Research Policy 49 (2020) 103951

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

The Matthew effect of a journal's ranking

Kyriakos Drivas[a,1,*], Dimitris Kremmydas[b,1]

[a] Department of Economics, University of Piraeus, Karaoli & Dimitriou 80, Piraeus 18534, Greece
[b] Department of Agricultural Economics & Rural Development, Agricultural University Of Athens, Iera Odos 75 118 55 Athens, Greece

ARTICLE INFO

Keywords:
Journal rankings
AJG/ABS list
Scientific diffusion;Matthew effect
Difference-in-differences
Coarsened exact matching

JEL:
L51
I23
O38

ABSTRACT

We examine the impact of a journal's ranking on the citations its papers garner. Our testbed is journals included in the early annual publications of the Academic Journal Guide, published by the Association of Business Schools from 2007 to 2010. By focusing on a small subset of these journals, we provide causal evidence that an increase in a journal's ranking will increase citations to its papers. We argue that this increase can be attributed to (i) more authors learning about and viewing these outlets and their publications and (ii) researchers signalling their paper's own impact by citing highly ranked journals. We find some evidence for the former, though not robust, and substantial evidence for the latter. Given that signalling is deliberate and associated with incentives to publish in highly ranked journals, we decompose this channel by citing researchers' characteristics. Except for senior academics, all types of researchers are consistent with signalling. The policy implications of our results relate to the pervasive use of journal rankings and recent initiatives to evaluate research.

1. Introduction

Many researchers today rely on journal rankings to set goals for the outlets they wish to publish. The main incentives for doing so are to achieve recognition and increase the likelihood of getting hired or promoted (Heckman and Moktan, 2018). Furthermore, funding agencies include journal ranking in their reward criteria to evaluate research output. Similarly, research administrators use such rankings to render their institution's research output measurable. The use of rankings by funding agencies and research administrators is meant to secure objectivity and transparency within academia and the general public (Judge et al., 2007).[2] The argument in favour of using journal rankings is to promote competition amongst researchers and therefore create one of the pillars for the 'entrepreneurial university' (Clark, 1998; Stensaker and Benner, 2013).

However, if the incentive structure outlined to the research community is heavily based on journal rankings, then there is potential to abuse it (Parker, 2014). In doing so, a research work may be evaluated primarily based on the journal in which it was published, neglecting the actual policy or real-world problem and its impact (Espeland and Sauder, 2007; Osterloh, 2010; Rafols et al., 2012; Alberts, 2013; Brooks et al., 2019). This type of abuse can have implications for the citations each given research work garners. Researchers who extensively employ journal rankings when reviewing previous literature in their work might be more likely to cite papers from highly ranked journals. Therefore, the citations a paper garners may be influenced by the journal it was published in rather its own merits.

This important distortion has not gone unnoticed in the literature. Studies have found a positive relationship between a journal's ranking, measured by a variety of metrics, and the citations its papers garner (Starbuck, 2005; Knothe, 2006; Judge et al., 2007; Ashton et al., 2009; Lariviere and Gingras, 2010; Lozano et al., 2012; Brooks and Schopohl, 2018).

The objective of this paper is twofold: first, to substantiate the past literature by providing causal evidence on whether journal rankings influence citations. Second, to identify and test two potential channels for why journal rankings influence citations, that is, to examine why researchers react to these journal ranking changes.

For the first channel, we argue that when a journal receives a high ranking, more researchers may learn about or review the outlet. Subsequently, increased viewership can potentially result in increased citations. We denote this as an 'information effect'. There is also a second, more nuanced channel. Researchers may cite highly ranked journals to signal that their own paper is part of the literature in such journals. We denote this as a 'signalling effect'. We borrow arguments from signalling theory and recent studies that examine the peer review

* Corresponding author.
E-mail address: drivas@unipi.gr (K. Drivas).
[1] Both authors have contributed equally.
[2] For an excellent discussion of how academics, policymakers and the public use journal ranking lists today, see Aguinis et al. (2020).

https://doi.org/10.1016/j.respol.2020.103951
Received 2 July 2019; Received in revised form 25 February 2020; Accepted 25 February 2020
Available online 03 March 2020
0048-7333/

process to substantiate this channel. Given the importance of this latter channel for policy, we examine whether the signalling effect is differentiated by a citing researcher's characteristics.

To causally test the effect of a journal's ranking on its papers' citations, we exploit a unique empirical setup. Our testbed is the Academic Journal Guide published by the Association of Business Schools (hereafter AJG/ABS list) which ranks journals in economics, business and management-related fields. Our research focus is explicitly on journal ranking lists rather than a journal's impact factor.[3] We do so because reliance on such lists in the social sciences by researchers, funding agencies and research administrators is becoming pervasive (Willmot 2011; Parker 2014; Hussain 2015), which can lead to abuse (Osterloh and Frey 2015).[4]

Methodologically, we examine papers published between 2000 and 2003 and the annual citations they received until 2018, performing a difference-in-differences (dif-in-difs) strategy between the two groups. The AJG/ABS list was first published in 2007 and then annually until 2010.[5] Ideally, we would look for journals that barely gained or lost the highest ranking. As we cannot observe this marginal ranking, we compiled two groups of journals that likely fall into this category. The first group comprises journals upgraded from the second-highest to the highest ranking during 2007–2010. The second group comprises journals that followed the opposite route. The geographic variation in the impact and use of the list in the early years allows us to control for each paper's underlying citedness, therefore alleviating any endogeneity concerns. We further match each paper from an upgraded journal with a paper from a downgraded journal based on their past citation rates employing the Coarsened Exact Matching (CEM) procedure by Iacus et al. (2012).

Our initial data collection includes a compilation of detailed bibliographic information for the focal 4000 papers and the roughly 200,000 papers that cite them. Additionally, we downloaded bibliographic information for the approximately 64,000 authors who cite the focal papers to identify the researchers' characteristics.

Our results show that papers from upgraded journals experience an increase in citations compared to papers from downgraded journals. Empirical results provide some evidence for the information effect, while the evidence for the signalling effect is more conclusive.

Of these two effects, the one that warrants the most attention is that of signalling. The information effect is generated organically and unintentionally. On the contrary, the signalling effect takes place intentionally to increase the probability of acceptance in highly ranked journals. To this end, we examined whether the signalling effect differs by types of researchers. We employed the recent survey of UK academics by Walker et al. (2018) concerning the AJG/ABS list to guide us for the types of researchers. We find that all types except senior researchers engage in signalling. We posit that senior researchers may not exhibit such an effect as they may be less likely to track changes in journal rankings.

Note this paper is not a criticism of the AJG/ABS list. There are several empirical studies that have provided refinements of lists or proposals how to generate more objective lists (Lane 2010; Hudson 2013; Mingers and Yang 2017).[6] Our motivation for employing the AJG/ABS list is that it provides a unique setting enabling inference of causality. There have been several respectable and rigorously designed journal ranking lists, and scholars have also employed a multitude of metrics to rank journals or provide methodologies to do so (Kalaitzidakis et al., 2003, 2011; Ritzberger, 2008; Laband, 2013; Sgroi and Oswald, 2013; Hole, 2017). However, the AJG/ABS list has had a profound impact as a tool for evaluating research by administrators in the UK (Willmot, 2011; Hussain, 2015). As a result, UK-based researchers overwhelmingly adopted the list. This source of exogenous variation is ideal to address the objectives of this paper.

2. Theoretical background

2.1. The effect of journal ranking on citations

Previous literature points to a positive relationship between a journal's status, evident by its impact factor or other ranking metrics, and its papers' citations (Starbuck, 2005; Judge et al., 2007; Lozano et al., 2012).[7] Studies by Knothe (2006) and Lariviere and Gingras (2010) identify papers that have been published in more than one journal – noted as duplicate papers – and examine whether their citation rates differed by the journal's impact factor. Both studies found that the paper's version published in a higher impact factor journal was associated with more citations. Our thesis is guided by these studies, suggesting that a journal's higher ranking will increase citations to its papers.

Nonetheless, our approach is different from the above studies. Instead of simply comparing citations between papers, we provide a dynamic analysis by first comparing citation differences within papers before and after the event of a journal upgrading or downgrading. Then we compare differences between papers in journals that were upgraded or downgraded.

This studied effect can be viewed as a Matthew effect for journals. The original Matthew effect, coined by Merton (1968), presents two pathways prestigious researchers can achieve higher levels of recognition. The first functions as a reward system, noted as a resource-based pathway. Prestigious researchers obtain more resources (e.g. PhD students, funding) and are therefore able to deliver higher-impact research. The second functions as a communication system, noted as a perception-based pathway. The prestigious status, researchers receive, alters peers' perceptions and, therefore, further elevates them amongst their peers. A notable implication of this latter pathway is that prestigious researchers' earlier work can garner more citations. Azoulay et al. (2014) focused exclusively on testing this perception-based pathway for researchers, earning the distinction of becoming Howard Hughes Medical Institute Investigators.

In our case of journals, we also focus on this pathway. A translation of the resource-based pathway could be that a highly ranked journal will attract the best and most impactful papers, which will attract further citations. The perception-based pathway, studied here, implies that a highly ranked journal will attract citations to its papers independently of their own impact but rather due to the journal's status. While it is only one of two components of the Matthew effect, it is arguably the most important for the case of journals. The resource-based pathway is correlated with the paper's own independent impact. However, the perception-based pathway implies that citations to a research work are independent of its merits.

[3] Naturally, journal ranking lists consider the impact factor when ranking journals, but they also consider other criteria, including the review process and the generality of the journal (ABS 2015).

[4] Journal rankings could also hamper creativity, a building block of scientific research (Spangenberg et al. 1990; Stephan 1996; Amabile 1998).

[5] The AJG/ABS list was then published in 2015 and 2018. We discuss in the forthcoming sections how these later versions may influence our results.

[6] For a comprehensive overview of lists with journal rankings, see Harzing (2019).

[7] There is anecdotal evidence to support this. For instance, economists know about the simultaneous formulation of the seminal model of economic growth in two independent works by Robert Solow and Trevor Swan in 1956. While there can be multiple reasons for the different impact of the two works (e.g., different affiliation, follow-up research work by Solow), part of this difference could be attributed to the journal in which each work was published. Solow's work was published in the *Quarterly Journal of Economics* (QJE), while Swan's was published in the *Economic Record.* While the latter is a rigorous and respectable journal with multiple contributions over the years, it does not share the same prestige as the former.

2.2. Information and signalling effects

The discussion above predicts that citations to journals that have moved up in rank will increase due to researchers' revised perceptions of these journals. However, we do not know why researchers may react to this change. Therefore, how these perceptions translate to citing behaviour requires additional context.

The first, intuitive, behaviour is the following. When a journal receives a high ranking in a journal list, researchers will learn of this journal or researchers who already are aware of the journal may be more inclined to read its research. Increased viewership of the works in those outlets is likely to spur citations. As an example, Feenberg et al. (2017) found that the random ordering of National Bureau of Economic Research (NBER) papers matters for their viewership and subsequent citations. The reasoning is that due to researchers' time constraints to review the literature, attention will be limited to papers that show up first on the list. With that in mind, we should expect that part of the increased citations to highly ranked journals should be attributed to an 'information effect'.

A second, more subtle behaviour, could be that authors act strategically when citing past research. In signalling theory, the landmark study by Spence (1973) discusses how potential applicants in the job market signal their attributes to employers. Spence shows that education can be an investment by a job applicant to signal quality, thereby differentiating the applicant from the rest of the pool.

In the peer review process, signalling has been observed at the editorial level. Endenich and Trapp (2018a, 2018b) discuss how the composition of editorial boards at leading accounting journals sends signal to the research community regarding acceptable contributions. Moreover, as Salterio (2018) argues, signals are not only emitted by journals but also from researchers themselves, who may resort to a narrow set of methods to reveal the technical soundness of their work.

Regarding citation behaviour, Baum (2011) discusses how certain journal editors may preferentially treat papers that cite their journal. Moreover, authors have been coerced to add superfluous citations to journals to act strategically on their impact factor (Wilhite and Fong 2012). Relevant to our paper, Hussain et al. (2019) examined the references disclosed in papers from six leading accounting journals. They found that the percentage of references to journals highly ranked by AJG/ABS has increased significantly in recent years. They attribute this to signalling behaviour by researchers to reveal to editors that their research is comparable to work published in highly ranked journals. Therefore, we should also expect that part of the increased citations to highly ranked journals can be attributed to a signalling effect.

Hussain et al. (2019) describe the same signalling argument outlined here. However, their research design cannot conclusively attribute this citing behaviour strictly to signalling. The authors do not examine references by papers published in lower-ranked accounting journals, hence they cannot compare how citation behaviour in those outlets has changed over the years. If the change in recent years (after publication of the AJG/ABS list) is similar between low and highly ranked journals, then their findings should be interpreted as an information effect. That is, all researchers, regardless of the outlets they publish, would cite highly ranked journals due to increased viewership. However, if the recent spike in references to AJG/ABS-ranked journals is seen only for highly ranked journals, then the findings of Hussain et al., could be interpreted as a signalling effect.

We substantiate the insightful study Hussain et al. (2019) by focusing on citations received by follow-on papers and the exploitation of changes in rankings, thereby identifying the existence of a signalling effect. Additionally, by considering journals from a wider array of disciplines, we can generalize the existence of the signalling effect.

2.2.1. Heterogeneity of signalling effect by researcher's characteristics

Although the information effect is generated organically, the signalling effect is intentional and the most worrisome channel. In terms of science policy, the signalling effect is manifested due to incentives to publish in highly ranked journals and the underlying premise that citing such works can increase, even marginally, the probability of acceptance. Therefore, it is useful to examine how this effect differs by examining researchers' profiles. Walker et al. (2018) examined the attributes of UK academics who are associated with extensive use of the AJG/ABS list. The most notable categories with which we can test the signalling effect, are i) affiliation, ii) seniority, iii) productivity.

3. Employing the AJG/ABS for the identification strategy

The AJG/ABS list was first published in 2007. The initial list included journals that had three or more papers submitted to the business and management panel of the Research Assessment Exercise in 2001. The list was then published again in 2008, 2009 and 2010, adding more journals that were previously included in other UK lists.[8] This process effectively consolidated these journal ranking lists in the AJG/ABS list (Morris et al., 2011). As Walker et al. (2018) discuss, Aston's 2008 journal list was the last formal journal ranking list of its kind, thereby making the AJG/ABS list the 'go to' guide to learn of a journal's ranking.

The ranking was simple and intuitive: a journal could be ranked as Grade-1, Grade-2, Grade-3 and Grade-4. The Grade-4 journals were the highest-ranked journals that 'publish the most original and best executed research' while Grade-1 journals were referred to as 'modest standard journals within their field' and 'are refereed relatively lightly according to accepted conventions'.[9]

To examine causally the effect of journal rankings on their papers' citations, we first need to consider endogeneity. The intuitive endogeneity issue is the following: if we were to perform a cross-sectional analysis and find that papers from Grade-4 journals receive more citations than their Grade-3 counterparts, then this could simply be attributed to Grade-4 journals on average accepting better papers than Grade-3 journals. Therefore, any citation difference could be attributed to the paper's own impact rather than the journal's ranking.

To account for this endogeneity concern, the first step is to focus on papers that were published between 2000 and 2003, a few years before the AJG/ABS list was first published and perform a dif-in-difs approach. In its rudimentary form, we could compare citation rates of papers that were published in Grade-4 and Grade-3 journals before and after the first four publications of the AJG/ABS list.

However, endogeneity could still be present because these journals already have an inherent ranking before the AJG/ABS list and therefore better papers could have been accepted in the Grade-4 journals. For this potential source of endogeneity, we need to consider for the second step searching for journals that were 'barely' awarded or lost the Grade-4 ranking. While we cannot observe this marginal ranking, we can observe journals that in 2007 were ranked as Grade-3 and by 2010 were ranked as Grade-4; we denote these journals and their papers as 'upgrades'. We also observe journals that were ranked as Grade-4 in 2007 and by 2010 were ranked as Grade-3; analogously we call them 'downgrades'.[10]

[8] Two of the editors of the 2007 AJG/ABS list were also involved in the compilation of the Bristol list of 2004, which ranked academic journals from 1 to 5 (Harvey and Morris, 2005). To provide validation and transparency, the 2007 AJG/ABS list included with its own rankings those from other UK universities including Warwick, Imperial, Cranfield, Kent, Aston and Durham. For more history of the AJG/ABS list, see Hussain (2010).

[9] The 2010 list also saw the enactment of Grade-4* journals which are 'grade four journals that are recognized worldwide as exemplars of excellence within the business and management field broadly defined and including economics'. However, this grading adds no information in our subsequent analysis and, therefore, was not examined.

[10] Note we do not consider the handful of journals that switched their ranking more than once. For instance, we did not consider journals that had a Grade-4 ranking in 2007, Grade-3 in 2008 and Grade-4 in 2009. Such switches would only add noise to the observed citation rates.

With the above in mind, the third step is to examine citation rates of papers between downgraded and upgraded journals before the first four editions were published, 2007–2010, and after. Overall, we identified twelve downgraded journals and twelve upgraded journals.[11] The list of these journals is displayed in Table A1 in Online Appendix A.

The setting of the AJG/ABS lists provides us with a powerful tool to control for each paper's citedness and, therefore, infer causality. Research administrators quickly started using this list to evaluate the work of researchers (Willmott, 2011; Hussain, 2015). As a result, the list became, at least in part, a criterion for authors deciding where to submit their papers. However, the described events initially took place in a single country, the UK. For instance, in discussing the early publications of the AJG/ABS list, Mingers et al. (2012) point to evidence that the list was used by the Committee of Professors of Operational Research for hiring and promotion decisions. In a survey, Walker et al. (2015) found that nine of ten academics in the UK actively use the AJG/ABS list.

We focus exclusively on citation behaviour by UK-affiliated researchers and examine how they modified their citation behaviour after the change in rankings. Therefore, our core dependant variable is citations of papers from authors who are only affiliated with UK institutions, which ensures any effect is due to the list's overwhelming acceptance in the UK. Also, focusing on UK citations facilitates employing part of the other citations each paper receives to control for its citedness.

Citations by US-affiliated researchers can function as a plausible candidate. In the early years, the United States was mostly unaware of the list or it did not make use of it. Although this might have changed in recent years (Walker et al., 2019), the early years saw the list's impact mainly within the UK (Taylor 2011; Willmott 2011; Hussain 2015). Moreover, US citations are the most populous in our data; therefore, to the extent that some citations are related to the AJG/ABS changes, such an effect is likely to be diluted due to the large citation counts.[12] Therefore, as a control variable we include the number of citations of papers where all the article's authors disclose only US-located affiliations.[13]

There is an additional benefit to including US citations as a control. It could be the case that the movements of upgraded and downgraded journals are confounded with other changes, independent of the AJG/ABS list. Therefore, any effect due to the list would be attributed to these omitted changes. However, if such changes were to influence the UK citations, then they would also influence the US citations. By controlling for this variable, we net out any other changes not incorporated in the model. We re-visit this issue in Section 5.1. in 'Robustness'.

By including US citation rates, we confidently control for each paper's underlying citedness. However, to achieve identification, we further match each paper from an upgraded journal to a paper from a downgraded journal. First, denote $UKCites_{i,\ t}$ as the citations by UK affiliates paper i receives in year/period t. Then, for every paper from an upgraded journal, we match a single paper from a downgraded journal that has the same publication year and the closest values of $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$, and we denote these two latter measures as 'matching' variables. Note that these two measures capture the citation rate of each paper before and around the publication of the first four editions of the list. We employ the CEM procedure formulated by Iacus et al. (2012) to match papers based on these 'matching' variables.

The CEM procedure non-parametrically matches a paper from an upgraded journal with a paper from a downgraded journal based on a set of pre-selected variables. The advantage of the CEM procedure is that it is non-parametric and therefore assumes no restrictions in the generation mechanism of the papers in upgraded journals compared to the papers in downgraded journals. Our choice of the 'matching' variables ensures that papers from upgraded journals do not display any pre-trend compared to the matched papers from downgraded journals. Our econometric estimations take place on this matched sample to alleviate any remaining endogeneity issues (Azoulay et al., 2010). In Section 5.1, we discuss in detail the implementation of the CEM procedure in our data.

It is crucial to discuss the follow-up publications of the AJG/ABS list and why these do not influence our estimations. The list was again published in 2015 and 2018. As we stop observing citations in 2018, then the latter publication is not influencing our results. As for the 2015 list, citations are naturally lagging by a couple of years, and therefore any effect of the 2015 version should be small. To ensure that such noise remains small, we verify that all twenty-four journals hold the same ranking in 2015 as in 2010.

4. Econometric specifications

4.1. Testing the effect of journal upgrading on citations

The goal of our econometric estimation is to employ a standard dif-in-difs approach to examine the effect on citations of papers from upgraded *vis-à-vis* downgraded journals before and after the publication of the first four versions of the AJG/ABS lists (2007–2010). Rather than considering a full sample, we implement the dif-in-difs on a matched sample of papers from upgraded and downgraded journals.

In detail, our data employed for the baseline estimations comprise an equal number of papers from upgraded and downgraded journals, respectively. We observe citations of all papers until 2018, with the earliest paper from 2000 and the latest 2003. For every paper from an upgraded journal in the data, there is a paper from a downgraded journal matched with the former paper based on publication year, and then the closest values of $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$ via the CEM procedure. Since the data comprise matched paper pairs, we need to identify each of these pairs separately; therefore, we include in all the regressions $Pair_{i,\ j}$ fixed effects (Azoulay et al., 2010; Furman and Stern, 2011). These are, in effect, n dummy variables, as many as the pairs of papers. A dummy variable $Pair_{i,\ j}$ takes the value of 1 only for the cases where paper i and paper j belong in the same pair, and 0 otherwise.

With that in mind, the baseline econometric specification is the following:

$$\begin{aligned} UKCites_{i,t} = b_0 + b_1 USCites_{i,t} + b_2 Upgrade_i + b_3 Upgrade_i\_x\_Window_t \\ + b_4 Upgrade_i\_x\_After_t + CitingYear_t + Pair_{i,j} \end{aligned} \quad (1)$$

Note that a citation is included in the *UKCites* only if all authors are UK affiliated and disclose no other affiliation outside the UK. Also, note that *UKCites* is net of citations made by the same journal or by the same author.[14] Unless otherwise noted, we have subtracted these types of self-citations from all citation-related variables. $USCites_{i,\ t}$ is the number of US affiliated papers that cite paper i in year/period t. As mentioned previously, a citation is included in the *USCites* when all authors are US affiliated and disclose no other affiliation outside the United States.

[11] We should stress that such upgrades and downgrades are not common. In the AJG/ABS 2010 edition, 94 journals were Grade-4 while 230 journals, Grade-3. This implies that the upgraded/downgraded journals comprise seven percent of the population of these journals.

[12] Even if the US citations overwhelmingly respond to the AJG/ABS list, then including them should bias our results against a Matthew effect. That is, by including them, the estimated effect of upgrading compared to the downgrading should be insignificant.

[13] Note that for both types of citations we require for all the authors to disclose affiliation only from the United States and the UK, respectively. We do this to ensure that all the citations in each citing article have been added by an author located in each focal country of interest.

[14] We re-visit this choice of netting citations in Section 6.1. in 'Robustness'.

$Upgrade_i$ takes the value of 1 if paper i is from an upgraded journal and 0 if it is from a downgraded journal.

$Window_t$ takes the value of 1 for the citing years 2009–2012 and 0 otherwise. $Upgrade_i\_x\_Window_t$ is the interaction term between the two variables. This interaction term shows the differential citation effect of papers from upgraded journals versus papers from downgraded journals compared to the pre-2009 period. We consider this measure for two reasons. First, the citation patterns of the two groups of papers during these years may be somewhat transitory. Journals have been upgraded or downgraded either during the 2008, 2009 or 2010 editions. Therefore, for papers where the journal upgrade or downgrade took place early, citations are bundled with citations of papers from journals that had not yet switched grade level. Second, the acceptance of the AJG/ABS list might have been heterogeneous across institutions and researchers in the early years. With this consideration, we allow for such a window as we are largely agnostic of its sign since it encompasses significant noise in the short-term.[15] Finally, we note that although the AJG/ABS list was published annually from 2007 to 2010, we consider as a window the 2009–2012 period as journals take several months to be published and researchers may not immediately observe a research output. As a result, citations are likely to be lagging by few years.

$After_t$ takes a value of 1 for the citing years 2013–2018 and 0 otherwise. $Upgrade_i\_x\_After_t$ is the interaction term between the two variables and the measure of interest. Its coefficient shows how the difference between papers from upgraded and downgraded journals changes after the first four editions of the AJG/ABS list were published compared to before.

In addition to $Pair_{i,\,j}$ fixed effects, we also include $CitingYear_t$ in the estimations. The latter are citing year fixed effects, that is, a set of nineteen dummy variables, one for each year that we observe citations (2000–2018). Note that we include both the $Pair_{i,\,j}$ and the $CitingYear_t$ fixed effects for all estimations.

As *UKCites* are counts, we report results from a fixed effects Poisson estimator throughout. The disadvantage of the fixed effects Poisson is that it does not allow for overdispersion (since it forces the mean to be equal to the variance) in the data but delivers consistent estimates as long as the mean of the dependant variable is correctly specified (Gourieroux et al., 1984).

Another popular estimator for count data is the unconditional fixed effects Negative Binomial. This estimator accounts for overdispersion; however, it suffers from the incidental parameters problem, even though Allison and Waterman (2002) show in simulation studies that the resulting bias is small. Regardless, to provide robustness, our baseline results are estimated via both Poisson and Negative Binomial regressions. Moreover, we also perform fixed effects ordinary least squares (OLS) analysis as a reference point.[16] Finally, since we consider pairs of papers, we cluster standard errors at the pair level to account for possible serial correlation (Bertrand et al., 2004).

4.2. Testing information and signalling effects

4.2.1. Baseline specifications

To test the information and signalling effects outlined in Section 2.2, the empirical specification is the same as in Eq. (1). The difference is that we consider different dependant variables by decomposing *UKCites* by citations that are likely to either reveal an information or signalling effect or both.

Our analysis makes a simple yet essential assumption, namely, that authors who publish in AJG/ABS ranked journals, regardless of Grade-level, are more likely to actively use the AJG/ABS list than authors who publish in non-AJG/ABS listed journals. We expand the assumption to maintain that the above relationship also holds for each citing paper. That is, an AJG/ABS citing paper's authors are more likely to actively employ the list than a non-AJG/ABS citing paper's authors.[17]

We first considered three new citation variables: *UKCitesNonABS, UKCitesBooks* and *UKCitesABS*. Note that these citation counts are all mutually exclusive subsets of *UKCites*. Specifically, for a paper to be counted in *UKCitesNonABS* it needs to be in a journal that is not listed in the AJG/ABS 2010 or the 2015 list. Citations made by books or book series are included in the variable *UKCitesBooks*. Finally, for a citation to be included in *UKCitesABS* it needs to be in a journal listed in the AJG/ABS 2010. Subsequently, *UKCitesABS* can be decomposed by the Grade-level of the outlet. Specifically, we denote *UKCitesABS34* as citations made by journals ranked Grade-3 and above in the 2010 edition. We further denote *UKCitesABS12* as citations made by journals ranked as Grade-2 or Grade-1 in the 2010 edition. As *UKCitesABS12* + *UKCitesABS34* = *UKCitesABS*, we focus on the two former variables.

Starting with the less nuanced information effect, we should expect that scholars may become aware of papers from the upgraded journals simply due to the upgrading, whereas the visibility of papers from the downgraded journals may decrease after downgrading. This implies that researchers, regardless of the outlet that publish, will increase their references to these highly ranked journals. This citing behaviour should also transcend researchers that publish in non-AJG/ABS journals. While these researchers may not be aware or simply do not consider the list, they are more likely to read a paper from a highly ranked journal due to increased visibility. Perhaps the only caveat is for *UKCitesBooks*. Books, depending on the topic, may not pay as close attention to the recent literature and may be more likely to cite non-journal publications. Overall, we should expect for the information effect to be validated to estimate an increase in *UKCitesABS12, UKCitesABS34* and *UKCitesNonABS*.

The signalling effect is more nuanced. This behaviour predicts that researchers will cite predominantly highly ranked journals to signal their own paper's quality of work. This incentive is present when these authors wish to publish in a Grade-4 journal. However, it is entirely plausible that they may not get their first choice and therefore attempt to publish their work in Grade-3 journals. Therefore, we should expect for citations to Grade-4 and Grade-3 journals to increase.

Nevertheless, if we estimate an increase in *UKCitesABS34*, then this could be attributed to the information effect. However, researchers who wish to publish in lower-Grade AJG/ABS journals do not have the same incentives to cite Grade-4 and Grade-3 journals. On the contrary, the perception is that citing highly ranked research will not increase the chances of acceptance more than citing lower-ranked journals. Therefore, to find evidence for the signalling effect, we should expect an increase in *UKCitesABS34* and no effect in *UKCitesABS12*.

Note that by validating the signalling effect, we cast some doubt on the information effect. That is, the information effect should be prevalent in all journal outlets including Grade-1 and Grade-2 journals; therefore, to the extent that the signalling effect is validated, this indicates that the information effect is not valid for journals within the AJG/ABS list.

4.2.2. Test signalling effect by citing researcher's characteristics

To examine the signalling effect by researcher profile, we decompose the two variables of interest (*UKCitesABS34* and *UKCitesABS12*) by citing researchers' attributes and re-estimate Eq. (1) with these

[15] In a similar design, Furman and Stern (2011) allow for a window to capture potential unobserved announcement effects of papers certified and accessed via Biological Resource Centres (BRCs).

[16] Several studies in the area of research and follow-on innovation employ either OLS or Poisson estimators or both (Azoulay et al. 2010; Williams 2013).

[17] Of course, it is entirely plausible that authors that publish in AJG/ABS journals, also publish in non-AJG/ABS journals. Our assumption only needs to hold for the average citing AJG/ABS and non-AJG/ABS citing article.

variables. We discuss in detail how each attribute is computed in Section 5.2.

5. Data construction and descriptive statistics

5.1. Data construction and description for testing the effect of journal upgrading on citations

For the focal twenty-four journals, we collected the records of every paper published from 2000 to 2003 on Scopus[18,19] which is one of the most comprehensive databases of scientific publications.[20] After excluding non-research articles such as conference papers, editorials and errata, we collected 4019 papers. Of these papers, 1820 belong in the upgraded group and 2199 in the downgraded group.

For each of these focal papers, we obtained bibliographic records for each of the approximately 200,000 papers that cite them. In particular, the variables of interest for the citing articles were publication year, author identifiers, author affiliations, type of publication (book or journal articles) and, if applicable, associated journal information (issn and title).

We excluded papers that received no *UKCites* up to 2018 as they cannot add any information. This data cut is mostly similar for the two groups: we dropped 27.5% and 31% of the papers from the upgraded and downgraded journals, respectively.[21] Finally, we excluded outliers, that is, papers that have cumulatively received more than a hundred *UKCites* up to 2018. The sample reduction is again similar, with 0.4% of papers excluded for both groups.

Overall, our unmatched sample includes 1494 and 1729 papers from upgraded and downgraded journals respectively. Table A2 in Online Appendix A shows the counts of papers by publication year for each journal. Table 1 – Panel A shows the summary statistics for selected citation measures by group and time period. Papers from downgraded journals receive on average 2.76 citations per year on or before 2008, whereas their upgraded counterparts receive 3.17. After 2012, this difference is even more substantial because papers from downgraded journals receive 5.23 citations per period while their upgraded counterparts 7.95 citations per period. Although this overall measure shows that these two types of papers are somewhat different even before 2009, this is not our measure of interest. *UKCites*, our focal dependant variable, shows that papers from downgraded journals receive more citations before 2009 and during the window compared to papers from upgraded journals. However, in the after-period, papers from upgraded journals appear to be cited more than the papers from downgraded journals.

Fig. 1.A plots the *UKCites* by group and supports the descriptive statistics. More importantly, this Figure is the most obvious evidence of a causal effect of the change in journal rankings on *UKCites*. Without even matching papers based on pre-2012 citation patterns, the two groups behave very similarly before 2012. If anything, papers from upgraded journals have lower UK citation rates, implying a lack of any pre-trend in the data. However, post-2012 the UK citation rates of the two groups start diverging, hinting to the influence of the upgrade and downgrade events.

Next, we performed CEM analysis on the baseline sample of the 3223 papers. The first year a paper's citations were observed is its publication year, while all papers' last observed citing year was 2018. Our goal is to match each paper from an upgraded journal published in year t to a paper published from a downgraded journal in year t based on the 'matching' variables, $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$. That is, for every paper from an upgraded journal, we first considered all papers from downgraded journals published in the same year, then matched a single paper from a downgraded journal closest to the values of $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$ with a paper from the upgraded journal.

The way the CEM algorithm works is to create strata for each matching variable. All possible combinations of strata of the matching variables comprise the set where each observation is assigned. For $\sum_{t=2000}^{2008} UKCites_t$, we generated six strata based on the 25th, 50th, 75th and 95th percentiles as cut-off values, and for $\sum_{t=2009}^{2012} UKCites_t$, we generated five strata based on the 50th, 75th and 95th percentiles as cut-off values. We did not include the 25th percentile since almost 50% of the values of this matching variable are 0. We then have thirty possible strata to which we can assign each paper from an upgraded journal. The algorithm then searches for papers from downgraded journals assigned to the same strata. In cases where there is more than one such paper, one is randomly selected.[22]

We opted for these two variables to ensure that the matched group does not exhibit any pre-trend in citation rates prior to 2012. Any pre-trend could blur the interpretation of our results in that an estimated increase could merely be attributed to the pre-trend rather than the effect of the upgrading *vis-à-vis* the downgrading. Furthermore, as can be seen from Table A1, the journals come from an array of disciplines (e.g., economics, management and finance). By matching each paper from an upgraded journal to the citation rates of at least ten years, we also net out any 'discipline' effects.

Of the 1494 papers in upgraded journals, we matched 1493 papers, thus achieving a 99.9% matching rate. Table 1 – Panel B shows the summary statistics for this matched sample. The variable *UKCites* is even more similar on the before and window period, while the difference between papers from upgraded and downgraded journals after 2012 is now larger. Fig. 1.B paints a similar picture. The summary statistics thus far provide substantial indication that papers from upgraded journals receive more *UKCites* than their downgraded counterparts after 2012, which is consistent with our initial conjecture of a Matthew effect for journals.

5.2. Additional data to test information and signalling effects

As we already collected the outlet for each of the roughly 200,000 citing papers, we can decompose *UKCites* to the following variables: *UKCitesNonABS, UKCitesBooks, UKCitesABS34* and *UKCitesABS12,* which comprise 25%, 24%, 24%, and 18% of *UKCites*, respectively.

Table A3 in Online Appendix A displays summary statistics by group and time period. These simple summary statistics demonstrate that non-ABS citations, either by books or journals, appear to be similar between the two groups. *UKCitesNonABS* appear to increase, though marginally, post-2012 for papers in the upgraded versus downgraded journals. The evidence, however, is more pronounced for the signalling effect: the difference of *UKCitesABS34* between upgraded and downgraded increases post-2012, whereas there is no significant change in the difference of *UKCitesABS12*.

[18] For automating the massive data retrieval process we used the rscopus R package (https://github.com/muschellij2/rscopus) and the Scopus API.

[19] www.scopus.com

[20] There is a large debate on which database for article records has more coverage (Falagas et al., 2008; Mongeon and Paul-Hus, 2016). There are arguments both in favour and against Scopus similarly to Web of Science (WoS). However, WoS has a smaller number of journals listed (Chadegani et al., 2013) than Scopus as its core focus is more on basic research (Goldfarb, 2008). Further, Kazakis et al. (2014) in a case study of researchers in medical schools found, according to their CVs, that Scopus listed all the publications of more than 90% of the authors. Also, Archambault et al., (2009) showed that there is high degree of correlation between Scopus and WoS; further, crucial to our econometric strategy, they also found high degree of correlation when looking at cross-country variation.

[21] Note that many of the papers that were dropped have citations, but no *UKCites*, which is the focal dependent variable. In fact, of the 4,019 papers only 49 have no citations.

[22] A process of how to implement the CEM procedure employing the routine by Blackwell et al., (2009) in cohorts of papers by citation rates is available upon request

Table 1
Summary statistics of unmatched and matched groups.

Unmatched Sample	Before (2000–2008) Downgraded	Upgraded	Window (2009–2012) Downgraded	Upgraded	After (2013–2018) Downgraded	Upgraded
N:	13,014	11,311	6916	5976	10,374	8964
CitesFull	2.76	3.17	5.11	6.88	5.23	7.95
	(4.10)	(4.28)	(7.10)	(8.76)	(8.43)	(12.31)
Cites	2.33	2.65	4.71	6.28	4.93	7.55
	(3.76)	(3.87)	(6.77)	(8.27)	(8.15)	(11.93)
UKCites	0.47	0.39	0.68	0.66	0.55	0.60
	(1.01)	(0.86)	(1.25)	(1.16)	(1.11)	(1.13)
USCites	0.61	0.89	1.08	1.78	0.94	1.68
	(1.39)	(1.67)	(2.30)	(2.91)	(2.36)	(3.19)
Matched Sample	Before (2000–2008) Downgraded	Upgraded	Window (2009–2012) Downgraded	Upgraded	After (2013–2018) Downgraded	Upgraded
N:	11,302	11,302	5972	5972	8958	8958
CitesFull	2.65	3.16	4.99	6.85	5.11	7.90
	(3.83)	(4.26)	(6.53)	(8.68)	(7.60)	(12.18)
Cites	2.23	2.64	4.60	6.26	4.80	7.51
	(3.51)	(3.86)	(6.22)	(8.20)	(7.33)	(11.81)
UKCites	0.41	0.39	0.65	0.66	0.52	0.60
	(0.92)	(0.86)	(1.17)	(1.16)	(1.03)	(1.12)
USCites	0.60	0.89	1.06	1.77	0.91	1.67
	(1.24)	(1.66)	(2.03)	(2.90)	(1.91)	(3.15)

Notes: The standard deviation is displayed below the average of each variable in parentheses. *CitesFull* is the number of total citations a paper receives per period (year). *Cites* is the number of total citations a paper receives per period after excluding self-author and self-journal citations. *UKCites* is the number of citations by UK affiliates a paper receives per period. A citation is included in the *UKCites* only if all authors are UK affiliated and disclose no other affiliation outside the UK. Note that *UKCites* is net of citations made by the same journal and/or by the same author(s). *USCites* is the number of citations by US affiliates a paper receives per period. A citation is included in the *USCites* only if all authors are US affiliated and disclose no other affiliation outside the United States. Note that *USCites* is net of citations made by the same journal and/or by the same author(s).

5.2.1. Additional data to examine signalling effect by citing researcher's characteristics

To examine the signalling effect by citing researchers' attributes, we download bibliographic information for the roughly 64,000 authors that cite the focal papers. As in Walker et al. (2018), we classified authors' institutions based on the 2014 Research Excellence Framework of the Unit of Assessment 19 (Business and Management). We assigned an institution as Top if it was ranked in the highest 20 institutions in the overall score and as Low or Middle (hereafter LM) if it was not ranked in these 20 institutions. We then decomposed *UKCitesABS34* to *TOP_UKCitesABS34* and *LM_UKCitesABS34*. *TOP_UKCitesABS34* are *UKCitesABS34* where all the authors are in a Top institution, and *LM_UKCitesABS34* are *UKCitesABS34* where at least one author is from an LM institution. Similar decomposition took place for *UKCitesABS12*.

Next, as it is impossible to observe the academic position held by the 64,000 authors in each of the years they were citing one of the focal papers, we classified them based on the year they first published. We therefore collected the publication history of each of the 64,000 authors. We classify a citing paper as junior if all the authors in the citing paper had their first publication within the preceding ten years, and we classify papers as senior if all the authors in the citing paper had their first publication more than ten years earlier. Finally, we classify papers as a mix for any other case. As before, we decomposed *UKCitesABS34* and *UKCitesABS12* by researcher's seniority.

To compile researcher productivity, there is no single criterion that is paramount. The h-index has been used extensively in the literature to classify researcher productivity (Kolympiris et al., 2019). Therefore, we opted to classify an individual researcher's productivity based on her h-index, acknowledging any shortcomings associated with this measure. The advantage of the h-index is that it places weight on the overall research output of a researcher as it values both the quantity of work and the scientific impact of the overall work (Hirsch, 2005).

For each citing author, we collected the fifteen most-cited articles as of 2018. Then for every author we manually calculated an h-index as of 2018. Note that we compiled the h-index manually by calculating papers published for every author until 2018 and all the citations each paper had accrued until 2018 from raw Scopus data. We then decomposed *UKCitesABS34* to *LowH_UKCitesABS34* if at least one author in the citing article had an h-index smaller than 15 and to *HighH_UKCitesABS34* if all the authors in the citing article had an h-index larger than 15. Similar decomposition took place for *UKCitesABS12*. The summary statistics of all the newly constructed variables are displayed in Table A4 in Online Appendix A.[23]

6. Results

6.1. Test the effect of journal upgrading on citations

Table 2 – Columns 1–3 show the results from the baseline matched sample. Column 1 is estimated via OLS, Column 2 via Negative Binomial and Column 3 via Poisson.[24] All estimators reveal qualitatively the same results. The coefficient of *Upgrade* is negative and significant, indicating that papers from upgraded journals receive fewer *UKCites*

[23] Note that we have performed a similar decomposition for *UKCitesNonABS*. For completeness, in the Online Appendix A we also examine the information effect by citing researcher's characteristics.

[24] The coefficients from the OLS regressions can readily be interpreted; for instance, focusing on the coefficient of *Upgrade_x_After*, we can state that after 2012 papers from upgraded journals receive 0.0684 more *UKCites* per year compared to their downgraded counterparts. The coefficients of the Negative Binomial and Poisson need to be exponentiated to be interpreted as elasticities. Focusing on the same coefficient in Column 2, we can state that after 2012 papers from upgraded journals receive Exp(0.166)-1 = 0.18 = 18% more *UKCites* per year compared to papers from downgraded journals.

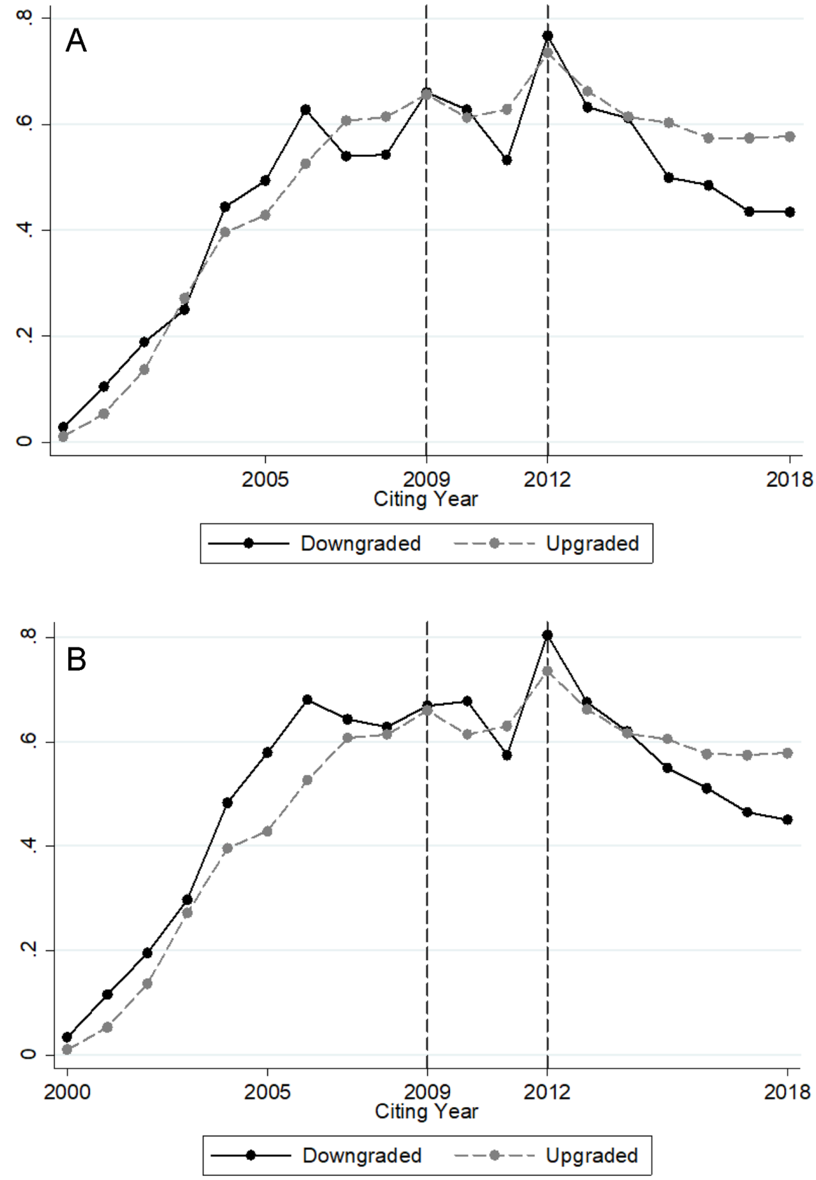

Fig. 1.A. *UKCites* by year and group. Unmatched sample.
Note: Each line shows the per-period average *UKCites* by upgraded and downgraded group, respectively.

before 2008 than their downgraded counterparts. The coefficient of interest, *Upgrade_x_After*, is positive and significant, which indicates that papers from upgraded journals receive more citations compared to papers from downgraded journals after 2012.

We should note that the coefficient of *Upgrade_x_Window* is also positive in the Negative Binomial and Poisson estimations. This is in line with our rationale of including a window where certain journals switch ranking early and others later. This interaction term was meant to capture any brief 'noisy' effect from the Grade changes. The coefficient of *Upgrade_x_Window* is at least four times smaller than the *Upgrade_x_After* and statistically significant only at 10% level; thus, it is not likely for this interaction term to imply any pre-trend in the data. Furthermore, the two coefficients are significantly different at the 1% level in both estimations.

To examine the dynamic effect, we re-estimated Eq. (1) via Poisson but instead of considering *Upgrade_x_Window* and *Upgrade_x_After*, we considered the interaction terms of *Upgrade* with each Citing Year dummy; that is, $Upgrade\_x\_CitingYear_{2000}$, $Upgrade\_x\_CitingYear_{2001}$, … and $Upgrade\_x\_CitingYear_{2018}$. We then plotted the coefficient estimates and their 95% confidence intervals in Fig. 2. As can be seen, the interaction terms before 2012 do not show a clear trend since they revolve around zero. However, starting in 2013, the coefficients of the interactions increase throughout 2018. This Figure does not indicate any pre-trend, though it also shows that the upgrade effect is not temporary and could increase even more in later years.[25]

Even though a pre-trend is not likely to drive the upgrade effect, to provide further robustness we performed a much stricter CEM procedure. Instead of matching on $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$, we matched based on $\sum_{t=2000}^{2002} UKCites_t$, $\sum_{t=2003}^{2004} UKCites_t$, $\sum_{t=2005}^{2006} UKCites_t$, $\sum_{t=2007}^{2008} UKCites_t$, $\sum_{t=2009}^{2010} UKCites_t$ and $\sum_{t=2011}^{2012} UKCites_t$. Given the large number of strata, the matching is poor ;[26] namely, we only matched 883 papers from upgraded journals, achieving a matching rate of 59%.

[25] This steady increase could also imply the diffusion of the list across UK universities, especially in the lead up to the Research Excellence Framework of 2014.

[26] This is a well-documented issue noted as the 'curse of dimensionality', that is, as the number of matching variables increases the number of strata makes matching between the two groups more challenging (Azoulay et al., 2010).

Table 2
Baseline results. Effect of the AJG/ABS upgrades/downgrades on *UKCites*.

	(1)	(2)	(3)	(4)	(5)	(6)
Upgrade	−0.0375***	−0.0539***	−0.0555***	−0.0189***	−0.0465***	−0.0491***
	(0.00711)	(0.0148)	(0.0156)	(0.00381)	(0.0154)	(0.0178)
Upgrade_x_Window	−0.00210	0.0324*	0.0331*	−0.0155**	0.00530	0.00925
	(0.0117)	(0.0190)	(0.0200)	(0.00724)	(0.0201)	(0.0225)
Upgrade_x_After	0.0684***	0.166***	0.161***	0.0501***	0.223***	0.217***
	(0.0197)	(0.0352)	(0.0357)	(0.0172)	(0.0544)	(0.0552)
USCites	0.0716***	0.0442***	0.0411***	0.0405***	0.0265***	0.0230**
	(0.00608)	(0.00498)	(0.00531)	(0.00558)	(0.00969)	(0.00972)
Observations	52,464	52,464	52,464	30,920	30,920	30,920
R-squared	0.405			0.395		

Notes: Columns 1-–3 include pairs of papers where each paper from an upgraded journal published between 2000 and 2003 is matched with a paper from a downgraded journal based on the same publication year and then on $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$ based on the CEM algorithm. Columns 4–6 include pairs of papers where each paper from an upgraded journal published between 2000 and 2003 is matched with a paper from a downgraded journal based on the same publication year and then on $\sum_{t=2000}^{2002} UKCites_t$, $\sum_{t=2003}^{2004} UKCites_t$, $\sum_{t=2005}^{2006} UKCites_t$, $\sum_{t=2007}^{2008} UKCites_t$, $\sum_{t=2009}^{2010} UKCites_t$ and $\sum_{t=2011}^{2012} UKCites_t$, based on the CEM algorithm. Columns 1 and 4 are estimated via OLS, columns 2 and 5 through a Negative Binomial and columns 3 and 6 via Poisson. All regressions include citing year fixed effect and paper pairs fixed effects. Standard errors are clustered at the paper pair level.

*** $p < 0.01$,.
** $p < 0.05$,.
* $p < 0.1$.

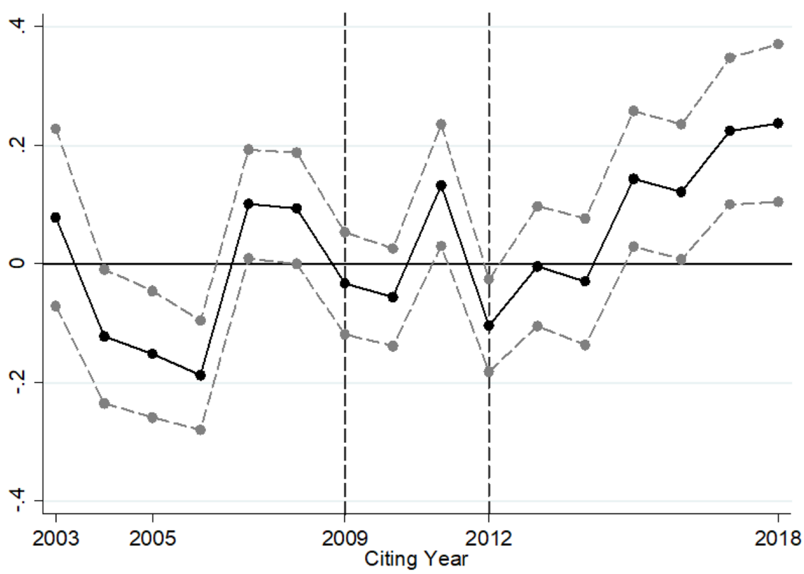

Fig. 2. Dynamic effects of upgrades on *UKCites*.
Notes: Effects of the upgrades on *UKCites*. We estimated Eq. (1) via Poisson regression but instead of considering *Upgrade_x_Window* and *Upgrade_x_After* we consider the interaction terms of *Upgrade* with each Citing Year dummy, that is, $Upgrade\_x\_CitingYear_{2000}$, $Upgrade\_x\_CitingYear_{2001}$,…, $Upgrade\_x\_CitingYear_{2018}$. The solid line represents estimates from these interaction terms. The 95% confidence interval (corresponding to robust standard errors, clustered around paper pairs) around these estimates is plotted with dashed lines.

However, the resulting pairs now have a much more similar citation pattern, as evident from Figure A1 in Online Appendix A. Results for this matched sample are displayed in Columns 4–6 of Table 2. For both the Negative Binomial and Poisson estimations, the coefficient of *Upgrade_x_Window* is even smaller than before and not statistically significant. Nonetheless, the coefficient of *Upgrade_x_After* has increased and remains significant at the 1% level.

6.1.1. Robustness

Note that in certain fields where the number of outlets is small, self-journal citations may be inevitable. Therefore, omitting them could lead to subject bias if such citations are numerous and display a different behaviour between downgraded and upgraded journals. To proof our results, we include self-journal citations back to *UKCites* and *USCites* and perform the same CEM and econometric analyses. A similar analysis to Table 2 is presented in Table A5 in Online Appendix A and shows qualitatively the same results.[27]

Also, we should note that the choice of papers from upgraded journals as the 'treated' group is random. To show that our results are invariant to the choice of group (upgraded vs. downgraded), we proceeded with an alternative matching procedure. Instead of matching a paper from an upgraded journal with a pool of papers from downgraded journals, we match each paper from a downgraded journal to a pool of papers from upgraded journals based on the same baseline criteria as in Table 2. For the variables of $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$, we matched 1717 papers from downgraded journals, achieving a matching rate of 99.3%. We also performed stricter matching by breaking down citation rates in more frequent time intervals and matched 945 papers

[27] For further robustness, we included self-author citations. Results are again similar and available upon request.

Table 3
Effect of the AJG/ABS upgrades/downgrades on *UKCites*. Match papers from downgraded to papers from upgraded journals.

	(1)	(2)	(3)	(4)	(5)	(6)
Upgrade	−0.0405***	−0.0465***	−0.0481***	−0.0173***	−0.0389**	−0.0399*
	(0.00742)	(0.0139)	(0.0148)	(0.00413)	(0.0180)	(0.0206)
Upgrade_x_Window	−0.0253**	0.00362	0.00478	−0.0158**	0.00524	0.00801
	(0.0120)	(0.0186)	(0.0200)	(0.00775)	(0.0205)	(0.0238)
Upgrade_x_After	0.0532***	0.145***	0.137***	0.0441***	0.207***	0.199***
	(0.0184)	(0.0325)	(0.0330)	(0.0167)	(0.0536)	(0.0551)
USCites	0.0819***	0.0370***	0.0325***	0.0410***	0.0307***	0.0274***
	(0.00566)	(0.00555)	(0.00532)	(0.00544)	(0.00913)	(0.00930)
Observations	60,178	60,178	60,178	33,082	33,082	33,082
R-squared	0.440			0.412		

Notes: Columns 1–3 include pairs of papers where each paper from a downgraded journal published between 2000 and 2003 is matched with a paper from an upgraded journal based on the same publication year and then on $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$, based on the CEM algorithm. Columns 4–6 include pairs of papers where each paper from a downgraded journal published between 2000 and 2003 is matched with a paper from an upgraded journal based on the same publication year and then on $\sum_{t=2000}^{2002} UKCites_t$, $\sum_{t=2003}^{2004} UKCites_t$, $\sum_{t=2005}^{2006} UKCites_t$, $\sum_{t=2007}^{2008} UKCites_t$, $\sum_{t=2009}^{2010} UKCites_t$ and $\sum_{t=2011}^{2012} UKCites_t$, based on the CEM algorithm. Columns 1 and 4 are estimated via OLS, columns 2 and 5 through a Negative Binomial and columns 3 and 6 via Poisson. All regressions include citing year fixed effect and paper pairs fixed effects. Standard errors are clustered at the paper pair level.

*** $p < 0.01$,.

** $p < 0.05$,.

* $p < 0.1$.

from downgraded journals, achieving a matching rate of 54.7%. Table 3 displays the results for these matched samples. The coefficient of *Upgrade_x_After* is positive and significant at the 1% level in all six specifications, whereas the coefficient of *Upgrade_x_Window* is not positive, and at the same time, statistically significant. Therefore, the choice of the matched group does not alter the qualitative results from the baseline analysis.

Another concern could be that the journals are in different fields and, therefore, part of the increase could be conflated with particular changes in a field. While our matching considers citation rates between matched papers for at least a decade; therefore, potentially netting out any field effects, it is useful to provide robustness. Note that due to the strict econometric design we only have twenty-four journals, hence, few degrees of freedom.

From a cursory review of the journals and the classification of the 2010 edition of the AJG/ABS list, we can assign journals in groups. *Brookings Papers on Economic Activity, European Economic Review, Journal of Development Economics, Journal of Health Economics* and *Review of Economics and Statistics* can be classified in an economics group and *British Journal of Sociology, Environment and Planning D: Society and Space, Industrial and Corporate Change, Journal of Law and Economics, Journal of Rural Studies, Regional Studies* and *World Development* can be classified in a social sciences group. The advantage of these two groups is that there are both upgraded and downgraded journals in each.

We considered only papers within each group to match papers based on publication year and $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$, employing the CEM method. Of the 238 papers from upgraded journals, we matched 236 papers in total. Note this is a much smaller sample, therefore, results should be viewed with caution. Table 4 – Columns 1–3 display the results. The coefficient of *Upgrade_x_After* for the OLS estimator is positive though not significant. However, for both the Negative Binomial and Poisson estimators, *Upgrade_x_After* is positive and significant at the 10% level. More importantly, in all cases, the magnitude of the coefficient is similar to Table 2, implying that the decrease in significance is due to the smaller sample size.

To provide further robustness, we considered the initial group of twenty four journals and exclude the journals in Psychology field for two reasons.[28] First, this field is the most distant from the rest of the journals, and second, all five journals included in the sample are upgrades; thus, we lack an appropriate control group for these journals. We then proceeded to a new matching procedure excluding these journals based on the same criteria as before. Of the 831 papers from upgraded journals, we matched them all. Table 4 – Columns 4–6 display the results. The coefficient of *Upgrade_x_After* is now positive and significant at the 1% level in all specifications. Although the coefficient of *Upgrade_x_Window* is positive and significant, it is not at the 1% level and, more importantly, the two coefficients *Upgrade_x_After* and *Upgrade_x_Window* are significantly different at the 1% level in all specifications.[29]

Finally, as we mentioned in Section 3, one could argue that the effect we are observing for *UKCites* could be attributed to other changes in these journals unrelated to the AJG/ABS list. There could be a few notable events. First, the University of Texas at Dallas (UTD) journal list is an important list for some UK universities. This list however is not influencing our results as none of our twenty-four journals is in the UTD list. Second, the journal list by the Financial Times (FT) is another list that academics may be employing. We review the FT's versions in 2007 and 2010 and observe if any journals are included. Only the *Journal of Consumer Psychology* shows up in the list; in fact, it is consistent with an upgrade. While it is not in the 2007 list it shows up in the 2010 list. However, in the previous Table we have already excluded the Psychology journals, including the latter, and results remain robust.

While these lists do not influence our results, there could be some other event, concurrent with the AJG/ABS list. While not likely, we propose a strategy that provides evidence that the AJG/ABS list is the ‘culprit’ of the observed effect. If there is another event, then that should have influenced citations by academics outside the UK as well. We have already included *USCites* and therefore should have netted out any such event. However, we go a step further and examine citations that disclose both UK and non-UK authors; we denote these citations as *PartUKCites*.[30] For these citations, it is safe to assume that the literature

[28] These journals are *Journal of Consumer Psychology, Journal of Experimental Social Psychology, Journal of Occupational and Organizational Psychology, Journal of Organizational Behavior* and *Journal of Vocational Behavior*.

[29] It could also be the case that few journals could have been viewed favourably in the AJG/ABS list, or by authors within the UK, thereby distorting the effects. Possible suspects could be the *British Journal of Management* and the *British Journal of Sociology* due to a home-bias effect. We excluded these two journals and re-ran the analysis. Results are displayed in Table A6 in Online Appendix A. They are qualitatively similar with the previous results.

[30] From these citations, we excluded those that disclose at least one US-affiliated author, to avoid any collinearity between the dependent and independent variables. Results, when including these citations, are similar and available upon request.

Table 4
Effect of the AJG/ABS upgrades/downgrades on *UKCites*. Considering groups of journals.

	(1)	(2)	(3)	(4)	(5)	(6)
Upgrade	−0.0217	0.00965	0.0120	−0.0221**	−0.0310	−0.0321
	(0.0210)	(0.0323)	(0.0325)	(0.0106)	(0.0194)	(0.0205)
Upgrade_x_Window	0.00352	0.0257	0.0385	0.0181	0.0505**	0.0481*
	(0.0401)	(0.0472)	(0.0501)	(0.0184)	(0.0242)	(0.0259)
Upgrade_x_After	0.0660	0.142*	0.141*	0.110***	0.202***	0.196***
	(0.0562)	(0.0827)	(0.0855)	(0.0280)	(0.0436)	(0.0441)
USCites	0.125***	0.0872***	0.0774***	0.0817***	0.0504***	0.0464***
	(0.0178)	(0.0153)	(0.0163)	(0.00756)	(0.00703)	(0.00706)
Observations	8344	8344	8344	29,366	29,366	29,366
R-squared	0.362			0.406		

Notes: Columns 1–3 include papers from two groups: (1) the economics group: *Brookings Papers on Economic Activity, European Economic Review, Journal of Development Economics, Journal of Health Economics* and *Review of Economics and Statistics* and (2) the social sciences group: *British Journal of Sociology, Environment and Planning D: Society and Space, Industrial and Corporate Change, Journal of Law and Economics, Journal of Rural Studies, Regional Studies* and *World Development*. A paper from an upgraded journal is matched to a paper from a downgraded journal based on the same publication year and then on $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$, based on the CEM algorithm from a paper from its own group. Columns 4–6 exclude all the journals in the Psychology field: *Journal of Consumer Psychology, Journal of Experimental Social Psychology, Journal of Occupational and Organizational Psychology, Journal of Organizational behaviour* and *Journal of Vocational behaviour*. We then treated the rest of the papers as one group and performed matching as previously. All columns are estimated via Poisson. Standard errors are clustered at the paper pair level.

*** $p < 0.01$,.

** $p < 0.05$,.

* $p < 0.1$.

review and any citations could have been added either by the UK or the non-UK author(s). The non-UK authors may have been influenced by the AJG/ABS but later than 2010. Bundling the citing behaviour by non-UK and UK authors we should expect any effect to be lagging by a few years. However, if there is another change that took place, then we should expect for *PartUKCites* to display the same behaviour as the *UKCites*.

Overall, of the initial sample of the 4019 papers, 1211 papers have at least non-zero *PartUKCites*. Of these, 475 papers have been upgraded. We then perform the same CEM procedure as previously for this group of papers. Results are displayed in Table 5 – Columns 1–3. The coefficient of *Upgrade_x_After* is insignificant, consistent with the premise that there was no other event that increased citations.

To explore whether there is lagging effect of the list, we decomposed the *After* dummy into two dummies: *After13_15*, which takes a value of 1 for the citing years 2013–2015 and 0 otherwise and *After16_18*, which takes a value of 1 for the citing years 2016–2018 and 0 otherwise. We then included the appropriate interaction terms in Columns 4–6. We observed that there is an increase in *PartUKCites* in the later years 2016–2018. For reference, we performed the estimations for the same specification for our baseline sample in Columns 7–9. Results show that the increase begins earlier, consistent with Fig. 2. This comparison of the dynamic effect shows that there was no concurrent event 'masked' with the upgrade or downgrade changes. Furthermore, the results show that any diffusion of the list in other countries is lagging the UK citing behaviour.

6.2. Test information and signalling effects

6.2.1. Baseline results

Table 6 tests the information and signalling effects, estimating Eq. (1) for citations by outlet. Columns 1 considers *UKCitesNonABS*. The coefficient of *Upgrade_x_After* is positive and significant; however, the coefficient of *Upgrade_x_Window* is positive and, unlike before, is significant at the 1% level. The two coefficients are statistically different at the 10% level, indicating that there is a weak statistically significant additional effect post-2012. In Column 2, when we consider *UKCitesBooks*, the coefficient of *Upgrade_x_After* is insignificant, indicating that books were not influenced by the Grade changes that occurred in the AJG/ABS lists. This latter finding is consistent with our conjecture that books have different citing behaviour than journals.

In Columns 3 and 4, we examine the response by *UKCitesABS34* and *UKCitesABS12*, respectively. These findings corroborate the signalling effect. The coefficient of *Upgrade_x_After* for *UKCitesABS34* is positive and significant at the 1% level, while it is insignificant for *UKCitesABS12*. More importantly, the coefficient of *Upgrade_x_Window* is positive for the case of *UKCitesABS34* but is only significant at the 10% level, and the two coefficients are statistically different at the 1% level.

This analysis provides strong evidence in support of the signalling effect. However, for the information effect, the lack of increase for *UKCitesABS12* does not conclusively support the existence of an information effect. We do find some, albeit weak, evidence that a journal ranking can transcend beyond listed journals, as evident by the increase of the *UKCitesNonABS*.

6.2.2. Heterogeneity of the signalling effect by citing researcher's characteristics

Table 7 examines the signalling effect by citing researcher's affiliation, seniority and productivity.[31] Panel A tests for the signalling effect by institution. For both types of institutions, the coefficient of *Upgrade_x_After* is positive and statistically significant at the 1% level (Columns 1–2), whereas there is no effect for Grade-1 and Grade-2 journals (Columns 3–4). These results indicate that researchers who publish in AJG/ABS journals, regardless of institution, display a strong signalling effect.

Panel B shows a significant difference in the signalling effect by seniority. While both senior and junior researchers display a positive effect post-2012, senior researchers also display a similar effect from 2009 to 2012, providing no clear evidence for this group. Overall, we find that junior researchers display a signalling effect, while for senior researchers the effect is not as strong. This could imply that senior researchers may be less likely keep track of all Grade changes the same way that junior researchers do.[32]

[31] For completeness, we also examined the effect on *UKCitesNonABS* by these researcher attributes. Results are displayed in Table A7 in Online Appendix A. Overall, we do not find any strong effect exclusively for the *Upgrade_x_After* coefficient, providing mixed evidence on any effect post-2012.

[32] We also performed the same analysis for citing papers that disclose senior and junior researchers. Results are displayed in Table A8 in Online Appendix A and resemble more the results by junior researchers. This result is intuitive since the junior member of the research team may track Grade changes.

Table 5
Effect of the AJG/ABS upgrades/downgrades on *PartUKCites*. Compare dynamic effects between *PartUKCites* and *UKCites*.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Upgrade	−0.00211***	−0.0111	−0.0111	−0.00210***	−0.0110	−0.0110	−0.0375***	−0.0538***	−0.0555***
	(0.000795)	(0.00874)	(0.00880)	(0.000794)	(0.00873)	(0.00873)	(0.00711)	(0.0148)	(0.0156)
Upgrade_x_Window	−0.00760***	−0.0507**	−0.0505**	−0.00759***	−0.0502**	−0.0502**	−0.00208	0.0325*	0.0332*
	(0.00210)	(0.0226)	(0.0229)	(0.00210)	(0.0228)	(0.0228)	(0.0117)	(0.0190)	(0.0200)
Upgrade_x_After	−0.0126	−0.0726	−0.0723						
	(0.00999)	(0.0667)	(0.0667)						
Upgrade_x_After13_15				−0.0584***	−0.395***	−0.395***	0.0316	0.0946**	0.0875**
				(0.0153)	(0.106)	(0.106)	(0.0225)	(0.0370)	(0.0374)
Upgrade_x_After16_18				0.0333**	0.231**	0.231**	0.105***	0.250***	0.248***
				(0.0144)	(0.0954)	(0.0955)	(0.0247)	(0.0475)	(0.0482)
USCites	0.00696***	0.0353***	0.0353***	0.00695***	0.0350***	0.0350***	0.0716***	0.0441***	0.0410***
	(0.00106)	(0.00579)	(0.00582)	(0.00106)	(0.00575)	(0.00575)	(0.00608)	(0.00499)	(0.00533)
Observations	16,732	16,732	16,732	16,732	16,732	16,732	52,464	52,464	52,464
R-squared	0.059			0.061			0.405		

Notes: Columns 1–6 consider the sample of papers where they have at least one citation included in *PartUKCites*. They include pairs of papers where each paper from an upgraded journal published between 2000 and 2003 is matched with a paper from a downgraded journal based on the same publication year and then on $\sum_{t=2000}^{2008} PartUKCites_t$ and $\sum_{t=2009}^{2012} PartUKCites_t$, based on the CEM algorithm. *After13_15* takes the value of 1 for the citing years 2013–2015 and 0 otherwise. *After16_18* takes the value of 1 for the citing years 2016–2018 and 0 otherwise. Columns 7–9 include the baseline matched sample of Table 2 (columns 1–3). Columns 1, 4 and 7 are estimated via OLS, columns 2, 5 and 8 through a Negative Binomial and columns 3, 6 and 9 via Poisson. All regressions include citing year fixed effect and paper pairs fixed effects. Standard errors are clustered at the paper pair level.

*** $p < 0.01$,.

** $p < 0.05$,.

* $p < 0.1$.

Table 6
Testing information and signalling effects.

VARIABLES	(1) *UKCitesNonABS*	(2) *UKCitesBooks*	(3) *UKCitesABS34*	(4) *UKCitesABS12*
Upgrade	−0.368***	0.0833*	−0.180***	0.310***
	(0.0593)	(0.0495)	(0.0507)	(0.0551)
Upgrade_x_Window	0.197***	−0.131**	0.127*	0.0318
	(0.0608)	(0.0592)	(0.0652)	(0.0691)
Upgrade_x_After	0.328***	−0.0132	0.392***	−0.118
	(0.0668)	(0.0697)	(0.0665)	(0.0733)
USCites	0.0315***	0.0261***	0.0533***	0.0527***
	(0.00663)	(0.00795)	(0.00905)	(0.00703)
Observations	52,464	52,464	52,464	52,464

Notes: All columns are estimated via Poisson. All columns include the baseline matched sample: pairs of papers where each paper from an upgraded journal published between 2000 and 2003 is matched with a paper from a downgraded journal based on the same publication year and then on $\sum_{t=2000}^{2008} UKCites_t$ and $\sum_{t=2009}^{2012} UKCites_t$ based on the CEM algorithm. All regressions include citing year fixed effect and paper pairs fixed effects. Standard errors are clustered at the paper pair level.

*** $p < 0.01$,.

** $p < 0.05$,.

* $p < 0.1$.

Finally, in Panel C, we observe that both productivity types of citing researchers display the signalling effect as we observe a significant increase post-2012 in Grade-3 and Grade-4 journals and no increase in the Grade-1 and Grade-2 journals.[33,34]

Overall, results show that all authors who publish in AJG/ABS journals, with the exception of senior researchers, may to an extent engage in signalling behaviour with respect to citing prior literature.

[33] To ensure that these results are invariant to the cut-off choice of 15 for the h-index, we considered different cut-off groups. First, we classify as low productivity citing papers that disclose at least one author with an h-index smaller than five (h-index ≤ 5). Results are displayed in Columns 1 and 3 of Table A9 in Online Appendix A. Second, we classify as low productivity citing papers those with least one author with 5 ≤ h-index < 15 and at least one author has an h-index larger than 15. Results are displayed in Columns 2 and 4 of Table A9 in Online Appendix A. Regardless of cut-off value, all researchers display similar signalling behaviours.

[34] We also computed the h-index based on the research output of an author at the time of citing the paper, not as of 2018. A counterpart of Table 7 – Panel C is estimated in Table A10 in Online Appendix A.

7. Discussion

7.1. Alternative explanations and relevance to past research

The results showed that journal upgrading will result in a citation uptick to its papers, a finding consistent with a Matthew effect for journals. The intuitive next step was to examine through which channels researchers react to this change in the perceived status of a journal. The first, an information channel, indicates that for upgraded journals, the visibility, readability and thereby their citedness will increase. Such behaviour of increased visibility accords with recent evidence of the random ordering of NBER papers mentioned above (Feenberg et al., 2017). We find some evidence to support this effect beyond the AJG/ABS-listed journals but not within.

The second, signalling channel, predicts that researchers will use the upgraded status of a journal to signal their own paper's research quality. This channel is a deliberate behaviour by researchers in contrast to the information channel, which is mostly unintentional. We find strong evidence validating such behaviour. We should stress that our empirical finding regarding signalling is not oblivious to alternative explanations.

Table 7
Testing signalling effect by citing researchers' characteristics.

Panel A: By institution				
	(1)	(2)	(3)	(4)
VARIABLES	*LM_UKCitesABS34*	*TOP_UKCitesABS34*	*LM_UKCitesABS12*	*TOP_UKCitesABS12*
Upgrade	−0.0865	−0.337***	0.468***	−0.135
	(0.0617)	(0.0900)	(0.0622)	(0.101)
Upgrade_x_Window	0.123	0.168	0.102	−0.0691
	(0.0794)	(0.119)	(0.0834)	(0.145)
Upgrade_x_After	0.435***	0.312***	−0.0983	−0.191
	(0.0804)	(0.114)	(0.0872)	(0.152)
USCites	0.0488***	0.0644***	0.0405***	0.0866***
	(0.00707)	(0.0161)	(0.00785)	(0.0115)
Observations	52,464	52,464	52,464	52,464

Panel B. By seniority				
	(1)	(2)	(3)	(4)
VARIABLES	*Junior_UKCitesABS34*	*Senior_UKCitesABS34*	*Junior_UKCitesABS12*	*Senior_UKCitesABS12*
Upgrade	−0.185**	−0.377***	0.486***	−0.0816
	(0.0838)	(0.0982)	(0.0749)	(0.114)
Upgrade_x_Window	0.0824	0.383***	0.0764	0.155
	(0.126)	(0.130)	(0.118)	(0.163)
Upgrade_x_After	0.263**	0.414***	−0.272**	0.0211
	(0.126)	(0.123)	(0.118)	(0.179)
USCites	0.0479***	0.0326***	0.0533***	0.0275**
	(0.0115)	(0.00859)	(0.0101)	(0.0136)
Observations	52,464	52,464	52,464	52,464

Panel C. By productivity				
	(1)	(2)	(3)	(4)
VARIABLES	*LowH_UKCitesABS34*	*HighH_UKCitesABS34*	*LowH_UKCitesABS12*	*HighH_UKCitesABS12*
Upgrade	−0.241***	−0.108	0.290***	0.255**
	(0.0720)	(0.0842)	(0.0664)	(0.112)
Upgrade_x_Window	0.113	0.130	0.0181	−0.0983
	(0.0932)	(0.137)	(0.0878)	(0.207)
Upgrade_x_After	0.396***	0.569***	−0.122	0.111
	(0.0886)	(0.157)	(0.0859)	(0.278)
USCites	0.0523***	0.0558***	0.0514***	0.0454***
	(0.0110)	(0.0107)	(0.00784)	(0.0160)
Observations	52,464	52,464	52,464	52,464

Notes: *LM_UKCitesABS34* and *TOP_UKCitesABS34* are subsets of *UKCitesABS34*. A citation from *UKCitesABS34* is included in *LM_UKCitesABS34* if at least one of the authors is from an LM institution. A citation from *UKCitesABS34* is included in *TOP_UKCitesABS34* if all the authors are from a TOP institution. An analogous definition applies to the rest of the variables of Panel A.
Junior_UKCitesABS34 and *Senior_UKCitesABS34* are subsets of *UKCitesABS34*. A citation from *UKCitesABS34* is included in *Junior_UKCitesABS34* if all the authors are classified as junior. A citation from *UKCitesABS34* is included in *Senior_UKCitesABS34* if all the authors are classified as senior. An analogous definition applies to the rest of the variables of Panel B.
LowH_UKCitesABS34 and *HighH_UKCitesABS34* are subsets of *UKCitesABS34*. A citation from *UKCitesABS34* is included in *LowH_UKCitesABS34* if at least one of the authors is classified as low productivity (h-index < 15). A citation from *UKCitesABS34* is included in *HighH_UKCitesABS34* if all the authors are classified as high productivity (h-index ≥ 15). An analogous definition applies to the rest of the variables of Panel C.
All columns have been estimated via Poisson. All regressions include citing year fixed effect and paper pairs fixed effects. Standard errors are clustered at the paper pair level.
*** $p < 0.01$,.
** $p < 0.05$, *$p < 0.1$.

Without considering them, we could leave the impression of Hypothesizing After the Results are Known or HARKing (Kerr, 1998; Bedeian et al., 2010; Starbuck, 2015). To avoid this, we consider alternative explanations, other than signalling, that could give rise to similar results.

First, as we already assumed, the authors publishing in AJG/ABS-journals may be more sensitive to rankings than non-AJG/ABS- journal authors. Therefore, they may simply react to higher attention Grade-4 journals. This could be interpreted as an unintentional information effect rather than a signalling effect. However, we should then expect an increase in *UKCitesABS12* as well, something that is not supported in the data. In brief, we found support for the signalling effect because both results hold, an increase in *UKCitesABS34* and no change in *UKCitesABS12*.

Second, it is entirely plausible that researchers aiming to publish their work in Grade-4 journals get rejected and then fail to publish again in Grade-3 journals. As a result, they publish their research in Grade-2 or Grade-1 journals. If this is a dominant behaviour, then such bias should work against validating the signalling effect. In other words, we should have also observed an increase in *UKCitesABS12*. It could be the case, however, that these researchers exclude Grade-4 journal references as they go down the list. Such behaviour to our knowledge is not common. Deleting a few references may be plausible but deleting a group of references could result in deletion of sentences and perhaps rewording of paragraphs for consistency. To the contrary, researchers may add citations after a rejection due to comments received or to make the manuscript more relevant to the new targeted outlet. If, however, researchers systematically delete references, then they should also delete the references of Grade-3 journals once they resort to publishing in Grade-2 journals, a painstaking process for the authors. Only if these, arguably unusual, conditions hold, then we would still have estimated a no effect in *UKCitesABS12*. However, this behaviour is also consistent with signalling, that is, the researcher only cites research relevant to the intended outlets for publication.

Our findings add to the recent study by Hussain et al. (2019). In their paper, the authors examined references of papers published in six accounting journals from 2002 to 2013. They posited that authors may signal the impact and quality of their work to editors and potential reviewers by citing appropriate research. In addition to evidence specific to the accounting field,[35] Hussain et al. (2019) found that the references these papers were citing, over time, increased towards papers published in Grade-3 and Grade-4 journals. This consistent signalling behaviour accords with works by Endenich and Trapp (2018a, 2018b), who examined signalling behaviour from the editors' viewpoint.

By examining forward citations, we can separate the signalling effect from a potential information effect, where in previous works it has not been examined simultaneously. Also, we substantiate this behaviour across a wider array of fields in addition to accounting, which was examined in Hussain et al. (2019). Finally, we were able to examine the signalling effect by citing researcher's attributes. The latter analysis embarks from the survey by Walker et al. (2018) where they examined, amongst other factors, seniority, productivity and affiliation. We find strong evidence to support that all types of researchers, except for senior academics, engage in signalling. These findings do not contrast with Walker et al., as we condition on researchers who target AJG/ABS-listed journals.

7.2. Policy implications

Citing practices, as those examined in this paper, are not likely to hamper the quality of research work.[36] Nonetheless, these distortions in citing patterns can have ramifications for research policy. Since the work by Garfield (1955), citations over the years have been used to measure scientific impact. Following his work, scholars, over the years, have highlighted several shortcomings (Price de Sola 1965; Gläser and Laudel, 2007; Macdonald and Kam, 2010; Hamermesh 2018). Despite these drawbacks, studies have shown that paper citations can still approximate scientific impact.[37] From a practical viewpoint, they leave a 'paper trail' and provide a wealth of readily accessible information. However, to the extent that citing patterns can be influenced by journal rankings, they can dilute their importance as a measure of scientific impact.

We should also highlight two additional potential negative implications of our findings. Though we do not formally test for them, their possibility adds further concerns that journal rankings can distort citation patterns. First, throughout the paper, we referred to an increase in citations due to journal upgrading. However, it could be the case that the effect is mainly driven by a decrease in citations of the downgraded journals. This is entirely plausible: downgrades could be more publicized than upgrades. If that is the case, then there might be a reverse Matthew effect for journals, namely, that papers in lower ranked journals receive a penalty in citations. Through this interpretation, our results could be viewed in a more negative way.

Second, the signalling effect could come at the expense of citing fewer works from lower-ranked journals or books. The recent study by Hussain et al. (2019) hints towards this. For the six journals the authors examined, Hussain et al. measured the percentage of references by type including journals and books. They found that citations to books have decreased over time. Therefore, the signalling effect could favour highly ranked journals at the expense of other works.

Our findings, and negative implications, relate to significant initiatives that point to distortions when journal-based metrics become the sole criterion for evaluation. The San Francisco Declaration of Research Assessment (DORA 2012), now signed by numerous organizations and individuals, highlights that the use of a journal's impact factor is accompanied with multiple deficiencies as a tool for research assessment. Instead, DORA brings the impact of any given research work to the forefront of research evaluation. DORA makes specific recommendations to all the parties involved to separate the journal metrics from the evaluation of research. Furthermore, the declaration stresses the importance of including qualitative indicators of research impact in addition to quantitative metrics. In a similar vein, the Leiden manifesto outlines ten principles to guide research evaluation. A running theme in many of these principles is that qualitative evaluation should be combined with quantitative metrics to flesh out the impact of a research work (Hicks et al., 2015).

Our paper's results, especially those pertaining to the signalling effect, relate to these initiatives. For our purposes, the first task is to identify the drivers of such behaviour. Two factors are the main culprits: i) the rising use of journal ranking lists, in this case the AJG/ABS, as the main tool for research assessment and recognition and ii) the perception that citing work in highly-ranked journals can increase the likelihood of acceptance in these journals. With respect to (i) the aforementioned initiatives have made efforts to separate each research work's impact from the outlet the work is published (DORA, 2012; Hicks et al., 2015; Wouters et al., 2019). While we do not offer additional recommendations, our findings support these initiatives pointing to the distortion of citations due to the overwhelming use of journal rankings.

With respect to (ii), the recommendations could be more nuanced. A first step towards resolving this issue would be explicit statements by editors that the reference list is not judged based on the outlets but based on relevance and individual merit. This, of course, will not be efficient in the short-run as authors will have nothing to lose by practising this behaviour; nonetheless, in the long run the culture may well change.[38] A probable experiment would be for researchers to submit their manuscripts for publication consideration with the full reference list but omitting the outlets. The journal information would appear only in the event of a publication. Reviewers, if they need to, can easily locate a specific reference with today's search engines. Furthermore, reviewers who judge that a relevant bibliography is missing can easily recommend to an author to include it. Such practices could convey to authors that they cannot employ signalling via their reference composition.

8. Conclusion

Journal rankings are used extensively by academics and research administrators. However, distortions are bound to arise when their use becomes pervasive, especially when employed to evaluate research.[39] Therefore, if the incentive structure for researchers is based heavily on the journal's ranking, then there is potential to abuse them.

[35] Hussain et al., (2019) examined whether papers in positivist journals draw from a narrower set of Grade-4 and Grade-3 journals compared to journals employing critical, interpretive and interdisciplinary methods.

[36] For a review of all the practices distorting the quality of research work see Martin (2013).

[37] For instance, in a study of patents, Roach and Cohen (2013) found that citations to papers approximate scientific impact better than citations to patents.

[38] One should not, however, underestimate such editorial statements. Blanco-Perez and Brodeur (2020) found that editorial statements targeted to referees made a material influence in accepting more papers that were not rejecting null hypotheses.

[39] Scholars have stressed that most papers published in prestigious journals will not reach high levels of citedness as a minority of papers within the same journal (Larivière et al. 2016; Callaway 2016). This is the well-known phenomenon of the extreme skewness of citations even within the same journal. The problem, however, begins when administrators do not consider this skewness and treat all papers from the same journal equally, thereby neglecting the impact of any given paper (Osterloh and Frey 2020).

A major unintended consequence is that researchers may cite papers because of the journals in which they are published rather their own merits. Given the importance of citations for evaluating scientific impact, such a distortion can have significant implications for science policy.

Providing causal evidence for this empirical question is challenging as we need to separate the journal's effect from the paper's underlying citedness. We employed a sample of journals in the early versions of the AJG/ABS list to provide identification. Our results showed that a journal's ranking indeed influences its papers' citations.

We proposed two explanations for these citation patterns: an information and a signalling effect. While the information effect is created unintentionally, the signalling effect is deliberate. We find more conclusive evidence for the signalling effect; therefore, we examined its heterogeneity by citing researcher's attributes. We find that researchers who publish in listed journals, regardless of institutional affiliation or productivity, will more likely signal their work by citing highly ranked work. In contrast to junior researchers, we find no evidence that senior researchers engage in signalling. This could indicate that senior researchers are less likely to keep track of all the upgrade or downgrade changes.

The signalling effect, identified here, is created due to i) increased incentives by researchers to publish in highly ranked journals[40] and ii) the underlying premise that this citation behaviour can increase, even marginally, the probability of acceptance. Both reasons have policy implications. As for increased incentives to publish in highly ranked journals, our paper supports recent initiatives and academic forums urging consideration of qualitative criteria in addition to quantitative metrics when evaluating research (DORA, 2012; Hicks et al., 2015). In a survey, Salter et al. (2017) found that academics prefer impact over publication. Therefore, aligning the metrics-based evaluation with intrinsic incentives can benefit academia in the long run.

For the premise that citing highly ranked publications can improve chances for acceptance, the research community could consider setting up an explicit contract between editors and researchers. Such a contract would highlight that citing research based on its relevance and merits should be the sole criterion for including it in the bibliography. There are already initiatives capable of supporting this issue. For instance, scholars and stakeholders across a variety of disciplines recently issued an open invitation to create a governing body to improve journal rankings and provide a roadmap for their responsible use (Wouters et al., 2019). Furthermore, several editors have already affirmed the ban of coercive citation practices and other such distorting practices.[41] Such an endeavour could potentially alter signalling practices in the long term. In the short run, we propose an experiment: for journals to ask for submissions with the full bibliography minus the outlet information. Outlets would only be disclosed in the event of journal publication. This could effectively cancel the incentive of signalling.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Christos Kolympiris, Vincent Larivière, Rossella Salandra, Ammon Salter, James Walker and Brian Wright for providing us with valuable comments in earlier drafts. All errors are our own.

[40] In analysis presented in Online Appendix B, we show that less productive researchers experience an even higher increase in citations if the journal is upgraded thereby amplifying their incentives to publish in such journals.

[41] See https://editorethics.uncc.edu/ (Accessed September 24, 2019)

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.respol.2020.103951.

References

ABS (2015) <https://charteredabs.org/academic-journal-guide-2015-view/> accessed 14 September 2019.
Aguinis, H., Cummings, C., Ramani, R.S., Cummings, T., 2020. An a is an A: The new bottom line for valuing academic research. Acad. Manage. Perspect.
Alberts, B., 2013. Editorial: Impact factor distortions. Science 340, 787.
Allison, P.D., Waterman, R.P., 2002. Fixed–effects negative binomial regression models. Sociol. Methodol. 32 (1), 247–265.
Amabile, T., 1998. How to kill creativity. Harv. Bus. Rev. 76, 76–87.
Archambault, É., Campbell, D., Gingras, Y., Larivière, V., 2009. Comparing bibliometric statistics obtained from the web of science and scopus. J. Am. Soc. Inf. Sci. Technol. 60 (7), 1320–1326.
Ashton, D., Beattie, V., Broadbent, J., Brooks, C., Draper, P., Ezzamel, M., Gwilliam, D., Hodgkinson, R., Hoskin, K., Pope, P., 2009. British research in accounting and finance (2001–2007): the 2008 research assessment exercise. Br. Account. Rev. 41 (4), 199–207.
Azoulay, P., Graff Zivin, J.S., Wang, J., 2010. Superstar extinction. Q. J. Econ. 125 (2), 549–589.
Azoulay, P., Stuart, T., Wang, Y., 2014. Matthew: Effect or fable. Manage. Sci. 60 (1), 92–109.
Baum, J.A.C., 2011. Free-Riding on power laws: questioning the validity of the impact factor as a measure of research quality in organization studies. Organization 18 (4), 449–466.
Bedeian, A.G., Taylor, S.G., Miller, A.N., 2010. Management science on the credibility bubble: cardinal sins and various misdemeanors. Acad. Manage. Learn. Educ. 9 (4), 715–725.
Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates. Q. J. Econ. 119 (1), 249–275.
Blackwell, M., Iacus, S., King, G., Porro, G., 2009. CEM: coarsened exact matching in stata. Stata J. 9 (4), 524–546.
Blanco-Perez, C., Brodeur, A., 2020. Publication bias and editorial statement on negative findings. Econ. J. Forthcoming.
Brooks, C., Fenton, E., Schopohl, L., Walker, J., 2019. Why does research in finance have so little impact. Crit. Perspect. Account 58, 24–52.
Brooks, C., Schopohl, L., 2018. Topics and trends in finance research: what is published, who publishes it and what gets cited. Br. Account. Rev. 50 (6), 615–637.
Callaway, E., 2016. Beat it, impact factor! publishing elite turns against controversial metric. Nature 535 (7611), 210–211.
Chadegani, A., Salehi, H., Yunus, M., Farhadi, H., Fooladi, M., Farhadi, M., Ale Ebrahim, N., 2013. A comparison between two main academic literature collections: Web of science and scopus databases. Asian Social Sci. 9 (5), 18–26.
Clark, B.R., 1998. Creating entrepreneurial universities: Organizational pathways of transformation. Issues in Higher Education. Elsevier Science Regional Sales, 665 Avenue of the Americas, New York, NY, pp. 10010.
DORA (San Francisco Declaration on Research Assessment), 2012.
Endenich, C., Trapp, R., 2018a. Signaling effects of scholarly profiles–The editorial teams of North American accounting association journals. Crit. Perspect. Account. 51, 4–23.
Endenich, C., Trapp, R., 2018b. Rejoinder: Toward an overarching signaling framework–The editorial teams of North American accounting association journals. Crit. Perspect. Account. 51 (C), 84–86.
Espeland, W.N., Sauder, M., 2007. Rankings and reactivity: How public measures recreate social worlds. Am. J. Sociol. 113 (1), 1–40.
Falagas, M.E., Pitsouni, E.I., Malietzis, G.A., Pappas, G., 2008. Comparison of PUBMED, Scopus, web of science, and Google scholar: strengths and weaknesses. FASEB J. 22 (2), 338–342.
Feenberg, D., Ganguli, I., Gaule, P., Gruber, J., 2017. It's good to be first: order bias in reading and citing NBER working papers. Rev. Econ. Stat. 99 (1), 32–39.
Furman, J.L., Stern, S., 2011. Climbing atop the shoulders of giants: the impact of institutions on cumulative research. Am. Econ. Rev. 101 (5), 1933–1963.
Garfield, E., 1955. Citation indexes for science: a new dimension in documentation through association of ideas. Science 122 (3159), 108–111.
Gläser, J., Laudel, G., 2007. 'The social construction of bibliometric evaluations. In: Whitley, R., Gläser, J. (Eds.), The Changing Governance of the Sciences. Springer, Dordrecht, pp. 101–123.
Goldfarb, B., 2008. The effect of government contracting on academic research: does the source of funding affect scientific output. Res. Policy 37 (1), 41–58.
Gourieroux, C., Monfort, A., Trognon, A., 1984. Pseudo maximum likelihood methods: applications to Poisson models. Econometrica: J. Econom. Soc. 701–720.
Hamermesh, D.S., 2018. Citations in economics: Measurement, uses, and impacts. J. Econ. Lit. 56 (1), 115–156.
Harvey, C., Morris, H., 2005. Classification of Academic Journals in the Field of Business and Management Studies. University of West of England, Bristol.
Harzing (2019) Harzing.com. <https://harzing.com/resources/journal-quality-list>accessed 10 June 2019.

Heckman, J.J., Moktan, S., 2018. Publishing and Promotion in economics: The Tyranny of the Top Five (No. w25093). National Bureau of Economic Research.
Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., Rafols, I., 2015. Bibliometrics: the Leiden Manifesto for research metrics. Nature News 520 (7548), 429.
Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. Proc. Natl. Acad. Sci. 102 (46), 16569–16572.
Hole, A.R., 2017. Ranking economics journals using data from a national research evaluation exercise. Oxford Bull. Econ. Stat. 79 (5), 621–636.
Hudson, J., 2013. Ranking journals. Econ. J. 123 (570), F202–F222.
Hussain, S., 2010. Accounting journals and the ABS quality ratings. Br. Account. Rev. 42 (1), 1–16.
Hussain, S., 2015. Journal list fetishism and the 'sign of 4'in the ABS guide: a question of trust. Organization 22 (1), 119–138.
Hussain, S., Liu, L., Miller, A., 2019. Accounting as a dichotomised discipline: an analysis of the source materials used in the construction of accounting articles. Crit. Perspect. Account.
Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: coarsened exact matching. Political Anal. 20 (1), 1–24.
Judge, T.A., Cable, D.M., Colbert, A.E., Rynes, S.L., 2007. What causes a management article to be cited: article, author or journal. Acad. Manag. J. 50 (3), 491–506.
Kalaitzidakis, P., Mamuneas, T.P., Stengos, T., 2003. Rankings of academic journals and institutions in economics. J. Eur. Econ. Assoc. 1 (6), 1346–1366.
Kalaitzidakis, P., Mamuneas, T.P., Stengos, T., 2011. An updated ranking of academic journals in economics. Can. J. Econ. 44 (4), 1525–1538.
Kazakis, N.A., Diamantidis, A.D., Fragidis, L.L., Lazarides, M.K., 2014. Evaluating the research performance of the Greek medical schools using bibliometrics. Scientometrics 98 (2), 1367–1384.
Kerr, N.L., 1998. HARKing: hypothesizing after the results are known. Pers. Soc. Psychol. Rev. 2 (3), 196–217.
Knothe, G., 2006. Comparative citation analysis of duplicate or highly related publications. J. Am. Soc. Inf. Sci. Technol. 57 (13), 1830–1839.
Kolympiris, C., Hoenen, S., Klein, P.G., 2019. Learning by seconding: evidence from national science foundation rotators. Org. Sci.
Laband, D.N., 2013. On the use and abuse of economics journal rankings. Econ. J. 123 (570), F223–F254.
Lane, J., 2010. Let's make science metrics more scientific. Nature 464 (7288), 488.
Larivière, V., Gingras, Y., 2010. The impact factor's Matthew effect: a natural experiment in bibliometrics. J. Am. Soc. Inf. Sci. Technol. 61 (2), 424–427.
Lariviere, V., Kiermer, V., MacCallum, C.J., McNutt, M., Patterson, M., Pulverer, B., Swaminathan, S., Taylor, S., Curry, S., 2016. A simple proposal for the publication of journal citation distributions. BioRxiv:062109.
Lozano, G.A., Larivière, V., Gingras, Y., 2012. The weakening relationship between the impact factor and papers' citations in the digital age. J. Am. Soc. Inf. Sci. Technol. 63 (11), 2140–2145.
Macdonald, S., Kam, J., 2010. Counting footnotes: citability in management studies. Scand. J. Manage. 26 (2), 189–203.
Martin, B.R., 2013. Whither research integrity? plagiarism, self-plagiarism and coercive citation in an age of research assessment. Res. Policy 42 (5), 1005–1014.
Merton, R., 1968. The Matthew effect in science: the reward and communication systems of science are considered. Science 159 (3810), 55–63.
Mingers, J., Watson, K., Scaparra, M.P., 2012. Estimating business and management journal quality from the 2008 research assessment exercise in the UK. Inf. Process. Manage. 48 (6), 1078–1093.
Mingers, J., Yang, L., 2017. Evaluating journal quality: a review of journal citation indicators and ranking in business and management. Eur. J. Oper. Res. 257 (1), 323–337.
Mongeon, P., Paul-Hus, A., 2016. The journal coverage of web of science and scopus: a comparative analysis. Scientometrics 106 (1), 213–228.
Morris, H., Harvey, C., Kelly, A., Rowlinson, M., 2011. Food for thought? a rejoinder on peer-review and RAE2008 evidence. Account. Educ. 20 (6), 561–573.
Osterloh, M., 2010. Governance by numbers. does it really work in research. Analyse Kritik 32 (2), 267–283.
Osterloh, M., Frey, B.S., 2015. Ranking games. Eval. Rev. 39 (1), 102–129.
Osterloh, M., Frey, B.S., 2020. How to avoid borrowed plumes in academia. Res. Policy 49 (1).
Parker, M., 2014. University ltd: changing a business school. Organization 21, 281–292.
Price de Sola, D., 1965. Networks of scientific papers. Science 149 (3683), 510–515.
Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A., 2012. How journal rankings can suppress interdisciplinary research: a comparison between innovation studies and business & management. Res. Policy 41 (7), 1262–1282.
Ritzberger, K., 2008. A ranking of journals in economics and related fields. German Econ. Rev. 9 (4), 402–430.
Roach, M., Cohen, W., 2013. Lens or prism? patent citations as a measure of knowledge flows from public research'. Manage. Sci. 59 (2), 504–525.
Salter, A., Salandra, R., Walker, J., 2017. Exploring preferences for impact versus publications among UK business and management academics. Res. Policy 46 (10), 1769–1782.
Salterio, S.E., 2018. Signaling effects of scholarly profiles-the editorial teams of North American accounting association journals: A personal perspective. Crit. Perspect. Account. 51, 78–83.
Sgroi, D., Oswald, A.J., 2013. How should peer-review panels behave. Econ.c J. 123 (570), F255–F278.
Spangenberg, J.F.A., Starmans, R., Bally, Y.W., Breemhaar, B., Nijhuis, F.J.N., 1990. Prediction of scientific performance in clinical medicine. Res. Policy 19, 239–255.
Spence, M., 1973. Job market signaling. Q. J. Econ. 87 (3), 355–374.
Starbuck, W.H., 2005. How much better are the most-prestigious journals? the statistics of academic publication. Org. Sci. 16 (2), 180–200.
Starbuck, W.H., 2015. Issues and trends in publishing behavioral science: a quarrelsome crew struggling with a disintegrating boat on a stormy sea. In: Welpe, I.M., Wollersheim, J., Ringelhan, S., Osterloh, M. (Eds.), Incentives and Performance: Governance of Research Organizations. Springer International Publishing, Cham.
Stensaker, B., Benner, M., 2013. Doomed to be entrepreneurial: Institutional transformation or institutional lock-ins of 'new' universities? Minerva 51 (4), 399–416.
Stephan, P.E., 1996. The economics of science. J. Econ. Lit. 34, 1199–1235.
Taylor, J., 2011. The assessment of research quality in uk universities: peer review or metrics. Br. J. Manage. 22 (2), 202–217.
Walker, J., Salter, A., Salandra, R., 2015. Initial Findings from the Survey of UK Business Academics 2015.
Walker, J.T., Fenton, E., Salter, A., Salandra, R., 2018. What influences business academics' use of the association of business schools (ABS) list? evidence from a survey of UK academics. Br. J. Manage.
Walker, J.T., Salter, A., Fontinha, R., Salandra, R., 2019. The impact of journal re-grading on perception of ranking systems: exploring the case of the academic journal guide and business and management scholars in the UK. Res. Eval.
Wilhite, A.W., Fong, E.A., 2012. Coercive citation in academic publishing. Science 335 (6068), 542–543.
Williams, H.L., 2013. Intellectual property rights and innovation: Evidence from the human genome. J. Polit. Econ. 121 (1), 1–27.
Willmott, H., 2011. Journal list fetishism and the perversion of scholarship: reactivity and the ABS list. Organization 18 (4), 429–442.
Wouters, P., Sugimoto, C.R., Larivière, V., McVeigh, M.E., Pulverer, B., de Rijcke, S., Waltman, L., 2019. Rethinking impact factors: better ways to judge a journal. Nature 569 (7758), 621–623.