A "Trojan Horse" in the peer-reviewed economic journals. How is trustworthiness the editorial process in fee-charging journals?

Roberto Dell'anno[*], Roberto Dell'Anno[‡], Andrea Morone[§]

(Version: 17th September 2019)

Abstract

This paper aims to unmask the inadequate quality standards applied by editors of a sample of fee-charging journals in economics. We submitted a bait-manuscript to 104 academic economic journals to test whether there is a difference in the peer-review process between *Article Processing Charges (APC)-charging journals* and *Traditional journals* which do not require a publication fee. The submitted bait-article, was based on completely made up data, with evident errors in terms of methodology, literature, reporting of results and quality of language. Nevertheless, about half (40% in Sample 1 and 66.7% in Sample 2) of the APC journals fell in the trap. Their editors accepted the article in the journals and required to pay the publication fee. We conclude that the Traditional model has a more effective incentive-mechanism in selecting articles, based on quality standards. Accordingly, articles published on APC-charging journals cannot be indexed mechanically in scientific database indexes (e.g., Scopus, ISI Web of Science) as well as considered for bibliometric evaluations of research institutions or scholars' productivity.

JEL: *I20; A10.*

Keywords: *Predatory journal; Publication fees; Article Processing Charges; Peer-review.*

[*] Department of Economics, Management and Law, University of Bari A. Moro, Largo Abbazia Santa Scolastica, 53 - Bari 70124 - Italy. Email: rocco.caferra@gmail.com.

[‡] Department of Economics and Statistics - CELPE, University of Salerno, Via Giovanni Paolo II, 132 – Fisciano (SA) 84084 – Italy. Email: rdellanno@unisa.it.

[§] Department of Economics, Management and Law, University of Bari A. Moro, Largo Abbazia Santa Scolastica, 53 - Bari 70124 - Italy. Email: a.morone@gmail.com.

1 Introduction

Publishing articles on scientific journals is not only the foremost way for scholars to disseminate new ideas within the research community, but it also constitutes the source of scholar's bibliometric score.

Quantitative analysis of research production (i.e. bibliometrics), aiming to combine a "number" to each scholar, is increasingly used to rank researchers and academic institutions. According to the slogan that "numbers don't lie" - because they seem a more "objective", comparable and unbiased measure of scholar's research productivity - we are observing a structural change in assessment criteria of the researchers' activity. More and more decision makers use quantitative measure of research productivity and it is definitely becoming an unexceptional element to determine career paths of researchers and success of academic institutions. Indeed, on the one hand, research institutions encourage their affiliated to improve bibliometric performances because low ranking may penalize a department or a university reputation, and, especially for higher education, reputation affects student and scholar mobility as well as attractiveness in terms of sponsors and research funds. On the other hand, individual bibliometric scores are increasingly used by department boardrooms to decide which research projects and researchers should receive more support than others.

In this context, unsurprisingly, academics have been largely converted to the "publish or perish" religion, and particularly for younger academics, this religion is biasing the scholars' behavior with a negative impact on quality of research activity.[1]

Our study provides evidence from the field, that the combination of these incentive to increase bibliometric scores and the inclusion of Open Access Journals which require an Article Processing Charges (APC) (also known as publication fee) to publish, are conflicting with an effective decision process based on peer review to discriminate between valuable and useless research production.

In this sense, the issue of the paper is not whether bibliometric indicators can be considered useful to evaluate individual or institutional research – as they do provide a useful, even if not-exhaustive,

[1] Van Dalen and Henkens (2012) surveyed the high publication pressure perceived by researchers that negatively switched the interest of scientists from policy and knowing facts toward publication and citation within academic circles. Moreover Fanelli (2010) claimed that "it forces scientists to produce 'publishable' results at all costs."

assessment of research production – but if free-charging journals can be considered as trustworthy business model to prevent unethical and opportunistic conducts of publishers and scholars.

Two main (not alternative) business models to cover the cost of journal publications in scientific publishing exist: the "*Traditional model*" which refers to subscription-based journals requiring the reader to pay to access to the journal and the "*Open Access model*" of publication, in which journals publish with open access - i.e., the reader does not pay to access - but charge the author(s) of accepted articles of an APC.

There are also "*hybrid business models*" in which the publisher requires a payment of a fee when submitting of the manuscript (i.e. "submission fee")[2] or, to give the authors the opportunity to make their articles accessible to everyone on the web. However, differently from the Open access model with APC – traditional and hybrid models keep the editorial decision to accept manuscript separate from the payment of the fee.

We hypothesize that the existence of a publication fee is a sufficient condition to definitely change the relationship among authors, publishers, and readers. The basic intuition is that while in the *Traditional model*, there is a stronger editor incentive in applying an effective peer review process to assess the quality of submitted manuscript, this incentive disappears in the *APC journals*. It occurs because the APC model, transferring the role of funding the publications from users (i.e. scientific community) to the producers (authors) broke the incentive to look for the quality of the manuscripts because the publisher's (increasing the revenue through the fee) and the author's aim (increasing his/her bibliometric score through an indexed journal article) is immediately achieved, without the "final users'" evaluation (i.e. readers).

The paper is organized as follows. Section 2 reviews the literature. Section 3 describes the Experiment design, econometric models and reports the empirical outcomes. Section 4 concludes.

2 Literature

[2] In this hybrid category we include journals that surcharge authors of accepted papers that voluntary opt to pay charges for color figures in print.

Bibliometric scores serve as the basis for assessing the performance and impact of scholars, research projects and universities (Osterloh and Frey 2015). Extensive research investigates how this use of quantitative measures of research performances affects governance in academia and determine careers in universities.[3]

A widespread literature on the harmful consequences of bibliometric measures on research profession exists. Weingart (2005, p. 118) defined them "*theoretically unfounded, empirically crude and dependent on the data that we know to be imprecise*". Such types of metrics neither neutralize different customs of citing (e.g. article in biomedical research are cited six times more than the ones in mathematics) nor take into account the "*quantitative bias*" that could arise at the expense of the quality of research (e.g. the number of citations does not say everything about the quality of the paper that cited the considered one). Larivière and Gingran (2010) described this pattern recalling the "Matthew Effect": authors tend to read journals with a high impact factor and subsequently submit their article in such reviews, deflecting the attention to the quality of research they have previously read and simultaneously reducing the impact of a high level publication in a low-prestige journal.

Judging the quality of a journal is not always a trivial task, this is particularly important for the so-called "Predatory Journals" which try to mask their low reputation level by asserting a consistent peer-review, but they only aim to earn money through the publication fee (i.e. Article Processing Charge). In recent times, this process has been in the spotlight. Smith (2006) criticized the common misperception that "when something is peer-review is in some sense blessed", shedding light on the bias of non-standardized procedures which result in the subjective evaluation of the reviewers. Seidl et al. (2005) empirically found that in some cases peer-review lacks impartiality, validity and fairness, therefore an emerging body of research is testing the trustworthiness of the process in different journals. The vulnerability has been also fostered by the exceeding growth rate of online journals[4] that can publish as much as they like, increasing the pressure and assigning papers to reviewers who are not experts in the area (Arns, 2014). Our investigation is inspired by the aforementioned considerations and by the

[3] An example might be found in Durante et al. (2011), which evaluated university performance on the basis of familiarity by using the CIVR ("Steering Committee for the evaluation of research") score.

[4] The annual number of articles indexed in the publisher Elsevier's Scopus database increased from around 1.2 million in 2000 to roughly 2.7 million in 2013. That is an increase of 113%, but some of this rise is simply due to articles from more journals being included in the later count (Arns, 2014).

emergence of the individuation of a trustworthy distinguishing method. This is of the utmost importance to reduce asymmetric information in scientific evaluation. Bagues et al. (2017) surveyed the impact that predatory journals have on the Italian academia, finding that some of these journals have managed to be included in citation indexes and some researchers took advantage of this to get a promotion, succeeding especially when the committee (randomly selected) lack of expertise.
Baxt et al. (1998) used a fictitious manuscript to evaluate the peer-review performance founding that referees failed to check two-thirds of the major errors. Bohannons (2013) sent a bait-article to 304 journals, finding that more than half of the journals had accepted the paper, bringing into question the reliability of peer reviewing. Sorokowski et al. (2017) criticized Bohanons's selection process because it did not include non-open-access journals nor did it explicitly compare titles that did or did not have an impact factor. They re-designed the study in order to compare whitelist and blacklist journals, coming up with a similar result and highlighting a general tendency to capture some type of profit[5].
Following Bohannons (2013) approach, we aim to contribute to the current literature by providing a clear-cut normative approach to deal with predatory publishers. We have used the list of scientific journals utilized by the Italian national agency for the evaluation of universities and research institutes (ANVUR) to assess research production in the area of Economics (the so-called Area 13/A and 13/B in Italian classification). Differently from the previous research, we have used the request of a publication fee as distinctive factor, hence we have tested if the APC-funded model is a sufficient clue to infer the low reliability of journal peer-reviewing.

3 Experimental Design

3.1 Sample

The ANVUR ranked 2731 international journals covering the *Economics and Statistics* Area, on the basis of bibliometric indicators calculated until 2014. The sample selection consists in two steps. In the first step we select the treatment group by screening the 2073 journals indexed in the scientific area of

[5] In their research, they created a profile of a fictitious scientist named Anna O. Szust and applied on her behalf to the editorial boards of 360 journals.

"Administrative and Management" or "Economics" by the Italian Research Quality Assessment exercise (2011-2014) (VQR, 2017) in order to pick those which require a payment of APC.[6] We found that 53 journals match this condition, 31 of them are classified as "Administrative and Management" and 22 in "Economics".

In the second step, we select the Control group by picking, from the VQR List, for each journal included in the Treatment group, a "*NoAPC*" (i.e., "Traditional") journal that: (1) belongs to the same area (i.e. "*Administration and Management*", "*Economics*"); (2) it is indexed in the same bibliometric database (ISI Web of Science and /or Scopus); (3) minimizes the difference between NoAPC- and APC-journal metrics (i.e. IF5, AIS, IPP, SJR and the h-index). These bibliometric statistics of journals are computed by using a combination of the ISI Web of Science metrics (five-year impact factor - IF5; Article Influence Score AIS), Scopus metrics (Impact per Publication – IPP; SCImago Journal Rank - SJR) and h-index estimated by Google Scholar (53% of sample) or by "Publish or Perish" software (47%) over the period 2010-2014. For non-ISI and non-Scopus journals, the metrics was imputed from the estimation of the correlation between IF5, AIS, IPP, SJR and the h-index for the journal. These bibliometric data have been made public in the last release by ANVUR in February 2017.[7] Table 1 summarizes the bibliometric criteria used to define the units included in the Control Group.

Table 1. Selecting Criteria for Treatment (APC) and Control (No APC) Groups

Variable	*Description*
ISI	Dummy for the presence in ISI WoS database (1=present, 0=not present)
Scopus	Dummy for the presence in Scopus database (1=present, 0=not present)
Area	Scientific Area: Administration and Management (0) Economics (1)
IF5	5-year Impact Factor 2014. Source: ISI WoS.
AIS	Article Influence Score 2014. Source: ISI WoS.
IPP	Impact per Publication 2014. Source: Scopus.
SJR	SCImago Journal Rank 2014. Source: Scopus.
h-index	h-index over 2010-2014. Source: Google Scholar

[6] This selection is both automatically and manually implemented. We create different loops in R to web scrape information from Journal Guide and DOAJ website, such as publisher and APC, and to scan SCImago dataset, publishers' open access and pricing list and Beall's list of predatory publishers. The residual unidentified reviews have been manually checked. For the sake of experimental design, we exclude Journals of "Statistics" Area.

[7] Dataset and details on VQR exercise are available : http://www.anvur.it/attivita/vqr/vqr-2011-2014/gev/area-13-scienze-economiche-e-statistiche/

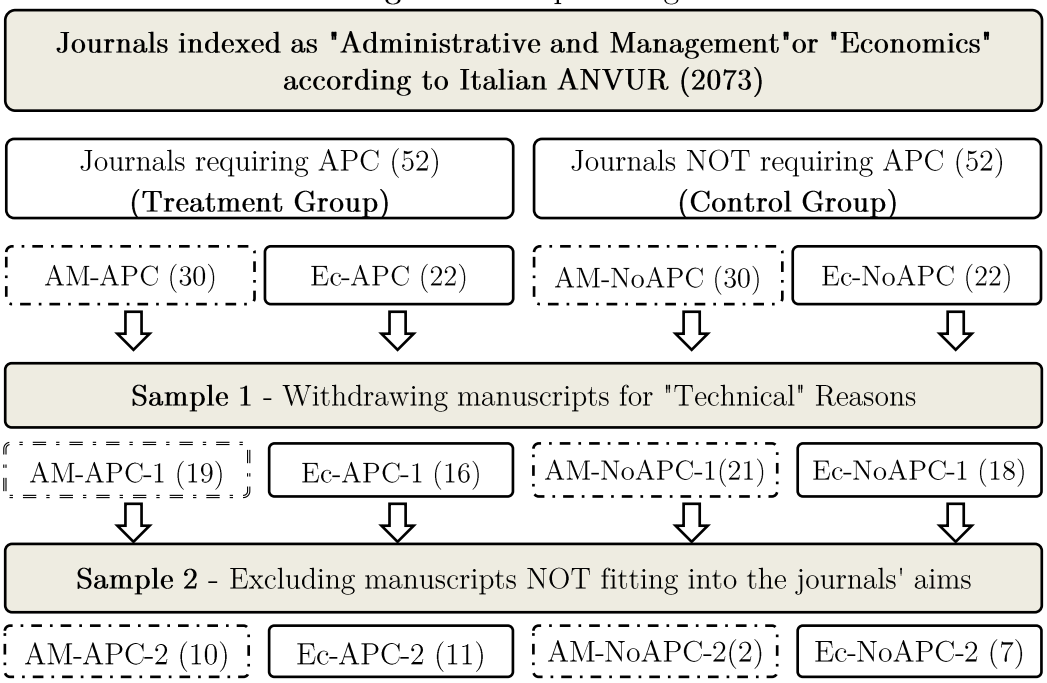
Accordingly, we have identified 104 journals from the official database collected by the ANVUR for the last (second) Italian Research Quality Assessment exercise (2011-2014).

Once the theoretical sample has been defined, we identify the final dataset through additional screening processes which consist in excluding journals due to "technical" reasons. In particular, this sample (hereinafter Sample 1) is derived by dropping 30 journals for six types of impediments: (1) we exclude 15 journals because the submission process is managed by the same editorial office, therefore we submit the bait-manuscripts only to one journal for each editorial office; (2) for 6 cases, editor/editorial office sent back the manuscript because it was not in accordance with the format style guidelines of the journal[8]; (3) for 6 journals, submissions are no longer allowed; (4) for 1 cases the submission was linked to the sending of personal documents of submitter; (5) for 1 journals submission are possible only with editor invitation and (6) for 1 journal payment of fee was required before submission. Accordingly, the empirical analysis in Sample 1 is based on 74 journals: 35 journals requiring APC and 39 journals not requiring APC (*No APC*).

A second test is based on a sample where we control for differences between the bait-manuscript's and the journals' topic (hereinafter Sample 2). Specifically, we drop from the Sample 1 those journals that, according to editors' emails sent to the "fake" submitters, reject the manuscript because it does not fit with the journal's aim. Sample 2 includes 31 journals: 22 of them requiring APC and 9 not requiring APC. Figure 1 summarizes the sample design.

[8] Frequently, these requests of reediting consisted in introducing new sections, inserting structured abstract, etc. If we had followed these editorial requests, we would have violated the hypothesis that all the Journals receive exactly the same manuscript, therefore we decided to withdraw the manuscript from these journals.

Figure 1: Sample Design



Note: AM = Administrative and Management; Ec = Economics. Number of journals in parenthesis.

Table 2 summarizes the Groups composition for the Sample 1 and 2. We report p-values of tests checking if Treatment (APC) and Control (No APC) groups have equal means in terms of bibliometric scores. These tests confirm that there are not statistically significant differences between groups at 5% level.

Table 2. Descriptive statistics of Sample/Groups composition

	T.S.×	Sample 1			Sample 2		
Variable	***All***	***All***	***(Treat.) APC***	***(Contr.) No APC***	***All***	***(Treat.) APC***	***(Contr.) No APC***
ISI *(Mean)*	**5.77%**	5.4%	5.7%	5.1%	0.0%	0.0%	0.0%
Scopus *(Mean)*	**26.0%**	27.0%	25.7%	28.2%	13.3%	19.0%	0.0%
IF5 *(Mean)*	**0.57**	0.64	0.69	0.59	0.54	0.62	0.34
*p-value**			0.49			0.097	
AIS *(Mean)*	***0.18***	*0.21*	0.24	0.19	*0.19*	0.22	0.12
*p-value**			0.33			0.097	
IPP *(Mean)*	***0.49***	*0.57*	0.55	0.59	*0.41*	0.48	0.26
*p-value**			0.79			0.14	
SJR *(Mean)*	***0.37***	*0.37*	0.35	0.40	*0.29*	0.31	0.25
*p-value**			0.36			0.42	
h-index *(Mean)*	***10.5***	*11.6*	12.14	11.13	*10.3*	11.4	7.78
*p-value**			0.59			0.15	
DOAJ§	*26.9%*	*25.7%*	*51.4%*	*0.03%*	*23.3%*	*28.6%*	*11.1%*
COPE§	*79.8%*	*81.1%*	*62.9%*	*97.4%*	*66.6%*	*57.1%*	*88.9%*
Beall's List§	*17.3%*	*18.9%*	*37.1%*	*0.3%*	*33.3%*	*47.6%*	*0.0%*
# Journals	***104***	***74***	**35**	**39**	***30***	**21**	**9**

*Notes: ×T.S.= Theoretical Sample; *p-value of two-sample t-test on equal means by assuming unequal variances (H_0); § share of journals included in Directory of Open Access Journals (DOAJ); share of journals belong to Committee on Publication Ethics (COPE); share of journals whose publishers are included in the Beall's List.*

3.2 The bait-manuscript

The bait-manuscript submitted to sample of journals follows Bohannon's (2013) scheme. The manuscript had a credible layout for an economic journal article. It included 5 sections (Introduction, Literature, Methodology, Results and Discussion, Conclusions), and a data appendix. This fake manuscript has been written in Italian and automatically translated to English using Google Translator[9]. In addition to the mistranslations we had also added an inconsistent and erroneous use of decimal separator, i.e. we use the comma to separate the integer part from the fractional part of numbers reported in main text and tables, while we use the point for the numbers reported in the appendixes. This kind of sloppiness is aimed to detect if there are differences in language checking between APC and not APC journals.

[9] https://translate.google.com/

In terms of content, we submitted practically identical papers to two groups of journals grouped in two partially overlapping areas of economic research (i.e. Administration and Management and Economics). The only difference was the submitter's name. Indeed, we have used two fake identities, in order to make it easier to manage these multiple communications between the "fake author" and the editors. The name chosen was the translation of "*Misunderstood Genius*" in Welsh and Haitian Creole languages. For the affiliations, we have combined the generic "*National University of*" with the capital cities of the Welsh and Haitian institutions.[10] In particular, the submitter was the Assistant Prof. Camddeall Athrylith - Department of Economics, Management and Statistics, National University of Aberystwth, Wales for the sample of "*Administrative Sciences and Management*" journals, and the Assistant Prof. Jeni Konpreyansyon - Department of Economics, Management and Statistics, National University of Port-au Prince, Haiti for the "*Economics*" journals.

The title of bait-manuscript is: "Crime and Economic Growth. An empirical analysis for Germany". The abstract of the submitted papers is as follows: "*This paper examines the role of crime, enforcement and taxation on the economic growth. These effects are studied by modified version of an endogenous growth model proposed by Loayza (1996). Econometric results are based on the German economy over the period 1992-2016. Empirical evidence confirms the theoretical model. We show that the relationship between crime and growth rate of GDP is negative in the long run equilibrium.*"

The manuscript counted 71 errors that a skilled peer reviewer should identify and suggest the editor to reject the manuscript. In particular, in addition to the slapdash quality of English, there were: 3 wrong first derivatives; 9 comments completely reversed compared to the empirical results reported in the table of manuscript; 16 erroneous interpretations of statistical significance of reported t-students (on 45 estimated coefficients); incorrect interpretations of Durbin Watson statistics, Cointegration and Unit Roots, as a consequence of the wrong interpretation the transformations applied to the variables included in the model were inappropriate; the conclusions and policy implications were in contrast with the empirical outcomes; some data sources and code of variables were fictitious; some descriptive statistics were false (e.g., we report some means which were lower than the minimum values, or larger than maximum values of the variables). Some examples of these errors are shown in Figure 2.

[10] As Bohannon (2013: 62) we used authors and affiliations from developing countries because "*authors and institutions would arouse less suspicion if a curious editor were to find nothing about them on the Internet.*"

Figure 2. Frames extrapolated from the bait-article

Frame 2.a: Examples of errors in analytical derivations (e.g., these two derivatives should be positive).

(5) $$\eta = p\lambda(q-\lambda)$$

where, by imposing the constraint on the quality of the institutions $q > 2\lambda$, the partial derivatives will be obtained that guarantee the desired functional characteristics: $\frac{\partial \eta}{\partial \lambda} > 0$; $\frac{\partial \eta}{\partial p} > 0$; $\frac{\partial \eta}{\partial q} < 0$; $\frac{\partial^2 \eta}{\partial \lambda \partial q} > 0$; $\frac{\partial^2 \eta}{\partial p \partial q} < 0$; $\frac{\partial^2 \eta}{\partial \lambda^2} < 0$. Indicating with S the relative size of the legal economy compared to the total one

Frame 2.b: Examples of errors in interpretation of empirical results

Table 1 - dependent Variable: growth rate of GDP per capita

	I	II	III	IV	V	VI	VII
Constant	-0,214 (-0,411)	-1,349** (-2,584)	-0,787** (-2,563)	-0,806 (-1,552)	-	-0,227 (-0,724)	0,242 (0,808)
Tax burden (τ_1)	-0,240 (-0,788)	-0,395 (-1,464)	-0,481 (-1,595)	-			
Tax burden^2 (τ_1)^2	-	-	0,225 (1,660)	-			
Pers. Inc.T. burden. (τ_2)			-	-0,533 (-1,902			
Pers. Inc.T. burden.^2 (τ_2)			-	-0,604* (-3,169)	(-5,200)	(-3,067)	(-2,191)
Criminal Economy (EC)	0,202*** (2,176)	0,158** (2,159)	0,090 (1,299)	0,212* (2,383)	0,218** (2,218)	0,206** (3,491)	0,197* (1,407)
Index of *enforcement* (λ)	-	-1,525** (-1,867)	-1,107** (-2,105)	-0,878* (-1,693)	-0,489* (-1,691)	-0,666** (-1,700)	-
Control Var.							
Internat. Trade Openness	0,285*	0,360**	0,338*** (3,054)				
Tertiary Education	(0,098)	(1,945)	-				
workforce participation rate	0,304 (0,632)	-0,152 (0,707)	-	(0,714)	(0658)	-	-
Observations	24	24	24	24	24	24	24
R^2-adjustd	0,190	0,604	0,623	0,665	0,631	0,685	0,253
Durbin-Watson st.						0,498	0,357
LM Test serial-correl.[a]						0,878	0,693
BPG Test Heterosched.[b]	0,726	0,998	0,997	0,978	0,603	0,673	0,979
JB Test Normality[c]	0,648	0,385	0,667	0,778	0,744	0,560	0,652

According to t-tests these coefficients are not statistically significant

This is the main result of the paper. In the Abstract, discussion and Conclusion we state that the "*relationship between crime and growth rate of GDP is negative*". At the contrary, the empirical result validates a positive correlation.

They are significant at 10%

The second conclusion of the paper is that the ". *The econometric analysis showed a positive correlation between the two variables*"[enforcement and GDP growth]. At the contrary, the empirical result shows a not statistically significant coefficient at 5%,(or negative at 10%).

Tertiary education counts 22 observations (as reported in the appendix) therefore these regressions cannot include 24

Frame 2.c. Examples of errors in Database

	Sources	Means	MAX	MIN	# obs
growth rate of GDP per capita (%)	World Bank – "GDP per capita growth (annual %)". World Development Indicators Online. Washington, DC. Codice: NY.GDP.PCAP.ZK.GD	1.6	4.1	-1.0	24 (‘92-‘16)
Tax burden (%) (τ_1)	Eurostat - tax revenue (including social contributions) in % of GDP, [gov_10a_taxag]	39.0	43.7	31.1	24 (‘92-‘16)
Pers.Inc. Tax/GDP (%) (τ_2)	Eurostat - Tax on personal income in % of GDP [gov_11a_taxag]	13.5	15.8	9.3	24 (‘92-’16)
Index of *enforcement* (λ)	Statistisches Bundesamt, (Destatis). Strafrechtliche Statistiken. (several years): Tab. 10.1 (1992, 1993); Tab. 3.1 (1994-2016) – numero persone denunciate per i quali l'Autorità giudiziaria ha iniziato l'azione penale / 100.000. La serie è trasformata in logaritmi.	5.6	5.7		24 (‘92-’16)
Criminal economy/GDP (%)	Jehle (2017). Reddito criminale prodotto in % of GDP (Tab. 1, pag. 11).	38.3	37.1	26.9	24 (‘92-’16)
International Trade Openness (%)	World Bank – "Merchandise trade as a share of GDP". World Development Indicators Online. Washington, DC. Codice: IMP.EXP.TR	37.4	47.8	[illegible]	24 (‘92-’16)
Tertiary Eduction (%)	World Bank – "School enrollment, tertiary (% gross). World Development Indicators Online. Washington, DC. Codice: SE.TER.ENRR (manca il valore per il 1998)	40.8	67.1	24.8	22 (‘92-’16)
Workforce participation rate (%)	World Bank – "Labor participation rate, total (% of total population ages 15+). World Development Indicators Online. Washington, DC. Codice: SL.TTF.CA.ZS	48.9	40.4	49.9	24 (‘92-‘16

This code does not exists. The right code is: NY.GDP.PCAP.KD.ZG

This variable and code do not exist in Eurostat database

This variable does not exist. The text is not in English, but German and Italian

MEAN<MIN

MEAN>MAX

These variables and codes do not exist.

3.3 Results

Between the 17th and 26th January 2019, each of two alter-ego of Assistant Prof. *Misunderstood Genius* sent his bogus manuscript to 104 journals, due to "technical reasons" the original sample size reduces to 74 journals by defining Sample 1. We disentangle the journal's evaluation process in three steps:

- the first step in which editor and/or reviewers evaluate whether the manuscript sufficiently suits the journal's aim. In favorable case, these journals are included in Sample 2.
- The second round in which editors and/or referees evaluate the scientific significance of the submitted manuscript.
- The third round in which editors and/or reviewers who had required revisions, evaluate the same paper without any further revision. Indeed, after few weeks from revision request, *Misunderstood Genius* re-submitted the original version without any changes but, in his cover letter for the editor and reviewers he stated: *"Dear Editor, please find attached the revised manuscript. Sincerely [author's name]".*

Figure 3 shows the outcomes of this analysis on Sample 1 and Sample 2.

Figure 3: Results based on Sample 1 - (APC Vs NoAPC)

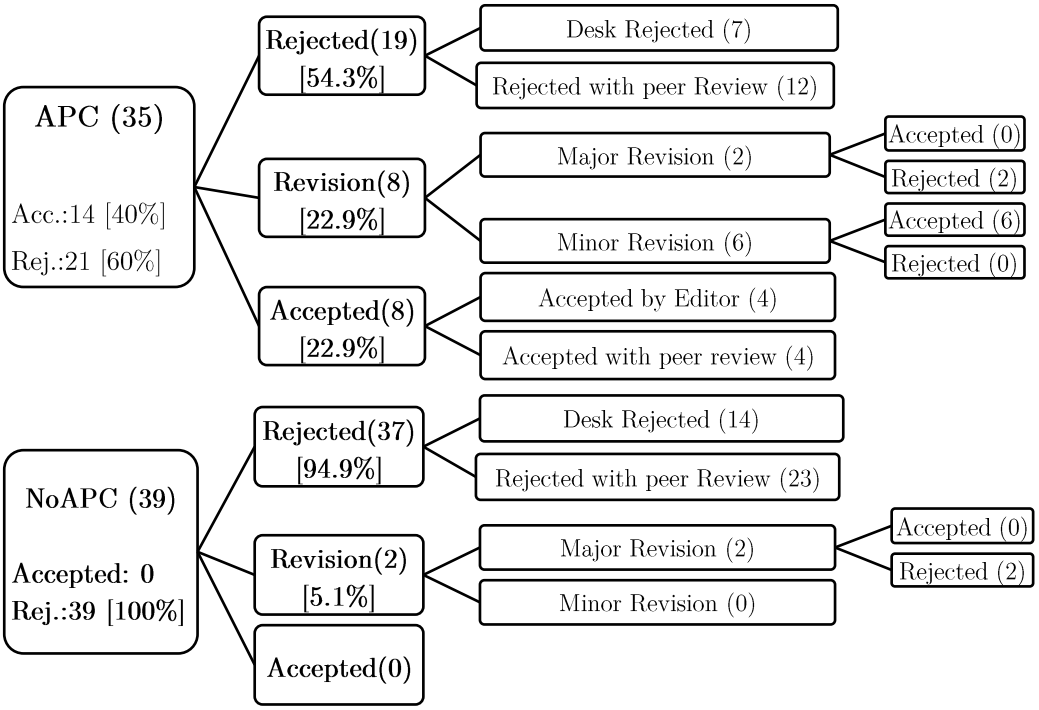


Figure 4: Results based on Sample 2 (APC Vs NoAPC)

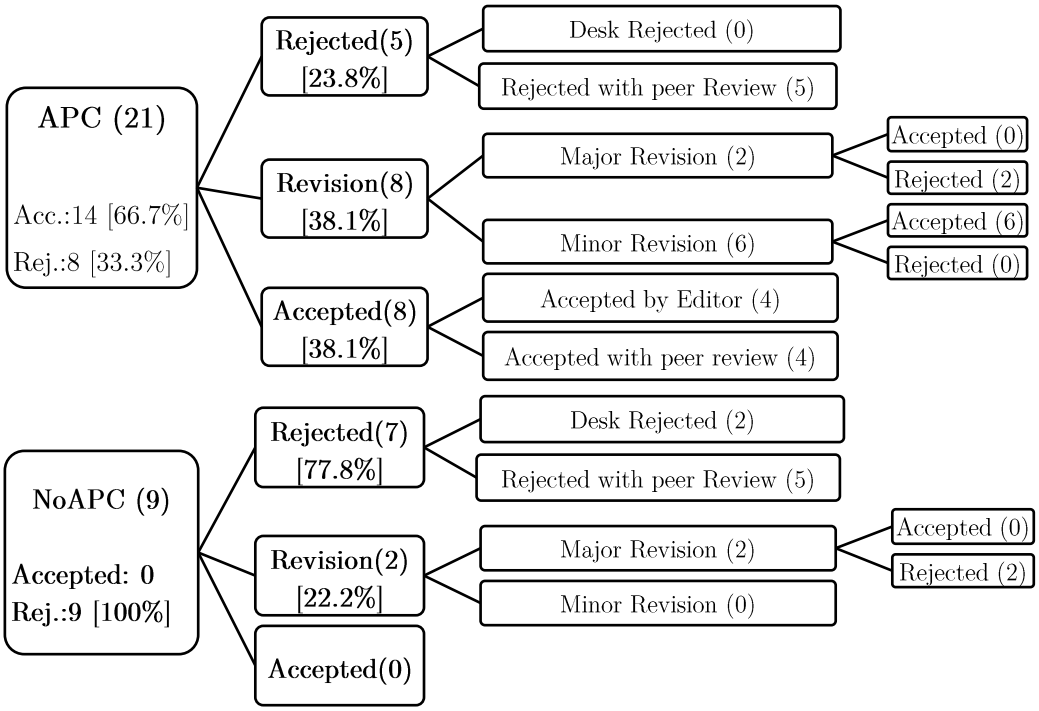
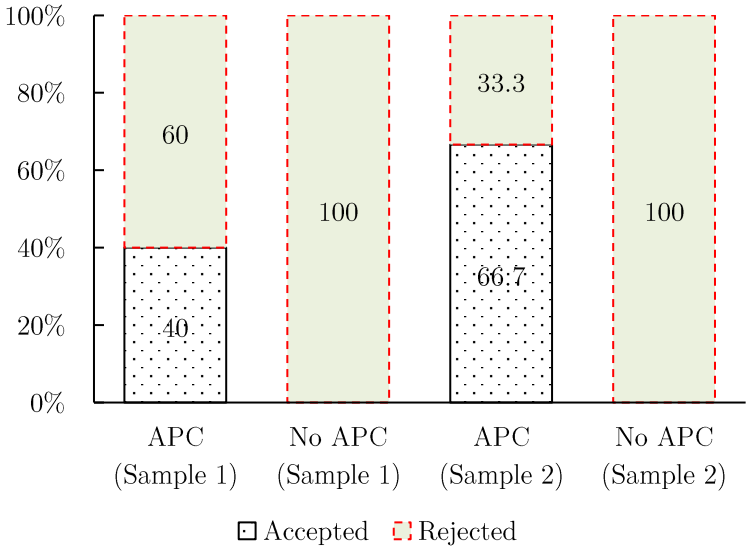


Figure 3 and Table 4 highlight as our quasi-natural experiment validates the hypothesis that Publication fees (APC) are a sufficient condition to deeply change the relationship among authors and editors. For the *Traditional model (No APC)*, an effective peer review process to assess the quality of submitted

manuscript exists, this incentive disappears for *APC* journals. In particular, APC journals exhibit a significantly higher acceptance rate (40% in Sample 1 and 66.7% in Sample 2) than No APC (0% in both Samples), marking a clear line between treatment and control groups (see Figure 5).

Figure 5. Acceptance/Rejection rates comparison between APC and No APC.



In order to test the validity of our findings, we conduct a binomial power test to verify the minimum sample size for each group attainable through our data given a significative level of 5% and power of 90% (Table 3).

Table 3. Power test on the Minimum sample size.

	Sample 1				Sample 2			
	N.		Acceptance r.		N.		Acceptance r.	
APC (Treatment Group)	35		0.4000		21		0.6667	
No APC (Control Group)	39		0.0000		9		0.0000	
significative level (α)	0.05				0.05			
power (1-β)	0.80	0.90	0.80	0.90	0.80	0.90	0.80	0.90
Alternative	One-sided		Two-sides		One-sided		Two-sides	
Minimum N for each Group	12	16	15	19	6	7	7	9

Table 4 reports statistical tests on the hypothesis that the acceptance rates are equal in two samples (APC Vs No APC).

Table 4. Tests on Difference between Acceptance rates

	Obs.	Sample 1	Obs.	Sample 2
APC (Treatment Group)	35	0.4000 (0.084)	21	0.6667 (0.105)
No APC (Control Group)	39	0.0000 (0.000)	9	0.0000 (0.000)
H_0: APC = No APC (p-value)		0.000		0.000
H_1: APC ≠ No APC (p-value)		1.000		1.000
H_1: APC < No APC (p-value)		1.000		1.000
t-stat on difference between groups		-4.761		-6.062
Welch's degrees of freedom		34		20

Note: Standard Error in parenthesis. Unequal variances between groups is assumed.

According to Table 3 results both the samples have a size sufficient to state that there is a statistical evidence in favor of the diversity between the two groups (Table 4).

In conclusion, peer review in APC journals is strongly questioned by the results. These journals accepted the paper without making a proper and accurate peer review, therefore there is sufficient evidence to state that more than 50% of APC journals accept publications without effectively screen for quality criteria of research.

4 Conclusion

The aim of this research is to analyze how the APC business model, independently from the Open Access nature of journals, biases the review process extraordinarily rising the acceptance rate. Accordingly, the peer review process instead of working as a mechanism aimed to recognizing quality, preventing plagiarism, expanding knowledge and promoting innovative research,[11] becomes an empty word recalled by predatory journals to catch, on the one hand, unexperienced (or unscrupulous) scholars and, on the other hand, to fulfil the standard requirements to be indexed on the most relevant databases of peer-reviewed publications.[12]

To carry out this research we submitted a bait-manuscript to 74 (out of 104) academic economic journals to test if there is a difference in the peer-review process between a treatment group (35 "*APC-*

[11] See Gans and Shepherd (1994) and Seidl et al. (2005) Smith (2016) for concerns in peer review process.

[12] For instance, the peer review process is among the required conditions to be included in the Scopus and ISI Web of Science database. We found that 2 Journals indexed by Scopus accepted the bait-article, while any Journals included in the ISI Web of Science fell into the trap.

charging journals) and a control Group *(39 Traditional journals* i.e. that do not require a payment of publication fee).

The submitted bait-article, was based on completely invented data, with evident errors in terms of methodology, literature, reporting of results, quality of language, policy implications and some fake references. Nevertheless 40% (Sample 1) and 66.7% (Sample 2) of APC journals fell into the trap: the editors accepted the manuscript and, required to us to pay the fee to proceeded to publish the article in their journals.

We find that the APC business model has shown its inappropriateness to preserve the standard of scientific publishing because the conflict of interests between journals and authors to publish poor research studies disappears. Accordingly, due to great consequences of bibliometric scores in determining career paths of researchers and performances of academic institutions, we conclude that relying exclusively on the editors' and authors' ethics, is not an effective method to preserve opportunistic behaviors.

We conclude that the *Traditional model* has a more effective incentive-mechanism to select based on quality standards. This model is able to adequately solve the publishers' commercial goal from the scientific community's aim to distinguish between good and poor quality of scientific studies. This is made possible because Traditional Journals through effectual peer-review tend to publish only accurate articles because they attract more journal citations, these quotations increase Impact Factors (IFs) and, in turn, increase revenue from subscriptions and attentiveness for sponsors. Journals with higher IF also increase the prestige of members of the editorial board and, in turn, the quality of peer-review process still increases in the long term.

This study has two main policy implications. The first one deals with the issue of efficient use of public resources in academic research. Indeed, payment of APC is usually sponsored by research department, and for useless articles, it constitutes a clear waste of research funds (for instance Assistant Prof. *Misunderstood Genius* had to pay about 4,900 US dollars as APC for the 14 accepted manuscripts). The second implication deals with the effect that these fee-charging journals have on scholars and institutions bibliometric scores. Bibliometric statistics are often used in a comparative way (e.g. the Italian national agency for the evaluation of universities and research institutes sets minimum standards of research production for candidates to the National Scientific Qualification System based

on median values of the overall research production of Italian academic community) therefore predatory journals may distort informative sets for decision making on governance in universities and academic recruitment.

References

[1] Bohannon J. (2013). Who's Afraid of Peer Review? Science, 342 (6154): 60-65.

[2] Baxt W.G., Waeckerle J.F., Berlin J.A., Callaham M.L. (1998). Who Reviews the Reviewers? Feasibility of Using a Fictious Manuscript to Evaluate Peer Reviewer Performance. Annals of Emergency Medicine 32(3): 310-317.

[3] Bagues, M. F., Sylos-Labini, M., Zinovyeva, N., (2017). A Walk on the Wild Side: 'Predatory' Journals and Information Asymmetries in Scientific Evaluations. IZA Discussion Papers 11041, Institute for the Study of Labor (IZA).

[4] Durante, R., Labartino, G., Perotti, R. (2011). Academic Dynasties: Decentralization and Familism in the Italian Academia. Available at NBER: http://www.nber.org/papers/w17572

[5] van Dalen, Hendrik P., Henkens, Kene, (2012). Intended and Unintended Consequences of a Publish-or-Perish Culture: A Worldwide Survey. CentER Discussion Paper Series No2012003. Available at SSRN: https://ssrn.com/abstract=1983205

[6] Gans J.J., Shepherd G. B. (1994). How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists. The Journal of Economic Perspectives, 8(1): 165-179.

[7] Grösche, P., Schmidt, U., Seidl, C. (2005). The Performance of Peer Review and a Beauty Contest of Referee Processes of Economics Journals. Estudios de Economía Aplicada. 23: 505-551.

[8] Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US states data. PLoS ONE, 5(4), e10271. doi: 10.1371/journal.pone.0010271

[9] Haustein, S., Larivière, V. (2014). The Use of Bibliometrics for Assessing Research: Possibilities, Limitations and Adverse Effects. À paraître dans (Eds.) I. Welpe, J. Wollersheim, S. Ringelhan & M. Osterloh, Incentives and performance: Governance of knowledge-intensive organizations, Cham: Springer International Publishing.

[10] Osterloh M., Frey B. S., (2015). Ranking Games. Evaluation Review 39(1): 102-129

[11] Seidl C., Schmidt U., Grösche P. (2005). The performance of peer review and a beauty contest of referee processes of economics journals. Estudios de Economía Aplicada, 23: 249-286.

[12] Sorokowski Piotr, Kulczycki Emanuel, Sorokowska Agnieszka, Pisansk Katarzyna (2017). Predatory journals recruit fake editor, Nature News, Mar 22.

[13] Smith R. (2006). Peer review: a flawed process at the heart of science and journals. Journal of Royal Society of Medicine 99(4): 178–182.

[14] Solomon D.J., Björk B. (2012). A Study of Open Access Journals Using Article Processing Charges. Journal of the American Society for Information Science and Technology 63(8):1485–1495,

[15] Starbuck W. H. (2005). How Much Better Are the Most-Prestigious Journals? The Statistics of Academic Publication. Organization Science 16(2):180-200.

[16] Wallace F.H., Perri T. J., (2018). Economists behaving badly: publications in predatory journals Scientometrics 115(2):749-766.

[17] Weingart, P. Impact of bibliometrics upon the science system: Inadvertent consequences? Scientometrics 62(1): 117–131.