Contributions of the Open Access modality to the impact of hybrid journals controlling by field and time effects

Pablo Dorta-González 1,* and María Isabel Dorta-González 2

1 Universidad de Las Palmas de Gran Canaria, TiDES Research Institute, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain. E-mail: pablo.dorta@ulpgc.es
2 Universidad de La Laguna, Departamento de Ingeniería Informática y de Sistemas, Avenida Astrofísico Francisco Sánchez s/n, 38271 La Laguna, Spain. E-mail: isadorta@ull.es

* Corresponding author

Abstract

Purpose: Researchers are more likely to read and cite papers to which they have access than those that they cannot obtain. Thus, the objective of this work is to analyze the contribution of the Open Access (OA) modality to the impact of hybrid journals.

Design / methodology /approach: The “research articles” in the year 2017 from 200 hybrid journals in four subject areas, and the citations received by such articles in the period 2017-2020 in the Scopus database, were analyzed. The hybrid OA papers were compared with the paywalled ones. The journals were randomly selected from those with share of OA papers higher than some minimal value. More than 60 thousand research articles were analyzed in the sample, of which 24% under the OA modality.

Findings: We obtain at journal level that cites per article in both hybrid modalities (OA and paywalled) strongly correlate. However, there is no correlation between the OA prevalence and cites per article. There is OA citation advantage in 80% of hybrid journals. Moreover, the OA citation advantage is consistent across fields and held in time. We obtain an OA citation advantage of 50% in average, and higher than 37% in half of the hybrid journals. Finally, the OA citation advantage is higher in Humanities than in Science and Social Science.

Research limitations: Some of the citation advantage is likely due to more access allows more people to read and hence cite articles they otherwise would not. However, causation is difficult to establish and there are many possible bias. Several factors can affect the observed differences in citation rates. Funder mandates can be one of them. Funders are likely to have OA requirement,
and well-funded studies are more likely to receive more citations than poorly funded studies. Another discussed factor is the selection bias postulate, which suggests that authors choose only their most impactful studies to be open access.

**Practical implications:** For hybrid journals, the open access modality is positive, in the sense that it provides a greater number of potential readers. This in turn translates into a greater number of citations and an improvement in the position of the journal in the rankings by impact factor. For researchers it is also positive because it increases the potential number of readers and citations received.

**Originality /value:** Our study refines previous results by comparing documents more similar to each other. Although it does not examine the cause of the observed citation advantage, we find that it exists in a very large sample.

**Keywords:** open access; open science; scholarly communication; hybrid journals; citation advantage.

**Introduction**

Researchers are more likely to read and cite papers to which they have access than those that they cannot obtain. Thus, since the emergence of the world wide web, scientists and scholarly publishers have used different forms of Open Access (OA), a disruptive model for the dissemination of research publications (Björk, 2004). In the last years, more and more scientists are making their research results openly accessible to increase its visibility, usage, and citation impact (Dorta-González et al., 2017; 2020).

The common characteristic of all different forms of OA is that the primary source of communication of research results, the peer reviewed article, is available to anybody with Internet access free of charge and access barriers (Prosser, 2003).

There are four main OA modalities. **Gold OA** refers to scholarly articles in fully accessible OA journals. **Green OA** refers to publishing in a subscription or pay-per-view journal, in addition to self-archiving the pre- or post-print paper in a repository (Harnad et al., 2004). **Hybrid OA** is an intermediate form of OA, where authors pay scholarly publishers to make articles freely accessible within journals, in which reading the content
otherwise requires a subscription or pay-per-view (Björk, 2017). And finally, *Delayed OA* refers to scholarly articles in subscription journals made available openly on the web directly through the publisher at the expiry of a set embargo period (Laakso & Björk, 2013).

As previously said, a hybrid journal is a traditional one, for which readers need a subscription or where readers can pay to view individual articles. However, the journal offers authors the possibility to open their individual article on condition of the payment of a price similarly than in a gold OA journal. The price level in the hybrid OA is typically around 3,000 USD, which many authors and their institutions perceive as high (Tenopir et al., 2017).

Hybrid journals are a risk free transition path towards full OA, in contrast to starting new full OA journals or converting ones, since the subscription revenue remains (Prosser, 2003). Thus, since Springer announced in 2004 the hybrid option “Open Choice” for their full portfolio of subscription journals, most big publishers have adopted similar modalities and the number of journals offering the hybrid possibility has increased in recent years.

The vast majority of subscription journals from the leading scholarly publishers are nowadays hybrid. The number of journals offering the hybrid option has increased from around 2,000 in the year 2009 to almost 10,000 in the year 2016, and the number of individual articles in the same period has grown from an estimated 8,000 in the year 2009 to 45,000 in the year 2016 (Björk, 2017).

Since Lawrence proposed in 2001 the OA citation advantage, this postulate has been discussed in depth without an agreement being reached (Davis et al., 2008). Furthermore, some authors are critical about the causal link between OA and higher citations, stating that the benefits of OA are uncertain and vary among different fields (Davis & Walters, 2011).

In this paper, as novel contribution, we take a journal-level approach to assessing the OA citation advantage, while many others take a paper-level approach. This is because research articles in both publication modalities in the same hybrid journal and publication year, are quite similar in discipline and with a priori the same citation potential.
Thus, based on citation data from the Scopus database, we provide longitudinal estimations of cites per article in both publication modalities in hybrid journals. Moreover, we answer the following questions:

1. Are hybrid OA research articles more highly cited than their paywalled counterparts?
2. How does this difference vary according to field and time?

**Theoretical framework on open access citation advantage**

Many researchers, starting with Lawrence (2001), have found that OA articles tend to have more citations than pay-for-access articles. This OA citation advantage has been observed in a variety of academic fields including computer science (Lawrence, 2001), philosophy, political science, electrical & electronic engineering, and mathematics (Antelman, 2004), physics (Harnad et al., 2004), biology and chemistry (Eysenbach, 2006), as well as civil engineering (Koler-Povh et al., 2014).

However, this postulate has been discussed in the literature in depth without an agreement being reached (Davis et al., 2008; Dorta-González & Santana-Jiménez, 2018; Norris et al., 2008; Joint, 2009; Gargouri et al., 2010; González-Betancor & Dorta-González, 2019; Wang et al., 2015). Furthermore, some authors are critical about the causal link between OA and higher citations, stating that the benefits of OA are uncertain and vary among different fields (Craig et al., 2007; Davis & Walters, 2011).

Kurtz et al. (2005), and later other authors (Craig et al., 2007; Moed, 2007; Davis et al., 2008), set out three postulates supporting the existence of a correlation between OA and increased citations. (1) OA articles are easier to obtain, and therefore easier to read and cite (*Open Access postulate*). (2) OA articles tend to be available online prior to their publication and therefore begin accumulating citations earlier than pay-for-access articles (*Early View postulate*). And (3) more prominent authors are more likely to provide OA to their articles, and authors are more likely to provide OA to their highest quality articles (*Selection Bias postulate*). Moreover, these authors conclude that early view and selection bias effects are the main factors behind this correlation.
Gaule & Maystre (2011) and Niyazov et al. (2016) found evidence of selection bias in OA, but still estimated a statistically significant citation advantage even after controlling for that bias. Regardless of the validity or generality of their conclusions, these studies establish that any analysis must take into account the effect of time (citation time window) and selection bias.

At journal level, Gumpenberger et al. (2013) showed that the impact factor of gold OA journals was increasing, and that one-third of newly launched OA journals were indexed in the Journal Citation Reports (JCR) after three years. However, Björk and Solomon (2012) argued that the economic model is not related to journal impact. This result has been confirmed by Solomon et al. (2013), concluding that articles are cited at a similar rate regardless of the distribution model.

The OA citation advantage is not universally supported. Many studies have been criticised on methodological grounds (Davis & Walters, 2011), and a research using the randomized-control trial method failed to find evidence of an OA citation advantage (Davis, 2011).

However, recent studies using robust methods have observed an OA citation advantage. McCabe & Snyder (2014) used a complex statistical model to remove author bias and reported a small but meaningful 8% OA citation advantage. Archambault et al. (2014) in a massive sample of over one million articles and using field-normalized citation rates, described a 40% OA citation advantage. Ottaviani (2016) reported a 19% OA citation advantage excluding the author self-selection bias and beyond the first years after publication.

In a recent study, Piwowar et al. (2018) used three samples, each of 100,000 articles, to study OA in three populations: all journal articles assigned a DOI, recent journal articles indexed in Web of Science, and articles viewed by users of the open-source browser extension Unpaywall. They estimated that at least 28% of the scholarly literature is OA, and that this proportion is growing mainly in gold and hybrid journals. Accounting for age and discipline, they observed OA articles receive 18% more citations than average, an effect driven primarily by green and hybrid OA.
**Methodology**

Since the end of 2020, Scopus has new Open Access filters providing information on the type of open access per article. With this new classification system, users can now filter their results or use specific OA tags, i.e. gold, hybrid gold, green, and bronze (delayed).

The source of OA information in Scopus is Unpaywall, an open-source browser extension that lets users find OA articles from publishers and repositories (run by OurResearch, a non-profit organization).

In this study, four subject areas in the Scopus database, one in each branch of knowledge, are considered: Arts & Humanities; Economics, Econometrics & Finance; Medicine; and Physics & Astronomy.

We decided a priori to take four subject areas. This number was set so that both figures and tables could be displayed in the paper. The subject areas were selected based on the previous experience of the authors and trying to cover fields as diverse as possible.

For each of these subject areas, the “research articles” in the year 2017 from 50 hybrid journals, and the citations received by such research articles in the period 2017-2020, were downloaded from the Scopus database (April 8, 2021).

Only 2017 was taken as the year of publication (census) in order to have a citation window of at least three full years for all documents (a full window of three years plus the time elapsed during the year of publication). Note that in most areas the maximum of the distribution of citations is reached before the third year from its publication. Articles published at the beginning of 2017 accumulate their citations for almost four years, while those published at the end of 2017 accumulate their citations for just over three years. This consideration has no consequences on the results obtained since the publication under the hybrid open access modality is distributed uniformly among all the issues of the same year.

The 200 journals were randomly selected from those with share of OA papers in 2017 higher than some minimal value: 5% in Medicine, 4% in Arts & Humanities, 2% in Physics & Astronomy, and 2% in Economics, Econometrics & Finance. Said threshold was set based on the prevalence of the OA modality in each subject area, so that this percentage is higher in areas where the OA modality in hybrid journals is more widespread. This information is detailed in the dataset in Annex A.
A total of 2,020,793 “research articles” were published in the Scopus database in 2017, of which 69,093 were in hybrid journals under the OA modality (3.4%). During that same year, the selected four subject areas published 874,556 research articles, of which 33,796 were in hybrid journals with OA modality (3.9%).

The distribution by subject areas is show in Table 1. The hybrid OA prevalence is 4.6% in Medicine, 3.7% in Arts & Humanities, 2.7% in Physics & Astronomy, and 2.5% in Economics, Econometrics & Finance. The four subject areas represents 43.3% of the database in 2017 by including the largest (Medicine) and the fourth largest (Physics & Astronomy) subject areas. Moreover, the OA articles in hybrid journals in the four subject areas represent 48.9% of the database by including also the largest (Medicine) and the fourth largest (Physics & Astronomy) subject areas in hybrid OA articles.

[Table 1 about here]

In the sample, the 62,608 research articles from 200 hybrid journals were analyzed. Of these, 8043 research articles were published under the OA modality. This represents 23.8% of the total OA research articles published in hybrid journals in the subject areas considered (33,796). This information disaggregate by subject areas is show in Table 2. The areas that are overrepresented in the sample in relation to the OA, in relative terms, are Economics, Econometrics & Finance (49.3%) and Arts & Humanities (40.8%). However, in absolute terms, the total number of OA articles included in these two areas are lower than in Medicine and Physics & Astronomy, due to the larger size of the journals in the latter.

[Table 2 about here]

Results

Cites per article in hybrid journals by modality

About the correlation between variables (Table 3), as expected, the size of the journal does not correlate with any other variable. The OA prevalence in hybrid journals, this is the proportion of research articles under the OA modality, does not correlate with the position of the journal in the citation ranking (best CiteScore percentile). As a particular case, it does weakly and negatively in Arts & Humanities (-0.69), that is, the best-
positioned journals in the citation ranking have a lower proportion of OA articles. This is due to some highly prestigious journals that are still in the initial stages of the hybrid publication model.

[Table 3 about here]

The OA prevalence either does not correlate with cites per article in the hybrid modalities. However, the position of the journal in the citation ranking (percentile) correlates weakly with cites per article in both hybrid modalities.

Note the only two variables that present high correlation, above 0.81 in three subject areas, are cites per article according to modality. That is, the higher cites per article in one modality, the greater in the other. Medicine highlight with a very high correlation (0.97). The exception is Physics & Astronomy, where the correlation reduces to 0.49.

As previously commented, there is a strong and positive linear correlation for cites per article in both hybrid modalities (see Figure 1). The coefficient of determination is generally high, with the exception of Physics & Astronomy. The hybrid journals with the greatest impact in one modality are also in the other. The bisector of the square, that is, the imaginary line that begins in the lower left corner and ends in the upper right corner of the square, separates the citation advantage for each modality. The bubbles below the bisector correspond to hybrid journals with citation advantage for the OA modality. Similarly, the bubbles above the bisector correspond to hybrid journals with citation advantage in the paywalled modality (citation disadvantage for the OA). Note in all the areas there is a majority of journals below the bisector, where the citation advantage corresponds to the OA hybrid modality. In fact, the regression line falls below the bisector in all cases, that is, the OA citation advantage in hybrid journals is observed even in the least squares estimate.

[Figure 1 about here]

In relation to the OA prevalence, this is the proportion of articles in the OA modality of the hybrid journal, indicated through the size of the bubble in Figure 1, there is a tendency for big bubbles to gravitate around the origin. This is especially evident for
Humanities and Physics. This means that hybrid journals with higher proportion of OA papers are usually cited less, which is in accordance with mostly negative correlation coefficients for these indicators in Table 3.

The box diagram for the average of cites per article in hybrid journals, according to modality and year of citation, is show in Figure 2. In all subject areas and each citation year, cites per article for those in the OA modality are clearly higher than the citations in the paywalled modality. These average citations for the OA modality are higher both in mean (indicated with the x symbol) and in quartiles of the distribution (box and whisker). Note that the mean of the distribution is considerably larger than the median. This is because the distribution is asymmetric with a long tail on the right.

![Figure 2 about here](image)

The increase in the number of citations over time relates to the shape of the citation distribution in each subject area. Thus, for example, in Physics & Astronomy the maximum of the distribution reaches in the third year. Beyond this logical increase in the number of citations over years, no clear time effect observes in Figure 2.

Open Access citation advantage in hybrid journals

The OA citation advantage (disadvantage when it is negative) for a journal in a particular year, is defined in relation to the sign of the subtraction as follows. If cites per OA article is greater than cites per paywalled article, then the OA citation advantage is:

\[
\frac{\text{Cites per OA} - \text{Cites per Paywalled}}{\text{Cites per Paywalled}}.
\]

However, if cites per OA article is less than cites per paywalled article, then the OA citation advantage (disadvantage because the result is negative) is:

\[
\frac{\text{Cites per OA} - \text{Cites per Paywalled}}{\text{Cites per OA}}.
\]

The OA citation advantage in relation to the journal percentile shows in Figure 3. There are differences in OA citation advantages between fields. For example, in Medicine there are few journals with a citation disadvantage for the OA, and in most cases the citation advantage is in the range 0–100%. However, in Economics, Econometrics & Finance the differences among journals are much greater and a big number of cases fall
into the range from -100% to 200%. Note the only two highly disadvantaged journals have medium percentiles. A more detailed analysis will follow.

[Figure 3 about here]

Figure 4 shows the OA citation advantage by subject areas, with and without outliers. Note the citation advantage of the OA modality in hybrid journals is clear for all subject areas. The data distribution, represented by the box and whiskers, displaces toward the positive part of the vertical axis. The median of the distribution (the inner line that divides the box into two parts) is in the range 25–50%, while the mean is in 40–60%. There is a citation advantage in more than 75% of the journals. Thus, the 25th percentile (the bottom line of the box) is located close to 0% in the worst case (Economics, Econometrics & Finance). Furthermore, the OA citation advantage is consistent across fields (Figure 4) and held in time (Figure 5).

[Figures 4 and 5 about here]

There is OA citation advantage in 80% of hybrid journals (Table 4). In the remaining 20% there are OA citation disadvantage or there are no differences. The results are relatively stable both across fields and over time. The subject areas where the number of journals with OA citation advantage is higher are Medicine (88%) and Arts & Humanities (82%).

[Table 4 about here]

The average of the OA citation advantage (Table 5) increases with time in the area where the OA prevalence is highest (Medicine), but has a U-shape in the area where the OA prevalence is lowest (Economics, Econometrics & Finance).

[Table 5 about here]

For the aggregate citations in 2017-2020, the average OA citation advantage varies in the range 41.4–62.4%, with a mean for the aggregate areas of 50.3%. The highest
average reaches in Arts & Humanities, while the lowest observes in Economics, Econometrics & Finance.

[Table 6 about here]

The outliers observed in the data distribution can skew the mean. However, half of the journals have OA citation advantage above the median of the distribution (and the other half below). Thus, the median (Table 6) is more robust measure of central tendency than the mean for data with such a high variance. The median OA citation advantage in 2017-2020 varies among fields in the range 26.9–49.4%, being 36.8% its value for the aggregate areas. The highest median reaches in Arts & Humanities, while the lowest observes in Medicine.

Thus, we can conclude that the citation advantage of the OA modality in hybrid journals, in relation to the paywalled modality, is around 50.3% in average for the 200 journals and four years in the sample, and higher than 36.8% in half of the journals. Moreover, this OA citation advantage held in time. Finally, the highest OA citation advantage is observed in Arts & Humanities.

Conclusions

Access to academic literature is a current debate in the research community. Research funders are increasingly mandating OA dissemination while, at the same time, the growth in costs have led more and more university libraries to cancel some subscriptions.

In this paper, the “research articles” in the year 2017 from 200 hybrid journals in four subject areas, and the citations received by such articles in the period 2017-2020 in the Scopus database, were analyzed. The journals were randomly selected from those with share of OA papers higher than some minimal value. More than 60 thousand research articles were analyzed in the sample, of which 24% under the OA modality.

Interestingly, we found that in general, the citations per article in both hybrid modalities strongly correlate. The hybrid journals with the greatest impact in one modality are also in the other. The evidence for this result is weaker in the field of Physics. However, there is no correlation between the OA prevalence, this is the proportion of
articles in the OA modality of the hybrid journal, and cites per article in any of the hybrid modalities.

We found that there is OA citation advantage in 80% of hybrid journals. This result is strong both across fields and over time. The number of journals with OA citation advantage is higher in Medicine (88%) and Humanities (82%).

We found that the average OA citation advantage increases with time in the field where the OA prevalence is highest (Medicine), but has a U-shape in the field with lowest OA prevalence (Economics). The average OA citation advantage in 2017-2020 varies among fields in the range 41–62%, with an aggregate mean of 50%. The highest average is obtained in Humanities, while the lowest is observed in Economics.

The median OA citation advantage in 2017-2020 varies in the range 27–49% according to fields, being 37% its value for the aggregate fields. The highest median is observed again in Humanities, while the lowest is obtained in Medicine.

Thus, we can conclude that the citation advantage of the OA modality in hybrid journals, in relation to the paywalled modality, is around 50% in average for the 200 journals and four years in the sample, and higher than 37% in half of the journals. Moreover, the OA citation advantage is consistent across fields and held in time. Finally, the OA citation advantage is higher in Humanities than in Science and Social Science.

There are some considerations in this regard. Some journals in the random sample have been cataloged by the Scopus database as Humanities, but are actually at the intersection with other areas. Notice that there are journals assigned to two different subject categories from two different areas. Indeed, these journals that employ scientific methods with applications to the Humanities receive more citations than pure humanistic journals. Therefore, the results obtained for this area must be taken with caution.

On the reliability of the data source, Unpaywall is indirectly used (through Scopus) to determine the publication modality in hybrid journals. Notice that Unpaywall is based on algorithms and not on indexing. This is the reason why, regardless of the discipline, the OA finder Unpaywall does not locate as many OA versions of journal articles as manual searches (Piwowar et al., 2018; Sergiadis, 2019).

Our study refines previous results by comparing documents more similar to each other, both in discipline and citation potential. Some of the citation advantage in the open
access modality is likely due to more access allows more people to read and hence cite articles they otherwise would not. However, causation is difficult to establish and there are many possible bias. Several factors can affect the observed differences in citation rates. Funder mandates can be one of them. Funders are likely to have OA requirement, and well-funded studies are more likely to receive more citations than poorly funded studies (Aagaard et al., 2020).

Another discussed factor is the selection bias postulate (Craig et al., 2007), which suggests that authors choose only their most impactful studies to be open access. Selection bias can occur in both paid open access journals (gold OA) and hybrid journals. This is due to researchers who have financial resources to publish their results prioritize the publication in open access those papers that they consider may have a greater impact. The current study does not examine the cause of the observed citation advantage, but does find that it exists in a very large sample.
References

Aagaard, K., A. Kladakis, & M. W. Nielsen (2020) ‘Concentration or dispersal of research funding?’, Quantitative Science Studies, 1/1: 117–149.

Antelman, K. (2004) ‘Do open-access articles have a greater research impact?’, College & Research Libraries, 65/5: 372–382.

Archambault, E., et al. (2014) ‘Proportion of open access papers published in peer-reviewed journals at the European and world levels - 1996–2013’. Report for the European Commission. Available at https://digitalcommons.unl.edu/scholcom/8/

Sergiadis, A. D. R. (2019) ‘Evaluating Zotero, SHERPA/RoMEO, and Unpaywall in an institutional repository workflow’, Journal of Electronic Resources Librarianship, 31/3: 152–176.

Björk, B. C. (2004) ‘Open access to scientific publications: an analysis of the barriers to change?’, Information Research, 9/2: 170.

Björk, B. C. (2017) ‘Growth of hybrid open access, 2009-2016’, PeerJ, 5: e3878.

Björk, B. C., & D. Solomon (2012) ‘Open access versus subscription journals: a comparison of scientific impact’, BMC Medicine, 10/1: 73.

Craig, I. D., et al. (2007) ‘Do open access articles have greater citation impact? A critical review of the literature’, Journal of Informetrics, 1/3: 239–248.

Davis, P. M. (2011) ‘Open access, readership, citations: a randomized controlled trial of scientific journal publishing’, FASEB Journal, 25: 2129–2134.

Davis, P. M., et al. (2008) ‘Open access publishing, article downloads, and citations: randomised controlled trial’, British Medical Journal, 337/7665: 343–345.

Davis, P. M., & W. H. Walters (2011) ‘The impact of free access to the scientific literature: a review of recent research’, Journal of the Medical Library Association, 99/3: 208–208.

Dorta-González, P., S. M. González-Betancor, & M. I. Dorta-González (2017) ‘Reconsidering the gold open access citation advantage postulate in a multidisciplinary context: an analysis of the subject categories in the Web of Science database 2009-2014’, Scientometrics, 112/2: 877–901.

Dorta-González, P., & Y. Santana-Jiménez (2018) ‘Prevalence and citation advantage of gold open access in the subject areas of the Scopus database’, Research Evaluation, 27/1: 1–15.

Dorta-González, P., R. Suárez-Vega, & M. I. Dorta-González (2020) ‘Open access effect on uncitedness: a large-scale study controlling by discipline, source type and visibility’, Scientometrics, 124/3: 2619–2644.
Eysenbach, G. (2006) ‘Citation advantage of open access articles’, *PLoS Biology*, 4/5: e157.

Gargouri, Y., et al. (2010) ‘Self-selected or mandated, open access increases citation impact for higher quality research’, *PLoS ONE*, 5/10: e13636.

Gaule, P., & N. Maystre (2011) ‘Getting cited: does open access help?’ *Research Policy*, 40/10: 1332–1338.

González-Betancor, S. M., & P. Dorta-González (2019) ‘Publication modalities ‘article in press’ and ‘open access’ in relation to journal average citation’, *Scientometrics*, 120/3: 1209–1223.

Gumpenberger, C., M. A. Ovalle-Perandones, & J. Gorraiz (2013) ‘On the impact of gold open access journals’, *Scientometrics*, 96/1: 221–238.

Harnad, S., et al. (2004) ‘The access/impact problem and the green and gold roads to open access’, *Serials Review*, 30/4: 310–314.

Joint, N. (2009) ‘The Antaeus column: does the “open access” advantage exist? A librarian’s perspective’, *Library Review*, 58/7: pp. 477–481.

Koler-Povh, T., P. Južnič, & G. Turk (2014) ‘Impact of open access on citation of scholarly publications in the field of civil engineering’, *Scientometrics*, 98/1033: 1033–1045.

Kurtz, M. J., et al. (2005) ‘The effect of use and access on citations’, *Information Processing & Management*, 41/6: 1395–1402.

Laakso, M., & B. C. Björk (2013) ‘Delayed open access: an overlooked high-impact category of openly available scientific literature’, *Journal of the American Society for Information Science and Technology*, 64/7: 1323–1329.

Lawrence, S. (2001) ‘Free online availability substantially increases a paper’s impact’, *Nature*, 411/6837: 521–521.

McCabe, M., & C. Snyder (2014) ‘Identifying the effect of open access on citations using a panel of science journals’, *Economic Inquiry*, 52/4: 1284–1300.

Moed, H. F. (2007) ‘The effect of open access on citation impact: an analysis of ArXiv's condensed matter section’, *Journal of the American Society for Information Science and Technology*, 58/13: 2047–2054.

Niyazov, Y., et al. (2016) ‘Open access meets discoverability: citations to articles posted to Academia.edu’, *PLoS ONE*, 11/2: e0148257.

Norris, M., C. Oppenheim, & F. Rowland (2008) ‘The citation advantage of open-access articles’, *Journal of the American Society for Information Science and Technology*, 59/12: 1963–1972.
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: e0159614.

Piwowar, H., et al. (2018) ‘The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles’, *PeerJ*, 6: e4375.

Prosser, D. (2003) ‘Institutional repositories and open access: the future of scholarly communication’, *Information Services and Use*, 23/2: 167–170.

Solomon, D. J., M. Laakso, & B. C. Björk (2013) ‘A longitudinal comparison of citation rates and growth among open access journals’, *Journal of Informetrics*, 7/3: 642–650.

Tenopir, T., et al. (2017) ‘Imagining a gold open access future: attitudes, behaviors, and funding scenarios among authors of academic scholarship’, *College & Research Libraries*, 78/6: 824–843.

Wang, X., et al. (2015) ‘The open access advantage considering citation, article usage and social media attention’, *Scientometrics*, 103/2: 555–564.
**Table 1. Description of the subject areas in the study**

| Subject Area               | OA Hybrid | %   | Other modalities * | %   | Total    |
|----------------------------|-----------|-----|--------------------|-----|----------|
| Arts & Humanities          | 2,821     | 3.7%| 74,458             | 96.3% | 77,279   |
| Economics, Econometrics & Finance | 1,097    | 2.5%| 43,376             | 97.5% | 44,473   |
| Medicine                   | 23,243    | 4.6%| 485,260            | 95.4% | 508,503  |
| Physics & Astronomy        | 6,635     | 2.7%| 237,666            | 97.3% | 244,301  |
| **Aggregate Areas**        | **33,796**| **3.9%**| **840,760**       | **96.1%** | **874,556** |
| Scopus database            | 69,093    | 3.4%| 1,951,700          | 96.6% | 2,020,793|
| **%**                      | **48.9%** |     | **43.1%**          |       | **43.3%** |

*Paywalled modality in hybrid journals, paywalled journals and OA journals

**Table 2. Representativeness of the sample**

| Subject Area               | Sample | Population | Sample % |
|----------------------------|--------|------------|----------|
|                            | OA Hybrid | Paywalled | OA Hybrid | OA Hybrid | Sample % |
| Arts & Humanities          | 1,151   | 5,759      | 2,821     | 40.8%     |
| Economics, Econometrics & Finance | 541     | 5,411      | 1,097     | 49.3%     |
| Medicine                   | 4,381   | 15,772     | 23,243    | 18.8%     |
| Physics & Astronomy        | 1,970   | 27,623     | 6,635     | 29.7%     |
| **Total**                  | **8,043**| **54,565** | **33,796**| **23.8%** |
Table 3. Pearson's linear correlation coefficient

| Arts & Humanities | Best CiteScore Percentile 2017 | Research Articles 2017 | OA Prevalence | OA Cites per Article | Paywalled Cites per Article |
|-------------------|--------------------------------|------------------------|---------------|----------------------|-----------------------------|
|                   | 1.00                           | 0.03                   | -0.69         | 0.50                 | 0.57                        |
| Research Articles 2017 | 0.03                           | 1.00                   | -0.21         | 0.30                 | 0.42                        |
| OA Prevalence     | -0.69                          | -0.21                  | 1.00          | -0.39                | -0.40                       |
| OA Cites per Article | 0.50                           | 0.30                   | -0.39         | 1.00                 | 0.81                        |
| Paywalled Cites per Article | 0.57                           | 0.42                   | -0.40         | 0.81                 | 1.00                        |

| Economics, Econometrics & Finance | Best CiteScore Percentile 2017 | Research Articles 2017 | OA Prevalence | OA Cites per Article | Paywalled Cites per Article |
|----------------------------------|--------------------------------|------------------------|---------------|----------------------|-----------------------------|
|                                 | 1.00                           | -0.16                  | 0.07          | 0.60                 | 0.57                        |
| Research Articles 2017           | -0.16                          | 1.00                   | -0.52         | 0.13                 | 0.11                        |
| OA Prevalence                    | 0.07                           | -0.52                  | 1.00          | -0.14                | -0.12                       |
| OA Cites per Article             | 0.60                           | 0.13                   | -0.14         | 1.00                 | 0.85                        |
| Paywalled Cites per Article      | 0.57                           | 0.11                   | -0.12         | 0.85                 | 1.00                        |

| Medicine                        | Best CiteScore Percentile 2017 | Research Articles 2017 | OA Prevalence | OA Cites per Article | Paywalled Cites per Article |
|---------------------------------|--------------------------------|------------------------|---------------|----------------------|-----------------------------|
|                                 | 1.00                           | -0.48                  | -0.22         | 0.29                 | 0.34                        |
| Research Articles 2017          | -0.48                          | 1.00                   | -0.14         | -0.18                | -0.20                       |
| OA Prevalence                   | -0.22                          | -0.14                  | 1.00          | 0.01                 | -0.02                       |
| OA Cites per Article            | 0.29                           | -0.18                  | 0.01          | 1.00                 | 0.97                        |
| Paywalled Cites per Article     | 0.34                           | -0.20                  | -0.02         | 0.97                 | 1.00                        |

| Physics & Astronomy              | Best CiteScore Percentile 2017 | Research Articles 2017 | OA Prevalence | OA Cites per Article | Paywalled Cites per Article |
|----------------------------------|--------------------------------|------------------------|---------------|----------------------|-----------------------------|
|                                 | 1.00                           | 0.12                   | 0.00          | 0.33                 | 0.52                        |
| Research Articles 2017           | 0.12                           | 1.00                   | -0.35         | 0.03                 | 0.15                        |
| OA Prevalence                    | 0.00                           | -0.35                  | 1.00          | -0.17                | -0.17                       |
| OA Cites per Article             | 0.33                           | 0.03                   | -0.17         | 1.00                 | 0.49                        |
| Paywalled Cites per Article      | 0.52                           | 0.15                   | -0.17         | 0.49                 | 1.00                        |

Note: (a) The OA prevalence is the proportion of articles in the OA modality of the hybrid journal. (b) We use the term ‘Best percentile’ because a journal may be assigned to several subject fields and have different percentiles in each of them.
Figure 1. Scatter plot for cites per article in both hybrid modalities. Average across all citation years for the 200 hybrid journals in the sample. Bubble size proportional to OA prevalence.
Figure 2. Box and whisker plot (without outliers) for the distribution of cites per article by hybrid modality and year of citation. Average in the citation years for the 200 hybrid journals in the sample.
Figure 3. OA citation advantage in relation to the best CiteScore percentile. We use the term ‘Best percentile’ because a journal may be assigned to several subject fields and have different percentiles in each of them.
Figure 4. OA citation advantage by subject areas
Figure 5. OA citation advantage along time
### Table 4. Number of journals with OA citation advantage

| Subject Area                          | 2017 | 2018 | 2019 | 2020 | 2017-2020 |
|---------------------------------------|------|------|------|------|-----------|
| Arts & Humanities                     | 33   | 41   | 40   | 39   | 41        |
|                                       | 66%  | 82%  | 80%  | 78%  | 82%       |
| Economics, Econometrics & Finance     | 35   | 38   | 34   | 40   | 37        |
|                                       | 70%  | 76%  | 68%  | 80%  | 74%       |
| Medicine                              | 39   | 43   | 44   | 40   | 44        |
|                                       | 78%  | 86%  | 88%  | 80%  | 88%       |
| Physics & Astronomy                   | 38   | 38   | 35   | 36   | 37        |
|                                       | 76%  | 76%  | 70%  | 72%  | 74%       |
| Aggregate areas                       | 145  | 160  | 153  | 155  | 159       |
|                                       | 73%  | 80%  | 77%  | 78%  | 80%       |

### Table 5. Mean of the OA citation advantage

| Subject Area                          | 2017 | 2018 | 2019 | 2020 | 2017-2020 |
|---------------------------------------|------|------|------|------|-----------|
| Arts & Humanities                     | 110.6% | 102.8% | 34.4% | 82.5% | 62.4%     |
| Economics, Econometrics & Finance     | 83.9%  | 28.4%  | 28.8% | 64.1% | 41.4%     |
| Medicine                              | 41.7%  | 43.5%  | 44.7% | 47.4% | 44.3%     |
| Physics & Astronomy                   | 41.4%  | 48.6%  | 57.0% | 56.3% | 53.1%     |
| Aggregate areas                       |       |       |      |      | 50.3%     |

### Table 6. Median of the OA citation advantage

| Subject Area                          | 2017 | 2018 | 2019 | 2020 | 2017-2020 |
|---------------------------------------|------|------|------|------|-----------|
| Arts & Humanities                     | 38.9% | 62.9% | 31.8% | 44.1% | 49.4%     |
| Economics, Econometrics & Finance     | 41.8%  | 37.9%  | 24.2% | 45.8% | 39.0%     |
| Medicine                              | 29.2%  | 32.1%  | 31.5% | 28.0% | 26.9%     |
| Physics & Astronomy                   | 30.5%  | 30.0%  | 30.7% | 24.8% | 27.9%     |
| Aggregate areas                       |       |       |      |      | 36.8%     |
## ANNEX A. Dataset

| Journal                                      | Best Citation Paywalled 2017 | Modality | Research Articles 2017 | % | Cites 2017 | Cites 2018 | Cites 2019 | Cites 2020 | Total Cites | Cites per Article | OA Citation Advantage |
|----------------------------------------------|------------------------------|----------|-------------------------|---|------------|------------|------------|------------|-------------|----------------------|-----------------------|
| Art and Humanities                           |                              |          |                         |   |            |            |            |            |             |                      |                       |
| Annales de Litterature                   | Paywalled                    | 28       | 45.5%                   |   | 0          | 0          | 0          | 0          | 0          | 0.00%                | 53.8%                 |
| Ramo                                          | OA                            | 10        | 15.3%                   |   | 0          | 0          | 0          | 0          | 0          | 0.00%                | -118.2%               |
| Archaeological and Anthropological Sciences | Paywalled                    | 102      | 90.3%                   |   | 86         | 157        | 202        | 212        | 662        | 6.69%                | 46.3%                 |
| Ramo                                          | OA                            | 12        | 10.5%                   |   | 6          | 27         | 20          | 25          | 78          | 3.13%                | 37.7%                 |
| Archaeology and Anthropology of Foods      | Paywalled                    | 13       | 11.6%                   |   | 10         | 14         | 7           | 14          | 41          | 4.00%                | 38.8%                 |
| Archaeometry                                | Paywalled                    | 64       | 89.5%                   |   | 35         | 62         | 83          | 93          | 265         | 4.14%                | 69.3%                 |
| Ramo                                          | OA                            | 8         | 11.5%                   |   | 2          | 14         | 14          | 21          | 45          | 2.63%                | 45.3%                 |
| Archives of Design Research                | Paywalled                    | 14       | 92.9%                   |   | 0          | 1          | 2           | 1           | 4           | 9.29%                | 219.3%                |
| Ramo                                          | OA                            | 6         | 11.8%                   |   | 10         | 21         | 12          | 12          | 45          | 3.83%                | 54.8%                 |
| Archives of Sexual Behavior                | Paywalled                    | 187      | 98.7%                   |   | 214        | 484        | 692         | 732         | 2137        | 11.05%               | 28.4%                 |
| Ramo                                          | OA                            | 11        | 6.3%                    |   | 30         | 45         | 61          | 66          | 136         | 12.20%               |                       |
| Arms and Art                                  | Paywalled                    | 31       | 97.6%                   |   | 0          | 0          | 0          | 0          | 0          | 0.00%                | 60.4%                 |
| Ramo                                          | OA                            | 11        | 97.6%                   |   | 0          | 0          | 0          | 0          | 0          | 0.00%                | 39.3%                 |
| American Exception and Psychopathesis        | Paywalled                    | 63       | 89.5%                   |   | 35         | 62         | 83          | 93          | 265         | 4.14%                | 69.3%                 |
| Ramo                                          | OA                            | 11        | 97.6%                   |   | 0          | 0          | 0          | 0          | 0          | 0.00%                | 60.4%                 |
| Mental Health and Disability Studies       | Paywalled                    | 50       | 18.6%                   |   | 52        | 165        | 225         | 315         | 755         | 28.72%               | 75.9%                 |
| Ramo                                          | OA                            | 11        | 6.3%                    |   | 30         | 45         | 61          | 66          | 136         | 12.20%               |                       |
| Anatomy                                     | Paywalled                    | 209      | 98.3%                   |   | 109        | 111        | 122         | 125         | 411         | 28.72%               | 65.3%                 |
| Ramo                                          | OA                            | 11        | 6.3%                    |   | 30         | 45         | 61          | 66          | 136         | 12.20%               |                       |
| Applied Economics                           | Paywalled                    | 410      | 97.6%                   |   | 356        | 506        | 732         | 712         | 2137        | 4.26%                | 158.2%                |
| Ramo                                          | OA                            | 11        | 6.3%                    |   | 30         | 45         | 61          | 66          | 136         | 12.20%               |                       |

The table above represents a dataset with various metrics related to different journals and their publication years, including the number of research articles published, the percentage of paywalled content, and various citation metrics.
| Title                                                                 | Paywall | OA     | OA     |
|----------------------------------------------------------------------|---------|--------|--------|
| Applied Economics Letters                                            | 96      | 57.9%  | 43     |
| Applied Health Economics and Health Policy                           | 63      | 20.3%  | 32     |
| International Economics                                              | 72      | 10.7%  | 10     |
| Ecological Economics                                                 | 62      | 7.9%   | 5      |
| Economic Geography                                                    | 64      | 7.6%   | 4      |
| Economic Journal                                                     | 91      | 12.5%  | 10     |
| Economic Theory                                                       | 65      | 13.7%  | 6      |
| Economics Letters                                                     | 58      | 95.8%  | 54     |
| Econometrica Netherland                                               | 63      | 52.8%  | 45     |
| Economy                                                               | 72      | 10.9%  | 7      |
| Empirical Economics                                                   | 64      | 96.1%  | 66     |
| Energy Economics                                                      | 93      | 94.6%  | 138    |
| Environmental and Resource Economics                                 | 61      | 68.4%  | 56     |
| European Journal of Health Economics                                 | 95      | 15.2%  | 78     |
| Experimental Economics                                                | 64      | 84.2%  | 55     |
| Finance and Electronics                                              | 67      | 37.6%  | 29     |
| Fiscal Studies                                                        | 63      | 25.6%  | 26     |
| Food Policy                                                           | 96      | 32.7%  | 23     |
| Forest Policy and Economics                                           | 92      | 96.9%  | 134    |
| Global Policy                                                         | 74      | 93.5%  | 80     |
| Globalisation                                                         | 68      | 74.9%  | 54     |
| International Environmental Agreements                               | 91      | 84.9%  | 51     |
| International Journal of Agricultural Sustainability                 | 89      | 79.7%  | 37     |
| International Journal of Game Theory                                | 94      | 9.6%   | 1      |
| International Journal of Production Economics                         | 91      | 3.4%   | 5      |
| International Journal on Macroeconomic Studies                        | 96      | 94.8%  | 218    |
| Journal of Development Economics                                     | 97      | 3.9%   | 4      |
| Journal of Economics                                                  | 82      | 9.4%   | 10     |
| Journal of Economic Behavior and Organization                        | 75      | 96.9%  | 51     |
| Journal of Economic Inequality                                       | 86      | 73.9%  | 12     |
| Journal of Environmental Economics and Management                     | 90      | 70.1%  | 43     |
| Journal of International Economics                                   | 90      | 37.7%  | 30     |
| Journal of International Money and Finance                            | 88      | 95.9%  | 32     |
| Journal of Public Economics                                           | 99      | 3.4%   | 4      |
| Journal of Risk and Uncertainty                                      | 82      | 9.4%   | 1      |
| Journal of the Academy of Marketing Science                          | 97      | 10.9%  | 3      |
| Journal of the European Economic Association                          | 99      | 13.9%  | 5      |
| Journal of Urban Economics                                            | 86      | 87.5%  | 17     |
| Letters in Spinal and Resource Science                               | 52      | 40.0%  | 9      |
| Marine Policy                                                         | 96      | 90.1%  | 144    |
| Medium Choice                                                         | 76      | 54.3%  | 14     |
| Quantitative Finance                                                  | 67      | 95.6%  | 36     |
| Resources Policy                                                      | 96      | 83.8%  | 30     |
| Review of Income and Wealth                                           | 72      | 0.9%   | 4      |
| Review of International Political Economy                            | 97      | 29.8%  | 32     |
| Small Business Economies                                              | 95      | 80.8%  | 54     |
| Social Choice and Welfare                                            | 54      | 29.3%  | 9      |
| Sociological and Economic Development of Economy                      | 93      | 12.3%  | 14     |
| Theory and Decision                                                   | 70      | 53.3%  | 10     |
| World Development                                                     | 97      | 98.1%  | 203    |
| Medicine                                                              | 87      | 39.0%  | 29     |

| Title                                                                 | Paywall | OA     | OA     |
|----------------------------------------------------------------------|---------|--------|--------|
| Advances in Therapy                                                  | 87      | 38.9%  | 26     |
| American Journal of Cardiology                                       | 63      | 84.5%  | 44     |
| American Journal of Preventive Medicine                             | 97      | 70.2%  | 56     |
| American Journal of Tropical Medicine and Hygiene                    | 66      | 61.6%  | 47     |
| Animal Ecology                                                       | 94      | 64.3%  | 53     |
| Annals of the Rheumatic Diseases                                     | 99      | 85.3%  | 83     |

| Title                                                                 | Paywall | OA     | OA     |
|----------------------------------------------------------------------|---------|--------|--------|
| Annals of the Rheumatic Diseases                                     | 99      | 85.3%  | 83     |
| Annals of the Rheumatic Diseases                                     | 99      | 85.3%  | 83     |
| Annals of the Rheumatic Diseases                                     | 99      | 85.3%  | 83     |
| Annals of the Rheumatic Diseases                                     | 99      | 85.3%  | 83     |
| Title                                                                 | Publisher       | Volume | Issue | Pages |
|----------------------------------------------------------------------|-----------------|--------|-------|-------|
| Communications in Mathematical Physics                              |                  | 89     |       | 211-216 |
| European Physical Journal B                                         |                  | 92     |       | 217-218 |
| European Physical Journal H                                         |                  | 87     |       | 219-220 |
| Experiments in Physics                                              |                  | 82     |       | 221-222 |
| Japanese Journal of Applied Physics                                 |                  | 91     |       | 223-224 |
| Journal of Chemical Physics                                          |                  | 93     |       | 225-226 |
| Journal of Computational Physics                                     |                  | 91     |       | 227-228 |
| Journal of Fluid Mechanics                                          |                  | 93     |       | 229-230 |
| Journal of Measurement Science and Journal of Thermophysical Physics |                  | 92     |       | 231-232 |
| Journal of Physics A Mathematical and Theoretical                   |                  | 91     |       | 233-234 |
| Journal of Physics: Condensed Matter                                |                  | 92     |       | 235-236 |
| Journal of Physics D Applied Physics                                 |                  | 93     |       | 237-238 |
| Journal of Quant Spectroscopy and Radiative Transfer                 |                  | 94     |       | 239-240 |
| Journal of Statistical Physics                                       |                  | 93     |       | 241-242 |
| Journal of Synchronous Radiation                                     |                  | 95     |       | 243-244 |
| Journal of the Acoustical Society of America                         |                  | 91     |       | 245-246 |
| Journal of the Physical Society of Japan                             |                  | 93     |       | 247-248 |
| Journal of Thermo Analysis and Calorimetry                          |                  | 95     |       | 249-250 |
| Measurement Science and Technology                                   |                  | 97     |       | 251-252 |
| Nano Letters                                                         |                  | 93     |       | 253-254 |
| Nuclear Fusion                                                       |                  | 97     |       | 255-256 |
| Nuclear Physics A                                                    |                  | 95     |       | 257-258 |
| Optical and Quantum Electronics                                      |                  | 52     |       | 259-260 |
| Optical Engineering                                                 |                  | 74     |       | 261-262 |
| Physical Review Applied                                              |                  | 91     |       | 263-264 |
| Physical Review C                                                    |                  | 69     |       | 265-266 |
| Physics of Fluids                                                    |                  | 74     |       | 267-268 |
| Physics of Plasmas                                                   |                  | 95     |       | 269-270 |
| Quantum Information Processing                                       |                  | 93     |       | 271-272 |
| Review of Scientific Instruments                                    |                  | 93     |       | 273-274 |
| Soft Matter                                                          |                  | 91     |       | 275-276 |
| Solar Physics                                                        |                  | 74     |       | 277-278 |
| Ultramicroscopy                                                      |                  | 67     |       | 279-280 |
| Soft Matter                                                          |                  | 91     |       | 281-282 |
| Superschool Science and Technology                                   |                  | 90     |       | 283-284 |
| Vibrations                                                          |                  | 94     |       | 285-286 |