

Eger, Thomas; Mertens, Armin; Scheufen, Marc

**Article — Published Version**

## Publication cultures and the citation impact of open access

Managerial and Decision Economics

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Eger, Thomas; Mertens, Armin; Scheufen, Marc (2021) : Publication cultures and the citation impact of open access, Managerial and Decision Economics, ISSN 1099-1468, Wiley, Hoboken, NJ, Vol. 42, Iss. 8, pp. 1980-1998, <https://doi.org/10.1002/mde.3429>

This Version is available at:

<https://hdl.handle.net/10419/284802>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>

## RESEARCH ARTICLE

## Publication cultures and the citation impact of open access

Thomas Eger<sup>1</sup> | Armin Mertens<sup>2</sup> | Marc Scheufen<sup>2,3</sup><sup>1</sup>University of Hamburg, Hamburg, Germany<sup>2</sup>Big Data Analytics, German Economic Institute (IW), Cologne, Germany<sup>3</sup>Ruhr-University Bochum, Bochum, Germany**Correspondence**

Thomas Eger, University of Hamburg, Hamburg, Germany.

Email: Thomas.eger@uni-hamburg.de

Marc Scheufen, Big Data Analytics, German Economic Institute (IW), Cologne, Germany.

Email: scheufen@iwkoeln.de;

marc.scheufen@rub.de

Does open access (OA) to journal articles foster citations to these articles? We compare the citation impact of gold and green OA in two disciplines: Biology and Economics & Management. The empirical analysis covers all articles of these disciplines included in the Web of Science “Journal Citation Reports” between 2000 and 2019. We show that, controlling for confounding variables pertaining to the journals and articles, gold OA increases citations across all articles. However, the individual disciplines feature starkly different effects: a 18.3% increase in Biology, compared to a decrease by 30.9% in Economics & Management. Also Green OA leads to an increase in citations to academic research. These results are confirmed by a number of robustness checks.

**JEL CLASSIFICATION**

L17; O33

## 1 | INTRODUCTION

With the advent of electronic publishing and the Internet, the traditional business model of academic publishing, based on subscription fees paid by the readers/libraries (closed access [CA] journals), has to some extent been replaced and to some extent complemented by different types of open access (OA). Regarding scholarly articles, two broad models of OA can be distinguished. *Gold* OA refers to scholarly journals that are financed not by subscription fees but by fees paid by the authors or their sponsors (so-called article-processing charges [APC]). Readers have free online access to such articles. A special case are *hybrid* OA (HOA) journals, where authors can either pay the APC and make their articles freely available, or not pay the APC, which means that readers will have to pay to access the full text. Issues of such journals typically comprise articles that are freely available as well as articles with a paywall.

The second broad model, *green* OA, refers to the posting of papers with a repository prior to or after their publication in a traditional journal. We may distinguish between institutional repositories, for example, on university websites, and subject-based repositories, such as PubMed Central for biomedical and life sciences, arXiv with a focus on physics, RePEc for economics, and

SSRN for the social sciences. Moreover, some authors post versions of their articles on their individual websites. The latter option, however, entails considerably larger search costs for potential readers compared to postings on institutional or subject-based websites.

The rise of the different types of OA triggered a discussion as to whether OA will foster citation counts, or more precisely, do OA articles receive systematically more citations than comparable CA articles? The answer is important to the authors of scholarly articles, especially in the natural and social sciences, as their reputation strongly hinges on citation counts—to some extent on their individual citation counts, but in particular on the average citation counts of the journals they publish in, as a proxy of the reputation of those journals within the corresponding discipline.<sup>1</sup>

While early studies, relying on cross-sectional data, found a large OA citation advantage of up to 580%, in more recent studies, the OA status constitutes only one of several explanatory variables in the regression to explain citation counts. If additional control variables are taken into account, such as the number of authors, self-citations, quality of the article, time of first availability of content, citation windows, and secular trends in citations, the calculated OA citation advantage turns out to be much smaller than in the early studies. The results still

-----  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Managerial and Decision Economics published by John Wiley & Sons Ltd.

vary widely: While many studies found an OA citation advantage of between 6% and more than 40%, others detected no OA citation advantage at all or even a negative one. These differences are partly due to the different objects of comparison. While some authors focus on OA and non-OA articles in the same HOA journal, others compare non-OA articles in traditional journals with pre- or post-prints in OA repositories, or they compare commercial online availability with OA. Moreover, there seem to be remarkable differences in OA citation advantages between different disciplines. Finally, different specifications of the econometric models with different confounding variables also yield different results.

In the following, we begin with a brief survey of the most important literature on the above question. Thereafter, we present the results of our own empirical research, focusing on systematic differences between two publishing cultures: Biology, a predominantly *gold culture* with strong OA journal usage (gold OA), and Economics & Management, a predominantly *green culture* with few OA journals but extensive use of OA repositories (green OA). The descriptive statistics reveal that compared to Biology, Economics & Management has a very small share of OA journals, with no OA journal among the discipline's Top 20 journals and the first few OA journals making it into the Top 50 only very recently. How come that OA journals enjoy so much more esteem in Biology than in Economics & Management? We think that the differences may be due to the timing of the establishment of OA journals in each discipline, combined with changes in the respective environment. Biology already featured three OA journals in 2000, which increased to 20 in 2019—about one-quarter of all journals in Biology. In October 2003, the Public Library of Science (PLOS) launched the OA journal *PLoS Biology*, which only 2 years later achieved an impact factor of 13.9, the highest in general biology (Suber, 2006, p. 151). There are several reasons why this OA journal succeeded so quickly and has been achieving excellent impact factors ever since (Eckdahl, 2004): (1) It was launched by outstanding experts in the field, allowing it to attract many excellent scholars as authors and reviewers; (2) this high level of quality in an area of great public interest has helped to attract large amounts of funding; (3) as a full-service journal, *PLoS Biology* also includes sections that communicate the results of high-profile original research to nonacademic audiences, such as physicians, patients, and policy makers. Thus, the early start of a very successful OA journal in Biology facilitated the establishment of subsequent OA journals in the field that were able to build on the strong reputation of *PLoS Biology*.

In Economics & Management, Web of Science (WoS) listed the first OA journal in 2000, and, following a period of slow growth, the number stood at 26 OA journals in 2019—4.69% of all journals in the field. In 2010, when the discipline featured just four OA journals (compared to 420 CA journals), Jeffrey Beall, a librarian at the University of Colorado in Denver, compiled a list of so-called predatory publishers (“Beall's List”). This term refers to publishers whose business model aims to either exploit inexperienced academic authors or allow experienced authors to publish low-quality articles in OA journals. The strategies employed by predatory publishers include charging APCs for articles with little or no peer review,

listing academics as members of editorial boards without their permission, appointing fake academics to editorial boards, mimicking the name or website of more established journals, and so on. Beall's List was attacked by the publishers concerned, as well as by OA advocates, who argued that Beall exaggerated the problems due to his ideological objection to OA (Anderson, 2015; Pinfield, 2015, p. 619), and it went offline in January 2017.<sup>2</sup> In any event, the publication of Beall's List likely impaired economists' perception of the reputation of OA journals. Consequently, the top authors in Economics & Management were deterred from submitting their papers to OA journals, so the negative expectations became self-fulfilling. In Biology, the rise of OA journals, with excellent quality, began already before the publication of Beall's List and led to a reinforcement of positive expectations.

We contribute to the existing literature in three respects: First, we compare two distinct publishing cultures—Biology (gold culture) and Economics & Management (green culture)—to find discipline-specific trends in the citation impact of OA publishing.<sup>3</sup> Second, we merge several different data sources to enrich the information on the article level, controlling for characteristics such as the publishing regime, the discipline regime and article information, as well as new measures, for example, author score. Third, we extend the methodology by applying a hurdle model for robustness to account for the large number of non-citations in our data. We find a highly significant effect of gold OA publishing on the number of citations of 18.3% in Biology. Interestingly, the publishing culture matters significantly, as OA articles in Economics have a large negative impact on citations, unlike in Biology. Regarding the green road of OA publishing, both disciplines reveal a positive impact on the number of citations. We check our results by applying several robustness checks.

## 2 | SURVEY OF THE LITERATURE

### 2.1 | Early studies

In the last 20 years, many contributions have measured the impact of OA on citation counts.<sup>4</sup> Using cross-sectional data, early studies found considerable OA citation advantages of up to several hundred percent. Here are a few examples: Lawrence (2001) analyzed 119,924 peer-reviewed conference articles in computer science and related areas published between 1989 and 1999, part of which were openly accessible online, while the others were available only in print. Focusing on 1,494 venues that contained at least five offline and five online articles, he found an average OA citation advantage of 336% (median of 158%). Harnad and Brody (2004) were the first to study the effects of green OA.<sup>5</sup> They compared 95,012 journal articles in physics and mathematics that were deposited as pre-prints in arXiv, an important natural science repository, to all the other 14 million articles that were published between 1992 and 2001 in the same journal and the same year as each of the green OA papers. They found average OA citation advantages between 250% and 580% across the years of publication.

Antelmann (2004) analyzed 2017 randomly selected articles published between 1999 and 2002 in high-impact journals across four disciplines (mathematics, engineering, political science, philosophy). She compared the citation counts of those articles whose full text is freely available online at a location other than the publisher's website to the CA articles, that is, all others. The mean OA citation advantage ranges from 45% in philosophy to 91% in mathematics. Hajjem et al. (2007) automatically analyzed more than 1.3 million articles across 10 disciplines published between 1992 and 2003. Between 5% and 16% of the articles were OA, depending on the discipline, year, and country. The authors compared the citation counts of the OA articles, that is, those whose full text was freely available online, to the CA articles, that is, articles that were only available in the subscription journal. They found average OA citation advantages between 36% and 172%, depending on the discipline.<sup>6</sup>

All of these studies assume that the differences in citation counts are primarily due to the articles' access status (see also Swan, 2010, p. 1). However, this assumption is far from plausible (Craig et al., 2007, p. 17). As a matter of fact, most citing authors are academics, who can access most of the journals offered by their institutions' libraries from their home computers. Consequently, a potential citing author does not care much whether her library has paid a subscription fee for a certain article or whether this article is freely available on the Internet, perhaps financed by an APC. All that matters is that the article is available somehow. However, with about 33,100 peer-reviewed English-language journals plus a further 9,400 non-English-language journals collectively publishing over 3 million articles a year as of mid-2018, even the richest universities cannot afford to subscribe to all journals.<sup>7</sup> With this in mind, rather than a general citation advantage of OA, we would expect a selective citation advantage, for two reasons. First, an OA citation advantage may persist with respect to articles published in less prominent journals that are not widely available in libraries and with respect to poorer universities with tight budgets and limited subscription potential.<sup>8</sup> Second, if the full text is not available, authors sometimes cite an article purely on the basis of the abstract. Facilitating access to the full text may lead to more informed citations, that is, citations to lower quality articles may be replaced with citations to articles of higher quality.<sup>9</sup> We would expect that article citation counts are in particular driven by factors such as the relevance of the topic (how many other authors deal with similar research questions?) and by the quality of the article (author reputation, journal ranking).

## 2.2 | More recent studies: The role of confounding variables

Measuring the impact of OA on citation counts is anything but trivial. Ideally, one would compare otherwise identical articles that only differ regarding their access status. Some studies therefore compare OA and CA articles in the same (issue of the same) journal, such as OA and CA articles in hybrid journals. This approach presumes that the articles in a given journal are sufficiently homogeneous.

However, this is unlikely to be the case. Even in the top journals with high average citation counts, only a few articles are cited very often, while most receive only a few citations, and many are never cited at all.<sup>10</sup> Thus, the distribution of citation counts is typically extremely skewed, which impairs the reliability of citation count averages across OA and CA articles within the same journal. Other studies refer to green OA and compare articles in traditional journals with and without pre- or post-prints posted in repositories. However, these articles are typically not homogeneous either as they are made public at different points in time or may differ in terms of content. Many publishers as the copyright holders condition their consent to green OA after publication on whether the archived article refers to the published, the accepted or the submitted version, and they may require embargo periods.<sup>11</sup>

In more recent studies of the determinants of citation counts, the OA status is only one of several explanatory variables. As a result, the OA citation advantage is typically much lower than in the early studies. In particular, two causes of this discrepancy are discussed.<sup>12</sup> The first is *selection bias*: Some studies assume that more prominent authors are more likely to publish their article OA and/or authors generally prefer to publish their better works OA. Selection bias implies reverse causality: Instead of the OA status causing citation counts, the perceived quality of the article and thereby the expected citation counts affect the probability of an article being published OA. To remedy this problem, the analysis should also control for author prominence, which could, for example, be proxied by prior citation history, publications in highly ranked journals, and the funding organization. To measure the quality of an article independently of its citation count is a challenge. Moreover, what is the rationale for the preference to publish only the best articles OA? This is by no means a trivial question. When an article that is (to be) published in a traditional journal can be made OA as a post-print or pre-print without additional costs to the author, why should they decide against OA? A quality bias could arise only if the repository employs some kind of quality filter, but this choice would not be based on the authors' self-selection. When an author is wondering whether to submit an article to a traditional journal or an OA journal, they will typically first of all consider the reputation of the journals. OA journals enjoy high reputation in some disciplines (e.g., Biology) but not in others (e.g., Economics).<sup>13</sup> In disciplines where OA is highly regarded, if the author's or the sponsor's APC budget is restricted, authors will tend to reserve their higher quality papers for OA journals—so we have a selection bias. The same holds when an article is accepted by an HOA journal, and the author must decide whether to pay the APC and publish OA. By contrast, when the author's institution automatically pays APCs for any article accepted by specific journals—as is the case, for example, with the consortium of German research institutions that have made summary deals with Wiley and Springer Nature (the “DEAL project”)—then this potential cause of selection bias vanishes.

A second potential source of discrepancy refers to the time when an article is first published. In some disciplines with long delays between submission and final publication of an article, it is common

practice to post a pre-print in a repository. If the pre-print is sufficiently similar to the published paper, the citation counts are subject to an *early view bias*, which must be distinguished from any effect of the OA status: Articles whose pre-prints are freely available long before their publication in a journal may receive more citations not because of their OA status per se but rather because of the early availability of the pre-print—in other words, due to mere *timing*. Thus, for a given window of time during which citations are counted, the longer the period between the upload of the manuscript and the publication of the final paper, the stronger the early view bias, which will also depend on the specific citation culture in each discipline. In some disciplines, such as traditionally Health Science and Biology, many journals used to follow the “Ingelfinger rule,” which bans the publication of articles that have previously been published elsewhere, so that pre-prints were typically not available. However, the application of the Ingelfinger rule appears to be eroding.<sup>14</sup> Journal publishers as the copyright holders restrict the self-archiving of post-prints in several ways (see the Sherpa/Romeo database<sup>15</sup>). Publishers often impose an embargo period of between 6 months and several years, which implies a *late view bias* (rather than an early view bias) for the OA version—provided that no self-archived pre-print is available. When comparing OA and CA articles in the same issue of an HOA journal, there is obviously no early or late view bias.

These methodological limitations apply not so much to more recent studies of the OA citation advantage that also consider other drivers of citation counts besides the OA status of an article. In the following, we discuss a selection of such studies. Eysenbach (2006) conducted a longitudinal bibliometric analysis of 1,492 articles published in the second half of 2004 in the “Proceedings of the National Academy of Sciences,” a multidisciplinary HOA journal. 14.2% of these articles were published OA. Eysenbach's logistic regression model controls for potentially confounding variables, such as the number of authors, the authors' lifetime publication counts and impact, number of days since publication, submission track (with different levels of rigor in peer review), the country of the corresponding author, funding organization, and discipline. He found an OA citation advantage of up to 42%, depending on the period over which the citations were counted. Furthermore, the number of authors and funding from competitive grants significantly predict the likelihood that an article is cited at least once within the applicable time window. A secondary analysis took into account that CA articles may also be self-archived in OA repositories, with the result that more openness is associated with higher citation counts.

Evans and Reimer (2009) investigated more than 26 million articles and the associated citations across 14 disciplines from the 8,253 most widely cited journals that came online since 1998. The citation data ends in 2005. The authors distinguish between commercial online availability and free online availability (i.e., OA). Controlling for commercial online availability, the average citation advantage of OA journals remains modest (about 8%). The influence of OA is more than twice as strong in the developing world than in the wealthy Northern and Western Hemisphere. However, it is weaker in the poorest countries with limited electronic access.

Gargouri et al. (2010) analyzed 27,197 articles from 1984 CA journals in several disciplines over the period 2002–2006. They compared citation counts (within the same journal/year) of CA articles to those of articles that were made OA as refereed post-prints, either by voluntary self-archiving or by mandatory self-archiving in the institutional repositories of four institutions. 6,215 of the articles originate from institutions with mandatory OA policies and 20,982 from institutions without such mandates. About 15% of the articles from institutions without mandate were voluntarily self-archived, while about 60% were self-archived when it was mandated. The logistic regression analysis controlled for article age, journal impact factor (JIF), number of coauthors, references, article length, discipline, article type, country, and other indicators. It found that OA articles are cited significantly more often than CA articles. Furthermore, there is no evidence that the citation advantage is smaller for mandated OA than for self-selected OA (i.e., there is no support for the hypothesis that authors tend to publish higher quality articles OA).

Gaulé and Maystre (2011) analyzed a sample of 4,388 biology papers published between May 2004 and March 2006 by the “Proceedings of the National Academy of Sciences,” an HOA journal. They compared the citation counts over the 2 years for those articles that were made OA in return for an APC of 1,000 USD (17% of the sample) to the remaining articles (CA).<sup>16</sup> The analysis controls for quality differences (selection effect) by taking into account characteristics of the last authors (typically the principal investigator in natural science) and article quality,<sup>17</sup> as well as for the availability of funds for the APC (proxied by publication in the last quarter of the financial year) and several other factors. The authors failed to find a significant difference in the citation counts.

Davis (2011) conducted a randomized controlled trial of 3,245 research articles and reviews published between January 2007 and February 2008 in 36 journals in various disciplines to explore the effects of OA on article downloads and citations. 712 articles (22%) were randomly assigned to the OA treatment group. The control group comprised the remaining 2,533 articles, which were only available to subscribing individuals and institutions.<sup>18</sup> The effects of the OA treatment on article downloads and citations were estimated by multivariate linear regression models that controlled for journal, article type and length, number of authors, and self-archiving of the articles. While making an article OA doubled the number of full-text downloads in the first year after publication, there was no effect on citation frequency within the first 3 years. Davis explains these results by social stratification: While the typical citing authors are affiliated with a small number of elite research universities with excellent access to the scientific literature and therefore do not depend on OA, many readers do not belong to the core research community. The latter are the real beneficiaries of OA to scientific literature.

McCabe and Snyder (2014) analyzed a sample of 100 top journals in Ecology, Botany, Multidisciplinary Science, and Biology between 1996 and 2005. They distinguished between *open access* and *online-availability* on the journal's own website or one of the major digital aggregators, such as JSTOR, EBSCO, ProQuest, Ingenta, Gale, and OCLC. The data set includes over 200,000 cited articles with 4.8

million cites contained in the over 8,000 WoS-indexed journals. Citations to the average journal volume peak in the second year after publication, and there is a significant secular upward trend in the sample citations. To control for unobserved quality and age and time effects, the authors used a Poisson estimator with volume fixed effects (FE), a journal-specific quadratic age profile of citation counts, and the secular trend in citations as controls. They obtained an OA citation advantage of 8.1%, the benefit being concentrated among the top-ranked journals, whereas the bottom-ranked journals receive significantly fewer cites. The authors explain this result by intensified competition for the readers' attention, which benefits high-quality articles at the detriment of low-quality articles.<sup>19</sup>

From the University of Michigan's institutional repository service "Deep Blue," Ottaviani (2016) randomly selected 3,850 papers that were made OA from 2006 onward via blanket licensing agreements between the publishers and the library. Thus, any self-selection bias is avoided. However, since the articles in the sample were only opened after some embargo period, there is a "late view bias" against the OA papers. These OA articles were matched with 89,895 CA articles from the corresponding journal issue. The original publication dates range from 1990 to 2013. The sample is dominated by physical science, health science, and engineering articles. Ottaviani focused on article-by-article differences<sup>20</sup> in citation counts, finding an OA citation advantage with a lower bound of approximately 20%. He also found that better (above median) articles gain more from being OA.

Mueller-Langer and Watt (2018) studied OA citation advantages with reference to HOA journals. They analyzed 1,944 research articles published between December 2006 and December 2012 in 28 HOA economics and interdisciplinary mathematics journals. 311 of the articles were published OA under HOA pilot agreements between two commercial publishers and several research institutions. Consequently, the OA status of the articles was exogenously determined by the agreements, so there was no risk of "author-driven" selection bias. Performing logistic regressions and controlling for article properties, author characteristics, institution properties, OA post- and pre-print properties, two instrumental variables, and HOA pilot agreements, the authors found that the overall HOA effect on citations was insignificant and small in magnitude. However, in cases where pre- or post-prints were not available, they obtained a significant positive effect of HOA. Furthermore, the analysis of a panel of 403 articles with pre-prints available via RePEc revealed a positive HOA citation effect ranging from 6.4% to 8.1%.

To summarize, over the last 15 years, research on the effect of OA on article citations has increasingly incorporated additional control variables and more sophisticated methodology. The examples we have presented reveal remarkable differences regarding the scope and the methods of these studies. Some focus on a single discipline, or even a single journal, with little more than 1000 articles; others cover multiple disciplines and up to 26 million articles. Some samples are restricted to articles published within half of a year; others comprise articles published over a period of up to 49 years. Some rely on all available data in the WoS or similar data sources;

others use a randomized controlled trial or a natural experiment. Most of them deal with quality selection bias, while some also account for early or late view bias. Comparing two different publishing cultures and applying new methodological approaches for robustness, we seek to offer a more differentiated picture on the citation effect of OA publishing.

### 3 | DATA EXPLORATION

#### 3.1 | Data

The data contain article information from all journals included in the WoS "Journal Citation Reports" (JCR) database in the categories Economics & Management (green culture) and Biology (gold culture). We cover the period 2000–2019, including all 83 journals in Biology and 563 journals in Economics & Management listed in the WoS. While the sample is unbalanced in terms of the number of journals in each field, this is to some extent compensated by the number of articles and authors: Biology comprises 267,340 articles written by a total of 1,248,992 authors, whereas Economics & Management accounts for 517,184 articles by 1,281,181 authors. The two fields and associated publication cultures thus differ with respect to the number of authors per publication and the number of articles published in each journal.

Besides the journal name and category, we also collected information on the total number of cites per journal and year, the JIF, and various other indicators of journal quality (e.g., immediacy score and eigenfactor score) from the JCR database. The JCR data were matched with article-level data and information on journal OA availability using the ISSN identifiers. Additionally, data on the OA status of the journals were acquired via the Directory of Open Access Journals (DOAJ) API. The two sources were merged using either the ISSN or journal titles.

The article and author information was gathered using Microsoft's Project Academic Knowledge API accessing Microsoft Academic Graph (MAG) data. We thus retrieved information on all the articles published in the 646 journals over the sample period. This includes the number of citations per article, the number of references and authors, and the date of publication, as well as information about the authors of each article (institutional affiliation, number of publications and citations of the institution, number of publications and citations of the author). The resulting data set comprises 2,530,173 author–article observations. Since the unit of analysis is not the author but the article, the information on all authors was summarized for each article, resulting in 784,529 observations on individual articles.

The MAG data also contain information about the URL of any online repository that provides access to each article. To construct the variable on potential green OA venues, using the Directory of Open Access Repositories (openDOAR) API, we checked whether any of the unique core URLs per article were listed in the openDOAR register. For an overview of the variables featured in the analysis, see

Table 1. MAG provides yearly data on the number of citations received by every research institution and every author, but not for individual research articles. Hence, the data are essentially of a cross-sectional nature since we only have information about the number of citations at the time of extraction.

### 3.2 | Exploration

By comparing two extreme cases, Biology (gold culture) and Economics & Management (green culture), we are able to investigate whether any citation effect of OA differs across these widely diverging publishing cultures. Looking at the share of OA versus CA journals listed in WoS, that is, only journals that meet a certain quality standard, we find considerable differences between the two fields. While in Biology the share of OA journals has been increasing almost monotonously since 2000, reaching beyond 30% in 2019, the share of OA journals in Economics & Management has remained very low with only slight increases since 2010, reaching about 5% in 2019 (Figure 1). Interestingly, this trend is not driven by journals switching from CA to OA, but by newly launched OA journals. In fact, for the whole period under study, only 1.2% (1.0%) of all OA journals in Biology (Economics) were formerly CA. Moreover, 94.7% (98.7%) of the journals in Biology (Economics & Management) as from 2000 stay in our sample throughout the entire sample period.

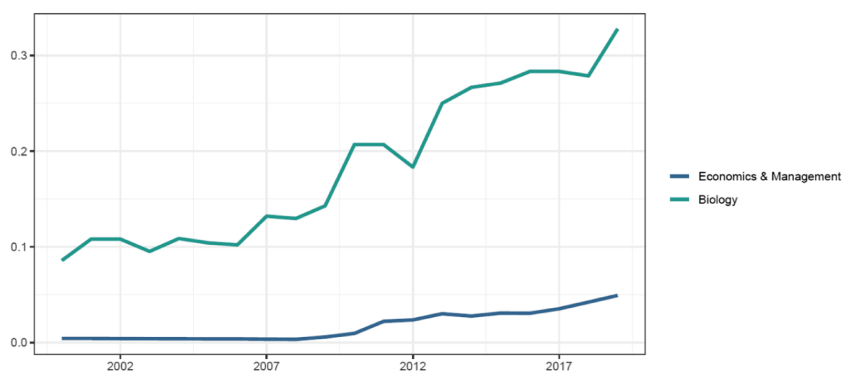
Most interestingly, we find that in Biology, the median impact factor of the OA journals has exceeded that of the CA journals since 2016 (Figure 2).<sup>21</sup> The reputation of OA journals thus equals

that of traditional journals in this discipline. By contrast, in Economics & Management, we find a substantial reputation gap between OA and CA journals. Bear in mind, however, that only very few OA journals are listed in WoS for Economics & Management.<sup>22</sup> The low esteem of OA journals in Economics & Management likely explains their low usage—an important point especially for young scholars seeking tenure.

If we reduce the sample to the top journals in terms of yearly impact factors, this picture of a reputation disadvantage of OA journals in Economics & Management as compared to the situation in Biology becomes even starker (Figure 3). The Top 10 journals in Economics & Management contain not a single OA journal, whereas in Biology, three OA journals made it into the Top 10 by 2019 (*eLife*, *PLoS Biology*, and *BMC Biology*). In Economics & Management, the *Journal of Innovation and Management* has been the only OA journal in the Top 20 journals since 2018. In Biology, the share of OA journals increased steadily to 25 (30) percent among the Top 20 (Top 50) journals by 2019.

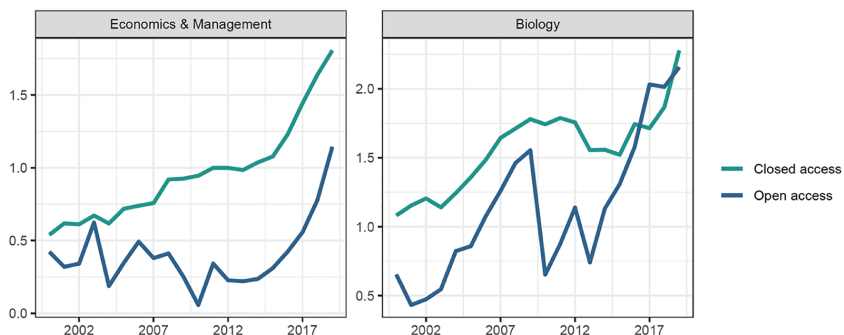
The number of citations, our dependent variable, clearly diverges between CA and OA journals (Figure 4). While the number of citations for OA journals in Economics & Management remained extremely low over the entire period, citations to CA journals increased throughout. By contrast, in Biology, the number of citations to OA journals increased steadily, catching up with the number of citations to CA journals by the end of the period.

Proceeding from the journal level to the article level as the unit of our analysis, Table 1 provides an overview on the summary statistics of all variables used in our regressions.



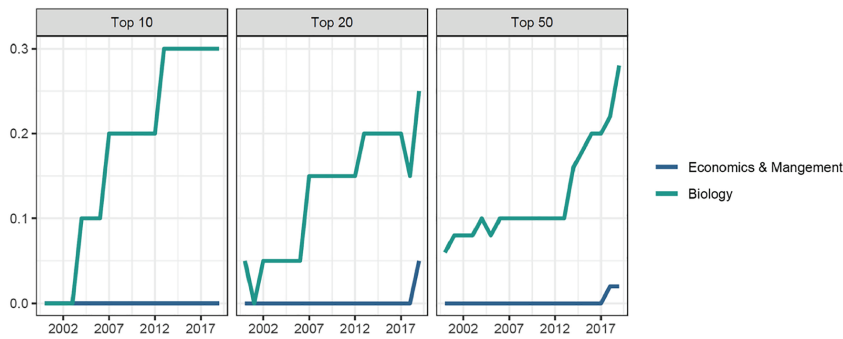
**FIGURE 1** Share of OA journals in Biology and Economics & Management [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Source: Authors' calculations based on WoS (2000-2019)



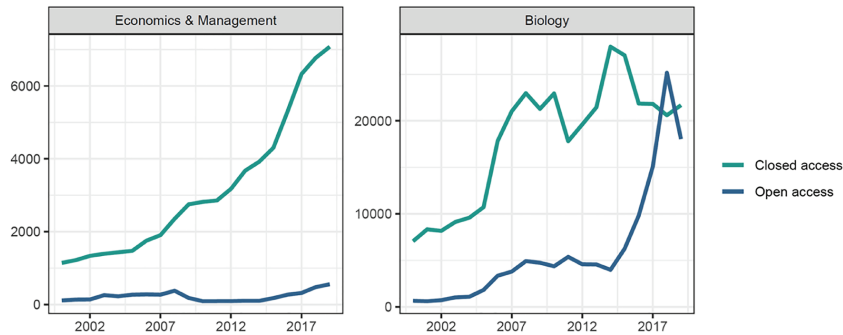
**FIGURE 2** Median impact factor of OA versus CA journals in the two disciplines [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Source: Authors' calculations based on WoS (2000-2019)



Source: Authors' calculations based on WoS (2000-2019)

**FIGURE 3** Share of top OA journals in Biology versus Economics & Management [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



Source: Authors' calculations based on WoS (2000-2019)

**FIGURE 4** Number of citations per OA versus CA journal in Biology and Economics & Management [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

	N	Mean	Std	Min	Median	Max
<b>Dependent variable</b>						
Number of citations	784,529	28.77	106.36	0	6	30,770
<b>Independent variables</b>						
<b>Regime</b>						
Gold OA	784,529	0.08	0.27	0	0	1
Green OA	784,529	0.47	0.50	0	0	1
<b>Discipline (culture)</b>						
Economics & Management	784,529	0.66	0.47	0	1	1
<b>Journal</b>						
JIF Bottom 25	784,529	0.26	0.44	0	0	1
JIF Bottom 50	784,529	0.29	0.45	0	0	1
JIF Top 50	784,529	0.24	0.43	0	0	1
JIF Top 25	784,529	0.21	0.41	0	0	1
<b>Article</b>						
Institution score	784,529	28.74	28.79	0	24.36	2099
Author score	784,529	11.99	47.28	0	3	13,673
Number of authors	784,529	3.23	3.3	1	2	500
Number of references	784,529	28.62	31.87	0	22	500
Publication year	784,529	2011.64	5.31	2000	2012	2019

**TABLE 1** Summary statistics

Source: Authors' calculations based on WoS (2000-2019), MAG, DOAJ, and openDOAR.



### 3.2.1 | Dependent variable

Our dependent variable is the number of citations a paper has received since its publication. The average paper in our sample receives around 29 citations. With a median of six citations and a maximum of 30,770 citations, however, we find considerable variation. As is commonly shown in the literature on the determinants of article citations (see, e.g., Baccini et al., 2014; McCabe & Snyder, 2015), few articles receive more than 100 citations (6.18%) and even fewer more than 1,000 citations (0.15%), while many articles are not cited at all (26.49%).<sup>23</sup> Articles in Economics & Management (32.98) on average receive more citations than Biology articles (20.63).<sup>24</sup> Anecdotal evidence indicates that articles in Biology have a shorter citation half-life, whereas findings in Economics & Management tend to maintain their relevance over time.

### 3.2.2 | Independent variables

The independent variables can be classified into four categories: (1) Regime, gold versus green OA; (2) Discipline (culture), Biology (gold culture) and Economics & Management (green culture); (3) Journal, four JIF groups as proxies of a journal's reputation in the year of publication; (4) Article, additional information on the article and its authors.

### 3.2.3 | Regime

“Gold OA” is a dummy variable that indicates whether an article is published in a pure<sup>25</sup> OA journal. This applies to 8% of the articles in our sample—20% in Biology versus only 1% in Economics & Management.<sup>26</sup> “Green OA” similarly indicates whether an article has been deposited with an OA repository.<sup>27</sup> 47% of all articles were posted in at least one OA repository.

### 3.2.4 | Discipline (culture)

We use a dummy variable for each of the two disciplines, Economics & Management (green culture, 66% of all articles) and Biology (gold culture, 34%).

### 3.2.5 | Journal

The JIF dummies indicate the group of journals that an article belongs to in term of its yearly impact factor (WoS).<sup>28</sup> In each of the two disciplines and for each year of publication, the journals were ranked according to their impact factor. The dummy variable JIF-Q1, for example, refers to the first quartile, or Top 25%, of the journals in either discipline in terms of JIF.

### 3.2.6 | Article

The “institution score” denotes the number of citations per publication for a given institution in a given year. Similarly, the “author score” indicates the sum of citations per publication for a given author in a given year. Other article information refers to the number of authors,<sup>29</sup> the number of references, and the year of publication. On average, a research article is written by 2.48 researchers (maximum of 228) in Economics & Management and 4.67 (maximum of more than 500<sup>30</sup>) in Biology.<sup>31</sup> The total annual publication output increases over time: Half of the articles in the sample were published in the first 13 years (2000–2012), whereas the other half took only 7 years to publish (2012–2019).

## 4 | EMPIRICAL ANALYSIS

### 4.1 | Methodology

As the number of citations constitutes a count variable, we use a Poisson model to estimate the effect of OA on citations. We estimate the equation with the following conditional mean:

$$E(Citations_i | OA_{it}, Discipline_i, X'_{it}) = \exp(\alpha + \beta_1 OA_{it} + \beta_2 Discipline_i + \beta_3 OA_{it} * Discipline_i + \gamma X'_{it}) \quad (1)$$

where  $Citations_i$  denotes the number of citations to article  $i$  at the time of data acquisition,  $OA_{it}$  is the OA dummy for article  $i$  at the time of publication  $t$ ,  $Discipline_i$  means the discipline of the journal in which article  $i$  was published, and  $X_{it}$  encompasses a set of controls (publication regime, JIF, number of references, authors' institution and reputation, and quarter and quarter-discipline FE). Besides quarter FE to control for overall time trends on article citations (see McCabe & Snyder, 2015), in line with Mueller-Langer et al. (2020), we also include quarter-discipline FE to control for discipline-specific citation trends—the assumption being that in Biology, recent articles are cited more often as scientific findings have a relatively short half-life, whereas in Economics & Management, the theoretical foundation was often established by older articles, which are therefore cited more often.

As described above, the distribution of citations is heavily skewed, with many research articles never being cited at all, as is typical of citation data (see, e.g., Baccini et al., 2014; McCabe & Snyder, 2015). This leads to over-dispersion, implying a larger variance compared to the mean and too many zeroes relative to a Poisson distribution. While the negative binomial model would constitute an adequate response to over-dispersion, it too does not provide accurate results when over-dispersion is mainly caused by an excess of zeroes (Baccini et al., 2014). Therefore, in line with McCabe and Snyder (2015), we use a Poisson quasi-maximum likelihood (PQML) model to estimate the conditional mean. It uses the mean regression and variance function from the regular Poisson model but leaves the dispersion parameter unrestricted and to be

**TABLE 2** The OA citation effect

	Dependent variable: Citation_count		
	(1)	(2)	(3)
Gold OA	0.292 <sup>***</sup> (0.270, 0.314)	0.369 <sup>***</sup> (0.346, 0.392)	0.183 <sup>***</sup> (0.160, 0.207)
Green OA	0.833 <sup>***</sup> (0.823, 0.842)	0.836 <sup>***</sup> (0.827, 0.845)	0.822 <sup>***</sup> (0.813, 0.832)
Economics	0.854 <sup>***</sup> (0.843, 0.864)	0.879 <sup>***</sup> (0.868, 0.890)	1.575 <sup>***</sup> (1.500, 1.651)
JIF Bottom 50	0.616 <sup>***</sup> (0.600, 0.631)	0.604 <sup>***</sup> (0.589, 0.620)	0.608 <sup>***</sup> (0.592, 0.623)
JIF Top 50	1.301 <sup>***</sup> (1.286, 1.316)	1.280 <sup>***</sup> (1.264, 1.295)	1.314 <sup>***</sup> (1.298, 1.329)
JIF Top 25	2.453 <sup>***</sup> (2.439, 2.468)	2.410 <sup>***</sup> (2.396, 2.424)	2.443 <sup>***</sup> (2.429, 2.458)
Number of references	0.010 <sup>***</sup> (0.010, 0.010)	0.010 <sup>***</sup> (0.010, 0.010)	0.010 <sup>***</sup> (0.010, 0.010)
Institution score	0.004 <sup>***</sup> (0.004, 0.004)	0.004 <sup>***</sup> (0.004, 0.004)	0.004 <sup>***</sup> (0.004, 0.004)
Author score	0.001 <sup>***</sup> (0.0005, 0.001)	0.001 <sup>***</sup> (0.0005, 0.001)	0.0005 <sup>***</sup> (0.0005, 0.001)
Number of authors	0.011 <sup>***</sup> (0.011, 0.012)	0.011 <sup>***</sup> (0.011, 0.012)	0.011 <sup>***</sup> (0.010, 0.012)
Gold OA* economics		-0.580 <sup>***</sup> (-0.686, -0.473)	-0.492 <sup>***</sup> (-0.597, -0.386)
Quarter FE	Yes	Yes	Yes
Quarter-discipline FE	No	No	Yes
Observations	784,529	784,529	784,529

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.

<sup>\*\*\*</sup>Significance level of 0.1%.

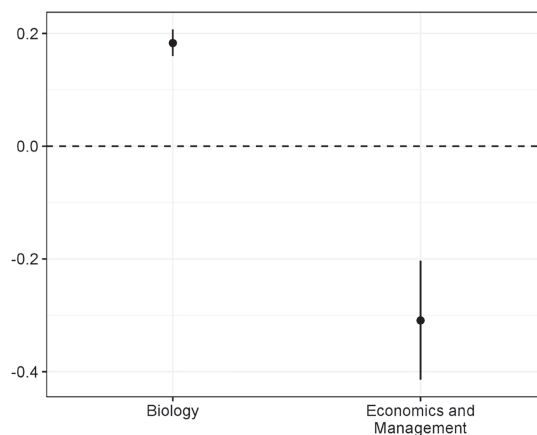
estimated from the data. We also estimate the more flexible hurdle model with Poisson and negative binomial distributions as a robustness check.

## 4.2 | Results

We start with our baseline model before turning to a discipline/culture-specific view on the citation effects of gold and green OA while focusing on the former. First, we estimate the impact of gold OA on the number of citations with three different model specifications, whose results are reported in Table 2. Column (1) refers to the basic model, which includes the regime and discipline (culture) indicators, as well as journal and article information. It also incorporates quarter FE to account for time trends. Specification (2) adds an interaction effect between the discipline and gold OA. Finally, we add quarter-discipline FE in Specification (3) to control for any discipline-specific citation trends over time.

All regression coefficients in Table 2 refer to marginal effects, which can be interpreted as proportionate changes. For instance, a marginal effect of +0.292 signifies that gold OA raises the number of citations by 29.2%. The numbers in parentheses refer to the 95% confidence intervals of the marginal effects.

We find a positive and robust citation effect across all specifications.<sup>32</sup> Moving from Model (1) to Model (2), the additional inclusion of the interaction term between gold OA and Economics raises the gold OA citation effect from 0.292 to 0.369. However, the interpretation of the gold OA coefficient consequently changes: It now indicates the effect of gold OA when the discipline variable (Economics) is zero. As such, the gold OA coefficient shows the increase in citations for



Note: The figure shows the estimated effects of open access publication on the number of citations by disciplines according to the interaction effect specified in specification (3) of table 4.1 (see also equation (1)).

**FIGURE 5** Marginal effects of open access publishing by disciplines

the discipline of Biology. Interestingly, however, the coefficient of the interaction term itself is negative, which means that gold OA does less to promote citations in Economics & Management than it does in Biology. In fact, gold OA even reduces citations in Economics & Management.

Our preferred specification is Model (3), which accounts for any discipline-specific citation trends by controlling for quarter-discipline FE. We find that gold OA increases the number of citations by 18.3% in Biology. The citation effect in Economics & Management is obtained by adding the gold OA effect to the interaction term: Publishing in gold OA reduces citations in Economics & Management by 30.9%.<sup>33</sup> Figure 5 highlights these findings of a reversed citation

effect when comparing Biology to Economics & Management. While gold OA boosts citations in Biology, it reduces citations in Economics & Management.

Hence, again, the research discipline matters to the magnitude of the gold OA effect on citations. This emphasizes the importance of controlling not only for quarter but also for quarter-discipline FE to account for any discipline-specific trends over time. The sharp decrease of the citation effect when including quarter-discipline FE suggests that citations increase over time in both disciplines. Accordingly, the impact of OA on the number of citations declines from 36.9% to 18.3% for Biology and from  $-21.1\%$  to  $-30.9\%$  in Economics & Management.

As expected, green OA has a positive impact on the number of citations (82.2%). We also find that the greater a journal's reputation (measured by its impact factor), the greater the effect of OA on the number citations. The control variables on the article level (number of authors and references, author reputation, and institution) also increase the number of citations by small amounts.<sup>34</sup>

### 4.3 | Robustness checks

#### 4.3.1 | Hurdle model

To check the robustness of the quasi-Poisson models of our main analysis, we also estimate the more flexible hurdle models introduced by Mullahy (1986). Hurdle models are two-component models that comprise, first, a hurdle component that models zero against larger counts and, second, a truncated count component (e.g., Poisson or negative binomial) for positive counts. For the hurdle (zero) component, we use a binomial logistic regression for its intuitive interpretation. We alternatively use the Poisson and the negative binomial distribution to estimate the truncated component (see Table A4). This estimation strategy, too, yields a significant citation effect of gold OA on the number of citations, with effects running in different directions depending on the field of study. The coefficient of gold OA in the zero component of the hurdle model shows a 25.3% higher chance of receiving the first citation if the article was published in a gold OA journal in Biology. The interaction term between gold OA and Economics is also both statistically and substantially negative. In line with our main analysis, both truncated component models also report a citation advantage of gold OA in Biology of +7.9% (negative binomial) and +11.8% (Poisson) and a large negative impact of gold OA on citations in Economics & Management. Given its much higher log likelihood score, we prefer the negative binomial to the Poisson model.

#### 4.3.2 | Alternative variables

In our main analysis, following Brogaard et al. (2020), we controlled for author reputation, as proxied by an author's total number of citations to their published papers. We then averaged these sums across

all the authors of a given article to obtain its author score. However, as Brogaard et al. (2020) argue, the citation count may be more closely related to the score of the first author or to the maximum score of the authors of a particular article. We therefore recreate Model (3) of Table 2 for all three author score measures (see Table A5), finding essentially the same outcomes. For example, raising the score of the mean (first, maximum) author by one point increases citations by 0.05% (0.03%, 0.03%). The effects of our main variable of interest (gold OA) remain similar.

Another important finding of our main analysis is the strong positive impact of a journal's reputation on citations. Clearly, journal reputation can also be proxied by measures other than the impact factor. As robustness checks of the gold OA citations effect obtained in the main analysis, we alternatively use the 5-year impact factor, the eigenfactor, and the immediacy index.<sup>35</sup> The results of all four measures are reported in Table A6. Again, the citation effect remains largely similar, regardless of which measure of reputation we employ. The marginal effect ranges from +10.5% (5-year impact factor) to +41.5% (eigenfactor).

Thirdly, in our main analysis, we controlled for time trends by using quarter FE and quarter-discipline FE (see Table 2). As another robustness check, we instead use the age of each journal article as a continuous variable (measured in days between publication and our acquisition of the data). The results are again quite similar (see Table A7).

Fourth, regarding the number of authors contributing to an article, following McCabe and Snyder (2015), we assumed a linear effect in our main analysis. This implies that every additional author increases the number of citations to a similar degree, as they introduce a new network of interested parties to the article. However, since the authors' networks arguably overlap, the marginal contribution of additional authors may be expected to decline. We reestimate all three models shown in Table 2 with the natural logarithm of the number of authors to control for this potential effect. Interestingly, the effect of OA for Biology declines but remains robustly positive (see Table A8).

Finally, Microsoft Academic uses an algorithm to determine the number of citations that one may expect to find for each article. We employ this estimated citation count as an alternative dependent variable in lieu of the actual number of citations (see Table A9). Again, the results are very similar to those reported in Table 2.

#### 4.3.3 | Sensitivity analysis

Last, we examine possible outliers in the dependent variable to demonstrate that our results are not driven by articles with a very large number of citations. The results shown in Table A10 mirror those of our main model (Model 3 in Table 2) but exclude articles with more than 1,000, 5,000, and 10,000 citations, respectively. Since the coefficients remain largely the same, our analysis is not overly sensitive to outliers.

## 5 | LIMITATIONS AND CONCLUSIONS

This paper has estimated and compared the citation impact of gold and green OA in two disciplines: Biology and Economics & Management. The empirical analysis covers all articles of the two disciplines that were included in the WoS “JCR” between 2000 and 2019: 267,340 articles written by a total of 1,248,992 authors in 83 Biology journals and 517,184 articles by 1,281,181 authors in 563 Economics & Management journals. Using a PQML model, we show that, controlling for quarter FE, journal ranking, number of references, institutional quality, and the reputation and number of authors, gold OA affects the number of citations in different ways: While citations in Biology are increased by 18.3%, citations in Economics & Management actually decline by 30.9%. This initially surprising result is due to the fact that in Biology, many OA journals are among the top journals in this discipline, whereas in Economics & Management, OA journals still have a poor reputation among the scholars. While in Biology, three of the Top 10 journals grant OA, in Economics & Management, the first OA journal has yet to make its way into the Top 20, and only in 2018 did one OA journal make the Top 50. The late start of OA journals in Economics & Management in conjunction with the emergence of predatory journals has likely prevented OA journals from achieving a good reputation in this discipline, while the early start of some excellent OA journals in Biology facilitated the rise of OA journals with high reputation. While our analysis of gold OA focused on pure OA journals, future analysis should look more closely at the increasing number of HOA journals.

Green OA increases overall citations for both disciplines, by 571.3% in Biology and by 59.9% in Economics & Management. The heterogeneous nature of and data on green OA make it difficult to find a satisfactory explanation for this enormous difference, which is why we focus on gold OA. Future research should devote more attention to the heterogeneity of green OA and its citation impact. Our general results are confirmed by several robustness checks.

Last but not least, some questions remain unanswered. First and foremost, we look at cross-sectional rather than panel data, since the citation counts are not available on a yearly basis. Panel data would permit a much deeper examination of the OA citation trends over time. However, we seek to capture these by controlling for quarter-discipline FE. Second, looking at green OA, we lack information on the specific type of repository used. In this context, it would be interesting to account for early and late view bias or biases that relate to systematic differences between the posted and the published versions of the articles. Third, we only look at pure OA journals but not the increasingly popular HOA journals. A natural follow-up would be to analyze the HOA effect for different publication cultures, as HOA may serve as a substitute for pure OA in some respects.<sup>36</sup> Finally, the decision of researchers to publish articles in OA journals may be driven by motives that we did not control for, possibly introducing endogeneity to our analysis. The perhaps most important omitted motive for publishing OA is the guarantee provided by some universities to pay the APC for accepted OA articles. Authors, who are for whatever reasons not

interested in OA, may opt out. In this context, the German DEAL project—two large deals between a consortium of more than 700 academic institutions on the one hand and Wiley (since 2019) and Springer Nature (since 2020) on the other hand—relies on APC instead of subscription fees for access to the publishers' academic journals. Whenever the corresponding author of an OA article is affiliated with the consortium, her institution automatically pays the APC. Most of the journals concerned are HOA but were formerly CA.<sup>37</sup> Other parts of the world feature similar though smaller agreements between publishers and academic institutions. In all these cases, the individual author neither faces a budget constraint nor is she, even in the case of economic journals, necessarily concerned about the journal's reputation. The HOA journals concerned cover a broad range of JIFs. Future research could analyze a sample of HOA journals published by Wiley and/or Springer Nature and investigate to what extent the decision to publish the article OA depends on whether the corresponding author is affiliated with one of the participating institutions.

### ACKNOWLEDGEMENTS

The authors are grateful for valuable suggestions from Sönke Häselser and an anonymous referee. We also wish to thank seminar participants at the University of Kassel and the LMU Munich as well as participants of the Annual Conference of the German Association of Law and Economics in Berlin for useful comments. Open access funding enabled and organized by Projekt DEAL.

### ENDNOTES

- <sup>1</sup> While citations are important for academic authors, many academic articles also benefit the broader public, who cannot easily access academic libraries. OA may thus increase the dissemination of content beyond what we see in the citation counts.
- <sup>2</sup> See also Anderson (2018, pp. 238–247) and Ritchie (2020, p. 184f).
- <sup>3</sup> Eger and Scheufen (2018) distinguish three so-called publishing cultures that determine the attitude of researchers toward the gold versus green road of OA: (1) the gold culture, with predominant usage of the gold road of OA publishing (i.e., OA journals), (2) the green culture, with predominant usage of the green road of OA publishing (i.e., online repositories), and (3) the gray culture, with little usage of either road.
- <sup>4</sup> See the excellent survey by Craig et al. (2007). More recent surveys are provided by Swan (2010), Davis and Walters (2011), and Lewis (2018).
- <sup>5</sup> See also the summary in Craig et al. (2007, p. 9).
- <sup>6</sup> There are several more recent studies along this line that focus on differences between disciplines and types of OA. Archambault et al. (2017) examine more than 3 million papers indexed in the Web of Science (WoS) from 22 disciplines published between 2007 and 2009 and the related citations (almost 35 million) in more than 12,000 WoS-indexed journals from 2007 to mid-2016. They identified gold and green OA papers that were published in peer-reviewed journals, finding that across all disciplines, OA papers of either type received 50% more citations than strictly paywalled papers. Green OA is more effective than gold OA in 20 of the 22 fields. Gold OA is more effective only in Biology and Biomedical Research; in Clinical Medicine, the difference is very small. According to the authors, these results are due to the NIH and Wellcome Trust OA mandates. Piwowar et al. (2018) analyzed a sample of 100,000 WoS-indexed articles and reviews from several disciplines published between 2009 and 2015. They found an OA citation

advantage of 18%, with differences regarding the way in which the papers are made OA. While the citation advantage amounted to 33% for self-archiving (green OA) and 31% for hybrid OA, gold OA yielded an average *negative* citation impact of 17% below the global average and 9% below that of strictly paywalled papers. The authors suggest that this negative effect of gold OA may be due to the increase in the number of newer and smaller OA journals that are considered less prestigious and are not always published in English and the continued growth of so-called mega journals such as PLoS ONE.

<sup>7</sup> See [https://www.stm-assoc.org/2018\\_10\\_04\\_STM\\_Report\\_2018.pdf](https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf). For 2015, Ulrich's journal database Ulrichsweb even listed 111,770 active scholarly journals: 66,734 peer-reviewed, 47,826 available online, 15,025 open access, and 10,916 included in Thomson Reuters Journal Citation Reports. See Moed (2017, p. 194).

<sup>8</sup> The main driver of the OA movement was the so-called serials crisis, that is, the drastic increase in the subscription fees for academic journals, which raised the financial pressure on the libraries (see, e.g., Eger & Scheufen, 2018). This development induced many libraries to forego some journals and, in particular, many books. The OA citation advantage could therefore also apply especially to citations to book chapters that are posted in OA repositories.

<sup>9</sup> See the contribution by McCabe and Snyder to this special issue.

<sup>10</sup> Haustein & Larivière (2015, p. 6) state that "as a rule of thumb, 80% of citations are received by 20% of documents and many are never cited, especially in the humanities."

<sup>11</sup> <https://v2.sherpa.ac.uk/romeo>

<sup>12</sup> See in particular Kurtz et al. (2005).

<sup>13</sup> See Eger & Scheufen, 2018 and the above remarks on predatory OA journals and "Beall's List."

<sup>14</sup> <https://science.sciencemag.org/content/357/6358/1344>; [https://polecopub.hypotheses.org/2017#identifer\\_1\\_2017](https://polecopub.hypotheses.org/2017#identifer_1_2017)

<sup>15</sup> <https://v2.sherpa.ac.uk/romeo>

<sup>16</sup> The CA articles are made open after 6 months.

<sup>17</sup> As a crude proxy for article quality, the authors used the evaluation given by biology professors on the website F1000 Biology (81% not evaluated, 12% recommended, 6% must read, and 1% exceptional).

<sup>18</sup> Many participating journals made all articles freely available after some embargo period.

<sup>19</sup> In another similar study, but with a focus on the impact of *online availability* on citations, McCabe and Snyder (2015) analyzed 260,000 articles from 100 journals in Economics and Business published between 1956 and 2005 and cited between 1980 and 2005 in ISI-indexed journals. Unlike in the earlier study on science journals, here citations to the average journal volume peak only in the fifth year after publication, while the secular upward trend in the frequency of citations is confirmed. Controlling for quality effects by journal and volume fixed effects, as well as for age and time effects, the authors no longer found a significant effect of online availability on citation counts. However, they did find substantial heterogeneity across platforms; for instance, a doubling of JSTOR subscriptions raised citations by 10% causes on average. Finally, they observed a larger impact of JSTOR availability on citations from developing countries and no effect on the citations from authors in non-English-speaking Europe. The decline of the JSTOR effect for more recent content may be due to the increasing importance of the self-archiving of pre- and post-prints.

<sup>20</sup> This is calculated in three ways: "each opened article is compared to the mean and the median article in the same journal which it appeared in and also to equivalent articles in that issue" (p. 3), where "equivalence" refers to citation counts.

<sup>21</sup> Note that comparing journal impact factors is only valid within but not between disciplines, due to different publishing cultures.

<sup>22</sup> The number of OA (CA) journals in Economics & Management increased steadily from 1 (232) in 2000 to 26 (528) in 2019.

<sup>23</sup> The distribution of the number of citations is shown in Figure A1.

<sup>24</sup> See Tables A1 and A2 for discipline-specific summary statistics.

<sup>25</sup> We follow the definition of the DOAJ to identify pure OA journals. In contrast to an HOA journal, a pure OA journal offers free online access to all journal content, rather than just to a selection of papers.

<sup>26</sup> See Tables A1 and A2 for discipline-specific summary statistics.

<sup>27</sup> We identify OA repositories in accordance with the definition of openDOAR.

<sup>28</sup> The JIF-dummies must be calculated anew in each year of publication using the impact factors at the time of publication to avoid circular arguments. Only thus can we ensure that our category formation does not interfere with the number of citations in subsequent periods, which is the basis of the independent variable.

<sup>29</sup> Note that the sum of the "number of authors" across all papers exceeds the number of authors in the sample as most authors appear on several articles.

<sup>30</sup> Microsoft Academic caps the number of authors and references at 500. Since that cap concerns two (66) articles with more than 500 authors (more than 500 references) from a total of 784,529 articles, it should not impede the analysis.

<sup>31</sup> See Tables A1 and A2 for discipline-specific summary statistics.

<sup>32</sup> Given the very large number of observations, tests for statistical significance have little merit since all coefficients become significant. Hence, we report confidence intervals for the coefficients and focus on the size of the effects, rather than on their statistical significance.

<sup>33</sup> Table A3 reports separate effects for the two disciplines.

<sup>34</sup> Table A3 also reports separate models for the disciplines of Biology and Economics & Management. While most coefficients are comparable in magnitude and direction, important differences can be found in the effect of gold OA (as discussed above) but also in the effect of green OA, which is considerably larger for Biology compared to Economics & Management.

<sup>35</sup> The eigenfactor estimates the importance of a journal by weighing citations from highly ranked journals higher than those of poorly ranked journal contributions. The immediacy index indicates how quickly a journal's articles are cited.

<sup>36</sup> Hybrid OA can be regarded as a form of price discrimination as the publishers distinguish between the markets for authors (author fee) and readers (subscription fee). Critics have been pointing to double-dipping as a major disadvantage of hybrid OA (e.g., Pinfield et al., 2016). Thus, it may be argued as to whether hybrid OA may serve as a substitute to pure OA from a social welfare perspective.

<sup>37</sup> For more details see the contribution by Haucap et al. to this special issue.

## DATA AVAILABILITY STATEMENT

Encourages Data Sharing

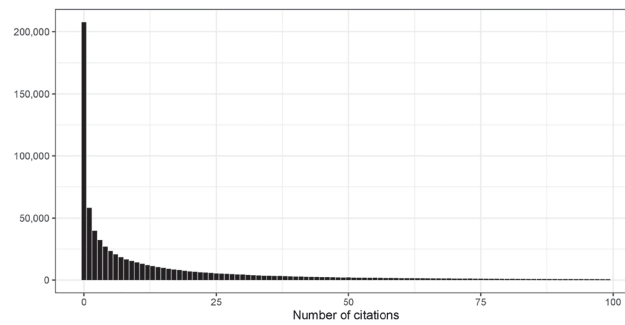
## REFERENCES

Anderson, R. (2015). Should we retire the term "predatory publishing"? The Scholarly Kitchen, May 11. <https://scholarlykitchen.sspnet.org/2015/05/11/should-we-retire-the-term-predatory-publishing>. Accessed 28 February 2021.

- Anderson, R. (2018). *Scholarly communication. What everyone needs to know*. Oxford University Press.
- Antelmann, K. (2004). Do open-access articles have a greater research impact? *College & Research Libraries*, 65(5), 372–382. <https://doi.org/10.5860/crl.65.5.372>
- Archambault, E., Côté, G., Struck, B., & Voorons, M. (2017). Research impact of paywalled versus open access papers. <https://www.1science.com/1numbr>
- Baccini, A., Barabesi, L., Cioni, M., & Pisani, C. (2014). Crossing the hurdle: The determinants of individual scientific performance. *Scientometrics*, 101, 2035–2062. <https://doi.org/10.1007/s11192-014-1395-3>
- Brogaard, J., Engelberg, J., Eswar, S., & Van Wesep, E. (2020). On the causal effect of fame on citations. Working Paper. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3565487](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3565487)
- Craig, I. D., Plume, A., Mcveigh, M. E., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact? A critical review of the literature. *Journal of Informetrics*, 1(3), 239–248. <https://doi.org/10.1016/j.joi.2007.04.001>
- Davis, P. M. (2011). Open access, readership, citations: A randomized controlled trial of scientific journal publishing. *FASEB Journal*, 25(7), 2129–2134. <https://doi.org/10.1096/fj.11-183988>
- Davis, P. M., & Walters, W. H. (2011). The impact of free access to the scientific literature: A review of recent research. *Journal of the Medical Library Association*, 99(3), 208–217. <https://doi.org/10.3163/1536-5050.99.3.008>
- Eckdahl, T. (2004). Review of: PLoS biology – A freely available, open access online journal. *Cell Biology Education*, 3(1), 15–17. <https://doi.org/10.1187/cbe.04-01-0026>
- Eger, T., & Scheufen, M. (2018). *The economics of open access. On the future of academic publishing*. Edward Elgar.
- Evans, J. A., & Reimer, J. (2009). Open access and global participation in science. *Science*, 323(5917), 1025.
- Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, 4(5), 692–698.
- Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-selected or mandated, open access increases citation impact for higher quality research. *PLoS ONE*, 5(10), 1–12.
- Gaulé, P., & Maystre, M. (2011). Getting cited: Does open access help? *Research Policy*, 40(10):1332–1338. <https://doi.org/10.1016/j.respol.2011.05.025>
- Hajjem, C., Harnad, S., & Gingras, Y. (2007). Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. [arXiv:cs/0701137](https://arxiv.org/abs/cs/0701137).
- Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10(6).
- Haustein, S., & Larivière, V. (2015). The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects. In I. M. Welpel, J. Wollersheim, S. Ringelhan, & M. Osterloh (Eds.), *Incentives and performance* (pp. 121–139). Governance of Research Organizations: Springer.
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & Murray, S. S. (2005). The effect of use and access on citations. *Information Processing & Management*, 41, 1395–1402. <https://doi.org/10.1016/j.ipm.2005.03.010>
- Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411(2001), 521. <https://www.nature.com/articles/35079151>
- Lewis, C. (2018). The open access citation advantage: Does it exist and what does it mean for libraries? *Information Technology and Libraries*, 37(3), 50–65. <https://doi.org/10.6017/ital.v37i3.10604>
- McCabe, M. J., & Snyder, C. M. (2014). Identifying the effect of open access on citations using a panel of science journals. *Economic Inquiry*, 52(4), 1284–1300. <https://doi.org/10.1111/ecin.12064>
- McCabe, M. J., & Snyder, C. M. (2015). Does online availability increase citations? Theory and evidence from a panel of economics and business journals. *Review of Economics and Statistics*, 97(1), 144–165. [https://doi.org/10.1162/REST\\_a\\_00437](https://doi.org/10.1162/REST_a_00437)
- Moed, H. F. (2017). *Applied evaluative informetrics*. Cham: Springer.
- Mueller-Langer, F., Scheufen, M., & Waelbroeck, P. (2020). Does online access promote research in developing countries? Empirical evidence from article-level data. *Research Policy*, 49(2), 1–20. <https://doi.org/10.1016/j.respol.2019.103886>
- Mueller-Langer, F., & Watt, R. (2018). How many more cites is a \$3,000 fee buying you? Evidence from a natural experiment. *Economic Inquiry*, 56(2), 931–954. <https://doi.org/10.1111/ecin.12545>
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33, 341–365. [https://doi.org/10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3)
- Ottaviani, J. (2016). The post-embargo open access citation advantage: It exists (probably), it's modest (usually), and the rich get richer (of course). *PLoS ONE*, 11(8), 1–11. <https://doi.org/10.1371/journal.pone.0159614>
- Pinfield, S. (2015). Making open access work. The 'state-of-the-art' in providing open access to scholarly literature. *Online Information Review*, 39(5), 604–636. <https://doi.org/10.1108/OIR-05-2015-0167> (accessed 28 February 2021)
- Pinfield, S., Salter, J., & Bath, P. (2016). The “total cost of publication” in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions. *Journal of the Association for Information Science and Technology*, 67(7), 1751–1766. <https://doi.org/10.1002/asi.23446>
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of open access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>
- Ritchie, S. (2020). *Science fictions: Exposing fraud, bias, negligence and hype in science*. London: The Bodley Head.
- Suber, P. (2006). Open access in the USA. In N. Jacobs (Ed.), *Open access: Key strategic, technical and economic aspects* (pp. 149–160). Oxford: Chandos.
- Swan, A. (2010). The open access citation advantage. Studies and results to date. University of Southampton Institutional Repository. <https://eprints.soton.ac.uk/268516>

**How to cite this article:** Eger, T., Mertens, A., & Scheufen, M. (2021). Publication cultures and the citation impact of open access. *Managerial and Decision Economics*, 42(8), 1980–1998. <https://doi.org/10.1002/mde.3429>

APPENDIX A.



Note: For the purpose of the figure, the distribution is cut off at 100 citations because the outliers, a few of which have received in excess of 30,000 citations, would have been difficult to depict.

Source: Authors' calculations based on WoS (2000-2019), MAG, DOAJ and openDOAR.

FIGURE A1 Distribution of the Number of Citations

TABLE A1 Summary statistics for "Economics & Management"

	N	Mean	Std	Min	Median	Max
Dependent variable						
Number of citations	517,189	32.98	123.07	0	7	30,770
Independent variables						
Regime						
Gold OA	517,189	0.01	0.12	0	0	1
Green OA	517,189	0.4	0.49	0	0	1
Journal						
JIF Bottom 25	517,189	0.25	0.43	0	0	1
JIF Bottom 50	517,189	0.25	0.43	0	0	1
JIF Top 50	517,189	0.26	0.44	0	0	1
JIF Top 25	517,189	0.24	0.43	0	0	1
Article						
Institution rank	517,189	28.8	27.6	0	25.81	2099
Author score	517,189	11.05	46.6	0	2	7856.67
Number of authors	517,189	2.48	1.68	1	2	228
Number of references	517,189	28.91	30.06	0	23	500

Source: Authors calculations based on WoS (2000-2019), MAG, DOAJ and openDOAR.

**TABLE A2** Summary statistics for "Biology"

	N	Mean	Std	Min	Median	Max
Dependent variable						
Number of citations	267,340	20.63	61.61	0	4	5912
Independent variables						
Regime						
Gold OA	267,340	0.2	0.4	0	0	1
Green OA	267,340	0.61	0.49	0	1	1
Journal						
JIF Bottom 25	267,340	0.26	0.44	0	0	1
JIF Bottom 50	267,340	0.36	0.48	0	0	1
JIF Top 50	267,340	0.21	0.41	0	0	1
JIF Top 25	267,340	0.16	0.37	0	0	1
Article						
Institution score	267,340	28.63	30.98	0	20.41	1503.48
Author score	267,340	13.82	48.52	0	5.17	13,673
Number of authors	267,340	4.67	4.83	1	4	500
Number of references	267,340	28.07	35.09	0	20	500

Source: Authors calculations based on WoS (2000–2019), MAG, DOAJ, and openDOAR.

**TABLE A3** Citation effect by disciplines

	Dependent variable:	
	Number of citations	
	(1) Biology	(2) Economics
Gold OA	0.133 <sup>***</sup> (0.111, 0.155)	-0.364 <sup>***</sup> (-0.475, -0.253)
Green OA	5.713 <sup>***</sup> (5.677, 5.749)	0.599 <sup>***</sup> (0.588, 0.609)
JIF Bottom 50	0.726 <sup>***</sup> (0.700, 0.751)	0.581 <sup>***</sup> (0.560, 0.601)
JIF Top 50	1.524 <sup>***</sup> (1.495, 1.553)	1.347 <sup>***</sup> (1.328, 1.366)
JIF Top 25	2.334 <sup>***</sup> (2.309, 2.359)	2.456 <sup>***</sup> (2.438, 2.475)
Number of references	0.008 <sup>***</sup> (0.008, 0.008)	0.011 <sup>***</sup> (0.011, 0.011)
Institution score	0.003 <sup>***</sup> (0.003, 0.003)	0.004 <sup>***</sup> (0.004, 0.004)
Author score	0.0004 <sup>***</sup> (0.0004, 0.0004)	0.001 <sup>***</sup> (0.001, 0.001)
Number of authors	0.010 <sup>***</sup> (0.010, 0.011)	0.022 <sup>***</sup> (0.020, 0.023)
Quarter FE	Yes	Yes
Observations	267,340	517,189

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.

<sup>\*\*\*</sup>Significance level of 0.1%.



**TABLE A4** Hurdle model using negative binomial and Poisson regression

	Zero component (first stage)		Count component (second stage)	
	Negative binomial	Poisson	Negative binomial	Poisson
Gold OA	0.253 <sup>***</sup> (0.219, 0.287)	0.253 <sup>***</sup> (0.219, 0.287)	0.079 <sup>***</sup> (0.064, 0.094)	0.118 <sup>***</sup> (0.116, 0.120)
Green OA	2.071 <sup>***</sup> (2.054, 2.088)	2.071 <sup>***</sup> (2.054, 2.088)	0.356 <sup>***</sup> (0.349, 0.363)	0.343 <sup>***</sup> (0.342, 0.344)
Economics	3.384 <sup>***</sup> (3.288, 3.481)	3.384 <sup>***</sup> (3.288, 3.481)	0.887 <sup>***</sup> (0.831, 0.943)	0.463 <sup>***</sup> (0.455, 0.470)
JIF Bottom 50	-0.218 <sup>***</sup> (-0.238, -0.198)	-0.218 <sup>***</sup> (-0.238, -0.198)	0.553 <sup>***</sup> (0.544, 0.561)	0.659 <sup>***</sup> (0.657, 0.660)
JIF Top 50	-0.264 <sup>***</sup> (-0.285, -0.243)	-0.264 <sup>***</sup> (-0.285, -0.243)	1.208 <sup>***</sup> (1.199, 1.218)	1.439 <sup>***</sup> (1.437, 1.440)
JIF Top 25	-0.421 <sup>***</sup> (-0.442, -0.400)	-0.421 <sup>***</sup> (-0.442, -0.400)	1.998 <sup>***</sup> (1.988, 2.008)	2.817 <sup>***</sup> (2.816, 2.819)
Number of references	0.090 <sup>***</sup> (0.089, 0.090)	0.090 <sup>***</sup> (0.089, 0.090)	0.015 <sup>***</sup> (0.014, 0.015)	0.009 <sup>***</sup> (0.009, 0.009)
Institution score	0.012 <sup>***</sup> (0.011, 0.012)	0.012 <sup>***</sup> (0.011, 0.012)	0.005 <sup>***</sup> (0.005, 0.005)	0.004 <sup>***</sup> (0.004, 0.004)
Number of authors	0.026 <sup>***</sup> (0.023, 0.029)	0.026 <sup>***</sup> (0.023, 0.029)	0.045 <sup>***</sup> (0.043, 0.046)	0.011 <sup>***</sup> (0.011, 0.011)
Author score	0.006 <sup>***</sup> (0.006, 0.007)	0.006 <sup>***</sup> (0.006, 0.007)	0.005 <sup>***</sup> (0.005, 0.005)	0.005 <sup>***</sup> (0.005, 0.005)
Gold OA* <i>economics</i>	-0.383 <sup>***</sup> (-0.452, -0.315)	-0.383 <sup>***</sup> (-0.452, -0.315)	-0.232 <sup>***</sup> (-0.273, -0.192)	-0.368 <sup>***</sup> (-0.378, -0.357)
Quarter FE	Yes	Yes	Yes	Yes
Quarter-discipline FE	Yes	Yes	Yes	Yes
Observations	784,529	784,529	784,529	784,529
Log likelihood	-2,580,085	-15,431,682	-2,580,085	-15,431,682

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.  
<sup>\*\*\*</sup>Significance level of 0.1%.

**TABLE A5** Baseline model with different specifications of author reputation

	Dependent variable		
	Number of citations		
	(1)	(2)	(3)
Gold OA	0.183 <sup>***</sup> (0.160, 0.207)	0.181 <sup>***</sup> (0.157, 0.204)	0.169 <sup>***</sup> (0.145, 0.192)
Green OA	0.822 <sup>***</sup> (0.813, 0.832)	0.827 <sup>***</sup> (0.818, 0.837)	0.825 <sup>***</sup> (0.815, 0.834)
Economics	1.575 <sup>***</sup> (1.500, 1.651)	1.573 <sup>***</sup> (1.497, 1.648)	1.574 <sup>***</sup> (1.498, 1.650)
JIF Bottom 50	0.608 <sup>***</sup> (0.592, 0.623)	0.604 <sup>***</sup> (0.589, 0.620)	0.604 <sup>***</sup> (0.588, 0.619)
JIF Top 50	1.314 <sup>***</sup> (1.298, 1.329)	1.313 <sup>***</sup> (1.298, 1.328)	1.305 <sup>***</sup> (1.290, 1.320)
JIF Top 25	2.443 <sup>***</sup> (2.429, 2.458)	2.450 <sup>***</sup> (2.436, 2.465)	2.428 <sup>***</sup> (2.414, 2.442)
Number of references	0.010 <sup>***</sup> (0.010, 0.010)	0.010 <sup>***</sup> (0.010, 0.010)	0.010 <sup>***</sup> (0.010, 0.010)
Institution score	0.004 <sup>***</sup> (0.004, 0.004)	0.004 <sup>***</sup> (0.004, 0.004)	0.004 <sup>***</sup> (0.004, 0.004)
Author score (mean)	0.0005 <sup>***</sup> (0.0005, 0.001)		
Author score (first)	0.0003 <sup>***</sup> (0.0003, 0.0003)		
Author score (max)	0.0003 <sup>***</sup> (0.0003, 0.0003)		
Number of authors	0.011 <sup>***</sup> (0.011, 0.012)	0.011 <sup>***</sup> (0.011, 0.012)	0.010 <sup>***</sup> (0.009, 0.010)
Gold OA* <i>economics</i>	-0.492 <sup>***</sup> (-0.597, -0.386)	-0.491 <sup>***</sup> (-0.597, -0.385)	-0.487 <sup>***</sup> (-0.593, -0.381)
Quarter-year FE	Yes	Yes	Yes
Quarter-year-discipline FE	Yes	Yes	Yes
Observations	784,529	784,529	784,529

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.  
<sup>\*\*\*</sup>Significance level of 0.1%.

**TABLE A6** Citation effect using alternative measures of reputation

	Dependent variable			
	Number of citations			
	(1) JIF	(2) 5-year IF	(3) Eigenfactor	(4) Immediacy
Gold OA	0.144*** (0.122, 0.167)	0.105*** (0.082, 0.127)	0.415*** (0.393, 0.438)	0.263*** (0.240, 0.287)
Green OA	0.812*** (0.803, 0.822)	0.999*** (0.989, 1.010)	0.809*** (0.799, 0.820)	0.833*** (0.823, 0.843)
Economics	1.473*** (1.398, 1.547)	1.534*** (1.473, 1.595)	1.454*** (1.392, 1.516)	1.567*** (1.490, 1.645)
Journal impact score	4.076*** (4.060, 4.092)	3.756*** (3.739, 3.773)	2.696*** (2.679, 2.713)	2.40*** (2.388, 2.419)
Number of references	0.010*** (0.010, 0.010)	0.009*** (0.009, 0.009)	0.010*** (0.010, 0.010)	0.010*** (0.010, 0.010)
Institution score	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)
Author score (mean)	0.0005*** (0.0005, 0.001)	0.0004*** (0.0004, 0.0004)	0.0004*** (0.0004, 0.0004)	0.0005*** (0.0005, 0.001)
Number of authors	0.011*** (0.011, 0.011)	0.010*** (0.010, 0.011)	0.010*** (0.010, 0.011)	0.012*** (0.011, 0.012)
Gold OA*economics	-0.440*** (-0.544, -0.336)	-0.354*** (-0.462, -0.246)	-0.595*** (-0.693, -0.496)	-0.605*** (-0.716, -0.493)
Quarter-year FE	Yes	Yes	Yes	Yes
Quarter-year-discipline FE	Yes	Yes	Yes	Yes
Observations	784,529	596,085	632,865	783,089

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.

\*\*\*Significance level of 0.1%.

**TABLE A7** Base model using age

	Dependent variable:	
	Number of citations	
	(1)	(2)
Gold OA	0.287*** (0.264, 0.311)	0.149*** (0.125, 0.173)
Green OA	0.883*** (0.873, 0.892)	0.897*** (0.887, 0.906)
Economics	0.896*** (0.884, 0.907)	0.220*** (0.194, 0.246)
JIF Bottom 50	0.593*** (0.577, 0.609)	0.583*** (0.567, 0.599)
JIF Top 50	1.239*** (1.223, 1.254)	1.235*** (1.220, 1.250)
JIF Top 25	2.284*** (2.269, 2.299)	2.278*** (2.263, 2.292)
Number of references	0.010*** (0.010, 0.010)	0.010*** (0.010, 0.010)
Age	0.0004*** (0.0004, 0.0004)	0.0003*** (0.0003, 0.0003)
Institution score	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)
Author score (mean)	0.001*** (0.001, 0.001)	0.001*** (0.001, 0.001)
Number of authors	0.011*** (0.010, 0.011)	0.010*** (0.010, 0.011)
Gold OA*economics	-0.584*** (-0.693, -0.476)	-0.519*** (-0.628, -0.411)
Age*economics		0.0001*** (0.0001, 0.0001)
Observations	784,529	784,529

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.

\*\*\*Significance level of 0.1%.

**TABLE A8** Main analysis with different specification for the number of authors

	Dependent variable: Number of citations		
	(1)	(2)	(3)
Gold OA	0.242*** (0.220, 0.264)	0.313*** (0.290, 0.335)	0.125*** (0.101, 0.148)
Green OA	0.787*** (0.777, 0.796)	0.790*** (0.781, 0.799)	0.772*** (0.763, 0.781)
Economics	1.037*** (1.026, 1.049)	1.063*** (1.052, 1.075)	1.901*** (1.825, 1.977)
JIF Bottom 50	0.586*** (0.570, 0.602)	0.575*** (0.560, 0.591)	0.579*** (0.564, 0.595)
JIF Top 50	1.219*** (1.204, 1.235)	1.200*** (1.184, 1.215)	1.231*** (1.215, 1.246)
JIF Top 25	2.271*** (2.257, 2.286)	2.233*** (2.218, 2.247)	2.264*** (2.250, 2.279)
Number of references	0.010*** (0.010, 0.010)	0.010*** (0.010, 0.010)	0.010*** (0.010, 0.010)
Institutions score	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)
Author score	0.001*** (0.0005, 0.001)	0.001*** (0.0005, 0.001)	0.0005*** (0.0005, 0.001)
Log of the number of authors	0.298*** (0.291, 0.305)	0.296*** (0.289, 0.303)	0.304*** (0.297, 0.311)
Gold OA* <i>economics</i>		-0.568*** (-0.674, -0.462)	-0.472*** (-0.578, -0.367)
Quarter-year FE	Yes	Yes	Yes
Quarter-year-discipline FE	No	No	Yes
Observations	784,529	784,529	784,529

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.  
 \*\*\*Significance level of 0.1%.

**TABLE A9** Citation effect using an alternative measure for citations

	Dependent variable: Number of citations		
	(1)	(2)	(3)
Gold OA	0.242*** (0.211, 0.274)	0.331*** (0.299, 0.364)	0.150*** (0.117, 0.184)
Green OA	0.848*** (0.836, 0.860)	0.851*** (0.839, 0.863)	0.839*** (0.827, 0.851)
Economics	1.363*** (1.348, 1.378)	1.397*** (1.382, 1.412)	2.125*** (2.022, 2.228)
JIF Bottom 50	0.724*** (0.703, 0.745)	0.711*** (0.690, 0.732)	0.715*** (0.694, 0.736)
JIF Top 50	1.575*** (1.555, 1.595)	1.549*** (1.529, 1.569)	1.585*** (1.565, 1.605)
JIF Top 25	3.066*** (3.047, 3.085)	3.013*** (2.994, 3.032)	3.055*** (3.036, 3.074)
Number of references	0.010*** (0.010, 0.011)	0.010*** (0.010, 0.011)	0.010*** (0.010, 0.011)
Institution score	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)
Author score	0.001*** (0.001, 0.001)	0.001*** (0.001, 0.001)	0.001*** (0.001, 0.001)
Number of authors	0.012*** (0.011, 0.012)	0.011*** (0.011, 0.012)	0.011*** (0.011, 0.012)
Gold OA* <i>economics</i>		-0.617*** (-0.767, -0.467)	-0.542*** (-0.692, -0.392)
Quarter FE	Yes	Yes	Yes
Quarter-discipline-FE	No	No	Yes
Observations	784,529	784,529	784,529

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.  
 \*\*\*Significance level of 0.1%.

TABLE A10 Sensitivity analysis

	Dependent variable: Number of citations		
	(1) <1000 citations	(2) <5000 citations	(3) <10,000 citations
Gold OA	0.175*** (0.157, 0.193)	0.185*** (0.163, 0.207)	0.184*** (0.161, 0.206)
Green OA	0.823*** (0.816, 0.831)	0.819*** (0.810, 0.828)	0.822*** (0.813, 0.831)
Economics	1.486*** (1.429, 1.544)	1.495*** (1.425, 1.566)	1.494*** (1.421, 1.566)
JIF Bottom 50	0.590*** (0.578, 0.602)	0.606*** (0.591, 0.620)	0.604*** (0.589, 0.619)
JIF Top 50	1.246*** (1.235, 1.258)	1.310*** (1.296, 1.324)	1.316*** (1.301, 1.330)
JIF Top 25	2.070*** (2.059, 2.081)	2.412*** (2.399, 2.426)	2.427*** (2.413, 2.441)
Number of references	0.010*** (0.010, 0.010)	0.010*** (0.010, 0.010)	0.010*** (0.010, 0.010)
Institutions score	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)	0.004*** (0.004, 0.004)
Author score	0.0005*** (0.0005, 0.0005)	0.0005*** (0.0005, 0.001)	0.0005*** (0.0005, 0.001)
Number of authors	0.012*** (0.011, 0.012)	0.011*** (0.011, 0.011)	0.011*** (0.011, 0.011)
Gold OA* <i>economics</i>	-0.493*** (-0.572, -0.414)	-0.493*** (-0.590, -0.394)	-0.492*** (-0.593, -0.391)
Quarter-year FE	Yes	Yes	Yes
Quarter-year-discipline FE	Yes	Yes	Yes
Observations	783,390	784,507	784,525

Notes: All coefficients refer to marginal effects. 95% confidence intervals in parentheses. Constant not reported.

\*\*\*Significance level of 0.1%.