	0	0	1
THE OPEN CRYSTALLOGRAPHY JOURNAL	1	100	0	0	0	0
THE OPEN ECONOMICS JOURNAL	1	100	0	0	1	0
THE OPEN FOOD SCIENCE JOURNAL	1	100	0	0	0	0
THE OPEN MEDICAL IMAGING JOURNAL	1	100	0	0	0	0
THE OPEN PSYCHOLOGY JOURNAL	1	100	0	0	0	0
UNIVERSAL JOURNAL OF MANAGEMENT AND SOCIAL SCIENCES	1	100	0	0	1	0
WORLD APPLIED SCIENCES JOURNAL	1	100	1	0	0	1
WSEAS TRANSACTIONS ON BIOLOGY AND BIOMEDICINE	1	100	1	0	0	0
WSEAS TRANSACTIONS ON CIRCUITS AND SYSTEMS	1	100	1	0	0	1
WSEAS TRANSACTIONS ON COMMUNICATIONS	1	100	1	0	0	1
WSEAS TRANSACTIONS ON MATHEMATICS	1	100	1	0	1	1
WSEAS TRANSACTIONS ON SIGNAL PROCESSING	1	100	0	0	1	0
JOURNAL OF ANIMAL AND VETERINARY ADVANCES	8	75	1	1	0	1
INTERNATIONAL MATHEMATICAL FORUM	7	71	0	0	1	0
JOURNAL OF LIFE SCIENCES	3	67	0	0	1	1
THE EUROPEAN JOURNAL OF ECONOMICS, FINANCE AND ADMINISTRATIVE SCIENCES	3	67	1	0	1	0
THE OPEN MECHANICAL ENGINEERING JOURNAL	3	67	1	0	0	1
WSEAS TRANSACTIONS ON COMPUTERS	6	50	1	0	1	1
WSEAS TRANSACTIONS ON INFORMATION SCIENCE AND APPLICATIONS	6	50	1	0	1	1
JOURNAL OF US-CHINA PUBLIC ADMINISTRATION	4	50	0	0	1	1
BRITISH JOURNAL OF MEDICINE AND MEDICAL RESEARCH	2	50	0	0	0	1
INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES	2	50	0	0	1	1
INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY (IJCSNS)	2	50	0	0	1	1
JOURNAL OF KNOWLEDGE MANAGEMENT, ECONOMICS AND INFORMATION TECHNOLOGY	2	50	0	0	1	1
JOURNAL OF MODERN ACCOUNTING AND AUDITING	2	50	0	0	1	1
JOURNAL OF PURE AND APPLIED MATHEMATICS: ADVANCES AND APPLICATIONS	2	50	0	0	1	0
PSYCHOLOGY	2	50	0	0	1	0

Continued on next page

Table B2 – Journals with at least one reported bad practice (continued)

Journal title	Replies	Any bad practice, %	Scopus 2012	WoS 2012	NSQ 2012	Google Scholar
SOCIOLOGY MIND	2	50	0	0	1	1
THE OPEN ZOOLOGY JOURNAL	2	50	0	0	0	0
THEORETICAL ECONOMICS LETTERS	2	50	0	0	1	0
US-CHINA FOREIGN LANGUAGE	2	50	0	0	1	1
APPLIED MATHEMATICAL SCIENCES	9	44	1	0	1	0
INTERNATIONAL JOURNAL OF ENGINEERING BUSINESS MANAGEMENT	5	40	1	0	1	1
WSEAS TRANSACTIONS ON SYSTEMS	8	38	1	0	1	0
INTERNATIONAL JOURNAL OF PURE AND APPLIED MATHEMATICS	12	33	1	0	1	1
HEAD AND NECK ONCOLOGY	3	33	1	0	0	0
INTERNATIONAL JOURNAL OF BUSINESS AND MANAGEMENT	3	33	0	0	1	1
JOURNAL OF BUSINESS AND POLICY RESEARCH	3	33	0	0	1	0
JOURNAL OF COMMUNICATIONS (JCM)	3	33	0	0	0	1
JOURNAL OF EARTH SCIENCE AND ENGINEERING	3	33	0	0	0	1
JOURNAL OF ENVIRONMENTAL SCIENCE AND ENGINEERING	3	33	0	0	1	1
JOURNAL OF INTELLIGENT LEARNING SYSTEMS AND APPLICATIONS	3	33	0	0	1	1
MEDITERRANEAN JOURNAL OF SOCIAL SCIENCES	3	33	0	0	1	1
NANOMATERIALS AND NANOTECHNOLOGY	3	33	0	0	0	0
CHINESE BUSINESS REVIEW	4	25	0	0	1	1
INVESTMENT MANAGEMENT & FINANCIAL INNOVATIONS	4	25	1	0	1	0
JOURNAL OF APPLIED SCIENCES	4	25	1	0	1	1
JOURNAL OF AGRICULTURAL SCIENCE AND TECHNOLOGY	5	20	0	0	1	1
JOURNAL OF ENVIRONMENTAL PROTECTION	5	20	0	0	1	1
FRONTIERS IN BIOSCIENCE	45	16	1	1	0	1
INTERNATIONAL JOURNAL OF ADVANCED ROBOTIC SYSTEMS	7	14	1	1	1	1
PHARMACOLOGYONLINE (PHOL)	8	13	1	0	1	0

Notes: The number of replies stands for the number of replies by respondents that could recall the experience with the journal. 'Bad practice' indicates cases when the author did not receive a referee report, received only a superficial referee report, or noticed something that made him/her distrust the integrity/professionalism of the journal.