

# No Deal: Investigating the Influence of Restricted Access to Elsevier Journals on German Researchers' Publishing and Citing Behaviours

Nicholas Fraser, Anne Hobert, Najko Jahn, et al.

## Abstract

In 2014, a union of German research organisations established Projekt DEAL, a national-level project to negotiate licensing agreements with large scientific publishers. Negotiations between the DEAL consortium and Elsevier, one of the world's largest scientific publishers, began in 2016, and broke down without a successful agreement in 2018; in this time, around 200 German research institutions cancelled their existing license agreements with Elsevier, leading to Elsevier restricting access to their journal portfolios at those institutions from July 2018 onwards. In this study we assess the effect of Elsevier access restrictions on the publishing and citing behaviour of researchers at the affected institutions. [Add a brief summary of results and conclusions].

## Background

In 2014 the [Alliance of Science Organisations](#) (*Allianz der Wissenschaftsorganisationen*; AWO), a union of the majority of German research organisations, established a national-level project named “Projekt DEAL” (herein referred to as “DEAL”), to negotiate licensing agreements for access to electronic journals of large scientific publishers. The key objectives of DEAL are:

1. To receive permanent, full-text access to the entire journal portfolio of the selected publishers.
2. To make all articles published by German authors automatically Open Access (OA) under CC-BY licenses.
3. To secure reasonable pricing according to a simple, future-oriented model based on publication volumes.

To date, DEAL negotiations have centred on three major publishers: Elsevier, Springer Nature and Wiley. Between 2012 and 2020, these three publishers were collectively responsible for publishing ~54% of scientific articles with at least a single author from a DEAL institutions (Elsevier ~23%, Springer Nature ~18%, Wiley ~12%; data according to [Dimensions](#)). Negotiations between DEAL and Elsevier officially began in 2016, with Springer Nature and Wiley negotiations beginning a year later in 2017. In January 2019, DEAL announced the signing of a Publish And Read (PAR) agreement with Wiley, fulfilling the defined negotiating objectives by allowing full access to Wiley's portfolio of journals for DEAL institutions, and automatic publishing of articles under OA licenses, for an annual fee equal to €2,750 per published article ([Sander et al., 2019](#)). In January 2020 a similar PAR agreement between DEAL and Springer Nature was signed, including the same per-article fee equal to €2,750 ([Kieselbach, 2020](#)).

Whilst negotiations with Wiley and Springer Nature have now concluded in publishing agreements, negotiations with Elsevier remain unresolved. At the end of 2016, ~70 German institutions ([archived list](#)) chose not to renew their contracts with Elsevier, leading to Elsevier restricting access to new journal issues at those institutions (and also restricting access to back-catalogues at some institutions) from the beginning of 2017 ([Vogel, 2017a](#)), although access was restored 6 weeks later ([Vogel, 2017b](#)). At the end of 2017, a further ~110 German institutions ([archived list](#)) decided not to renew their contracts with Elsevier, and at the beginning of July 2018, the [German Rectors' Conference](#) (*Hochschulrektorenkonferenz*; HRK), who are leading negotiations on behalf of AWO, [announced](#) the breakdown and cancellation of all ongoing negotiations with Elsevier. In mid-July 2018, authors at institutions which had cancelled their contracts with Elsevier had their access to new journal issues completely cut-off ([Else, 2018](#)). A further ~25 institutions, including the Max Plank Society and Fraunhofer Society ([archived list](#)) did not renew their contracts with Elsevier at the end of 2018.

## Motivation and Research Questions

As a provider of a large proportion of research published and cited by German researchers, restricted access to Elsevier’s article collections may have measurable effects on the publication and citation patterns of researchers at DEAL institutions. Attempts to quantify such effects have already been made through recent survey approaches: [A survey commissioned by Elsevier](#) found that 61% of German researchers agreed or strongly agreed that losing access made their research activities less efficient, whilst 54% agreed or strongly agreed that losing access delayed the speed that they produce their research output. A separate survey of 384 researchers at the Faculty of Medicine of the University of Münster showed an overall similar sentiment, with 66% of researchers reporting that they now require more time to acquire literature and 46% of researchers reporting that losing access was a competitive disadvantage, yet only 29% of researchers reported that they would no longer write or review articles for Elsevier journals ([Obst, 2019](#)).

The situation in Germany is not unique and breakdowns in negotiations between library consortia and Elsevier have been reported elsewhere. In Sweden, a number of universities, research institutes and government agencies were cut off from Elsevier between mid-2018 and the end of 2019 due to a breakdown in negotiations between Elsevier and the [Bibsam Consortium](#) (the national-level license negotiating body for Sweden). A [PAR agreement](#) was eventually signed between Bibsam and Elsevier at the end of 2019, to take effect from 1st January 2020. A large-scale survey of 4,221 Swedish researchers carried out by the Bibsam Consortium during the time period that Elsevier journals were inaccessible found that 51% of respondents were negatively affected in their desire to publish with Elsevier, and 54% had their work negatively impacted ([Olson et al., 2020](#)). The University of California (UC) system also recently signed a [Memorandum of Understanding](#) with Elsevier following the suspension of negotiations, and shutoff to Elsevier access, in 2019.

Despite these survey approaches, to our knowledge no ground truth on the effect of restricted access to Elsevier journals from a bibliometric perspective has yet been established. In this study we therefore aim to conduct a bibliometric investigation of the following research questions:

1. Did restricted access to Elsevier journals at DEAL institutions result in a change in researchers’ publishing behaviour?
2. Did restricted access to Elsevier journals at DEAL institutions result in a change in researchers’ citing behaviour?

For both of these research questions, we also consider variability with respect to research disciplines, collaboration patterns, and article OA status.

## Methods

### Data Sources

**DEAL institutions** We collected names and contract expiration dates of 210 universities, research institutions, higher education institutions and regional libraries that had their access to Elsevier articles restricted as part of the DEAL negotiations, using publicly available information on the DEAL website (see archived lists of institutions with contracts that expired at the end of [2016](#), [2017](#) and [2018](#)). We manually mapped institutions to identifiers in the [Global Research Identifier Database](#) (GRID), using the available search interface on the GRID website. Of the 210 institution names provided, 209 were matched to a GRID identifier, with the exception of “HS Villingen-Schwenningen”, for which we were unable to unambiguously determine the correct GRID identifier. The list extracted from the Project DEAL website contained 10 individual records for each campus associated with the Baden-Wuerttemberg Cooperative State University (“Duale Hochschule Baden-Württemberg / DHBW”), but all campus are collectively associated with a single ID in the GRID database (“grid.449295.7”). The Max Planck Society and Fraunhofer Society were listed individually on the DEAL website, but both are umbrella associations that consist of a number of individual research institutions. Thus, we also extracted the GRID information for all their constituents, according to “parent-child” relationship information stored in GRID. The Helmholtz Association and Leibniz Association are similar umbrella associations, but the lists from the DEAL website contained the names of individual Helmholtz Association and Leibniz Association institutions, thus we limited the dataset to those contained directly in the list and did not extract information for all other constituent institutions. We attempted to

verify the information contained on the DEAL website by manually searching for press releases or other informational web pages issued or maintained by the individual institutions that referred to restricted access to Elsevier articles. Of the original 210 institutions on the list, we found relevant information for 121 institutions; the information contained within each page matched well with the information contained in the DEAL lists (e.g. in terms of contract status and timing of restricted access).

**Article and author metadata** Article and author metadata used in this study were derived from three main bibliometric data sources: Dimensions, Crossref and Unpaywall. Initially, we retrieved article DOIs, complete author and affiliation details, fields of research, and reference lists (DOI-DOI links) for all articles with at least a single author from a DEAL institution, via the Dimensions Analytics API. These data were retrieved in the first two weeks of April 2021. Articles were limited to those with a publication date in years 2012 to 2020, and to “article” publication types. As the Dimensions Analytics API only allows a maximum of 50,000 records to be returned in a single query, we queried iteratively through each year and individual DEAL institution, using the associated GRID identifier, and extracted details of all articles that included an author at the respective institution. In a final step we combined all article records together and removed duplicates (e.g. where an article had authors from multiple DEAL institutions). Following these steps we created a set of 892,169 unique articles (Figure 1A).

Figure 1B shows the distribution of the number of authors per article in our original dataset. The figure shows that a high proportion of articles contain a large number of authors. In cases of articles written by large teams or consortia, the contribution of DEAL authors to the writing of the article or subsequent publication strategy may be small. Figure 1D shows the proportion of articles in our original dataset in each year divided by authorship types: “DEAL First Author” refers to articles where the first author is from a DEAL institution but the last author is not, “DEAL Last Author” where the last author is from a DEAL institution but the first author is not, “DEAL First and Last Author” where both the first and last authors are from DEAL institutions, and “DEAL Middle Author” where neither the first nor last authors are from DEAL institutions. These results show that the majority of articles in our dataset have a first or last author (or both) from DEAL institutions, yet there exist a number of articles where DEAL authors are only included as middle authors. Although practices for the assignment of author order are neither clear nor consistent across disciplines (Brand et al., 2015), for the purposes of this study we make the assumption that the publication strategy for the article is primarily determined by either the first or last author of an article. Thus, as our study aims to focus on the direct behaviour of researchers at DEAL institutions, we subsequently limited our dataset to articles with a first AND last author from a DEAL institution (i.e. the group “DEAL First and Last Author” in Figure 1D).

For each article retrieved from Dimensions, we also retrieved and parsed a complete list of references (total number of DOI-DOI reference links: 33,652,274). An overview of the distribution of references per article for our original dataset is shown in Figure 1C. We observed that an anomalously high number of articles in this distribution contained either zero references or a single reference. A manual investigation on a random sample of these articles revealed that articles with zero references often represent diverse types of editorial content (e.g. corrections, errata, tables of content) which is registered in Dimensions as “article” types. Articles with a single reference often represent abstracts which are linked to a single journal article, (see, for an example, abstracts published by the journal [ChemInform](#)). In a small number of cases, articles that did not contain references in our dataset did in fact contain a reference list on the journal page - highlighting a potential weakness in Dimensions as a data source. However, as a broad generalisation, we conclude that articles containing zero references or a single reference do not represent “true” research articles, which is the focus of this study. We thus decided to remove these articles as a source of uncertainty from our dataset for subsequent analyses.

Following the removal of articles without DEAL first and last authors, and articles with zero references or a single reference, our final dataset of articles from Dimensions was reduced to 410,084 articles (46% of the original dataset).

For determining subject classifications, we used the “Fields of Research” (FOR) scheme available in Dimensions, which is itself based on the Australian and New Zealand Standard Research Classification (ANZSRC) system (further details [here](#)). The classification scheme consists of 22 divisions at the upper level. Unlike classification

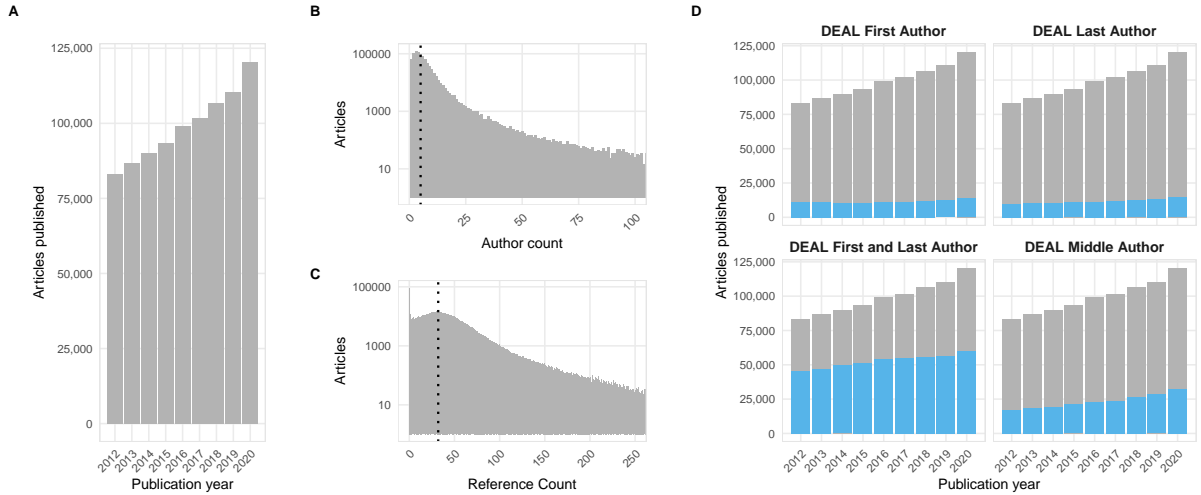


Figure 1: Properties of our initial dataset of articles published by DEAL researchers extracted from Dimensions. (A) Number of articles published per year. (B) Distribution of number of authors per article (only articles with <100 authors are shown). Note y-axis is on a log-scale. (C) Distribution of number of references per article (only articles with <250 references are shown). Note y-axis is on a log-scale. (D) Proportion of articles by authorship type: “DEAL First Author” refers to articles where the first author is from a DEAL institution but the last author is not, “DEAL Last Author” where the last author is from a DEAL institution but the first author is not, “DEAL First and Last Author” where both the first and last authors are from DEAL institutions, and “DEAL Middle Author” where neither the first nor last authors are from DEAL institutions.

systems in other bibliometric databases which classify articles on a journal level (e.g. Web of Science, Scopus), Dimensions firstly classifies articles on a single document level using a text-based classification approach. Where information is insufficient, Dimensions falls back to a journal-level classification. Some initial discussion of the strengths and weaknesses of the Dimensions approach has been conducted by Bornmann (2018) and Herzog and Kierkegaard Lunn (2018). Bornmann (2018) noted a number of inaccuracies in the classification of his own publication record in Dimensions. However, improvements to the classification system have since been implemented and Dimensions [reported an increase in the precision and recall of the method in August 2019](#).

Article records from Dimensions were matched to records in Crossref (for classification of Elsevier versus Non-Elsevier content, using the Crossref member ID of Elsevier, 78) and Unpaywall (for determination of article OA status). Matching was conducted through exact matching of DOIs: 99.7% and 99.8% of articles in our dataset from Dimensions were matched to articles in Crossref and Unpaywall via DOIs, respectively. Crossref data is based on an openly available Crossref database snapshot ([Crossref, 2021](#)) that contains all Crossref records registered until 7th January 2021. Relevant metadata fields were parsed applying the rcrossref parsers (Chamberlain et al., 2021), following the same approach documented in [Jahn et al., 2021](#). To reduce computation time and storage demands [**Check this is ok with Najko**], the Crossref dataset was subsequently limited to records registered after 1st January 2008. Unpaywall data is based on an openly available database snapshot (details available [here](#)) from February 2021. Processing of the Unpaywall dataset followed the same procedure as that documented in [Hobert et al., 2021](#).

**Data Processing, Storage and Analysis** To allow fast data processing and analysis, all large datasets described above (i.e. those from Dimensions, Crossref and Unpaywall) were imported to Google BigQuery (<https://cloud.google.com/bigquery>), a cloud data warehouse which allows querying of large datasets with an SQL vocabulary. All analysis of data was subsequently carried out in R, using the [DBI](#) and [bigquery](#) packages to interface R directly with Google BigQuery. [**To Do:** add citations for R packages used]

## Results

### Publishing behaviour of DEAL researchers

In this section we assess how restricted access to Elsevier articles has influenced publishing patterns of DEAL researchers. Whilst access restrictions have reduced the ability for DEAL researchers to read and download Elsevier articles, there exists no further barriers for DEAL researchers to *publish* in Elsevier journals beyond those that previously existed, e.g. meeting peer-review criteria, affordability of journal-specific fees, etc. However, we hypothesise that access restrictions may lead to negative sentiment amongst researchers which may influence their decision when choosing a suitable publication venue for their work; such negative desire to publish with Elsevier was reported by 51% of respondents of a survey conducted by the Bibsam Consortium when Elsevier restricted access to their journals in Sweden (Olson et al., 2020).

We assess changes in publishing behaviour of DEAL researchers through two related metrics: (1) the total number of articles published by DEAL researchers in Elsevier versus non-Elsevier journals each year, and (2) the annual change in the absolute proportion of articles published by DEAL researchers in Elsevier journals (i.e. the change in Elsevier’s market share of DEAL publications)<sup>1</sup>. With respect to both of these metrics, we consider variation with respect to the year of contract expiration between the DEAL institution and Elsevier, research disciplines, collaboration patterns and OA status of the published articles. We focus on the period between 2012 and 2020, allowing us to capture the long-term trends in the years prior to access restrictions, and the effect of access restrictions in 2017/2018 on publishing patterns in subsequent years.

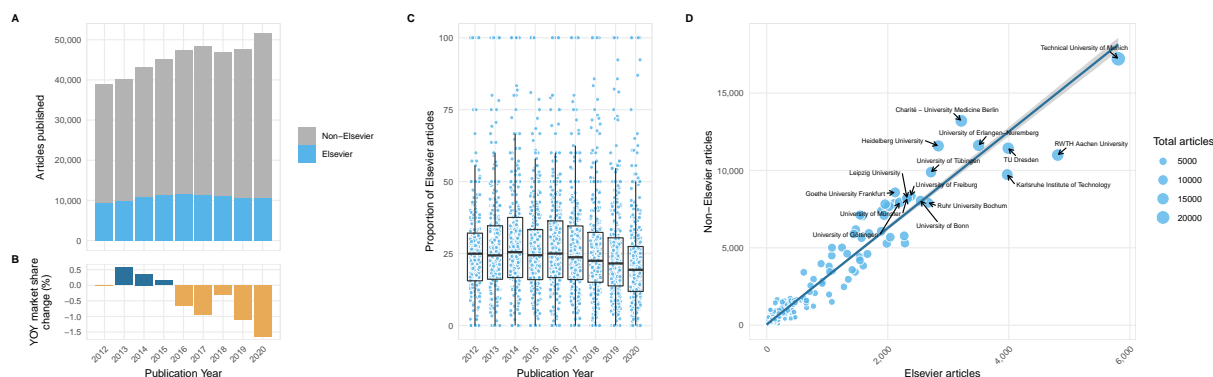


Figure 2: Publishing behaviour of DEAL researchers, 2012-2020. (A) Total number of articles published by DEAL researchers in Elsevier and non-Elsevier journals. (B) Year-on-year (YOY) change in Elsevier’s market share of articles published by DEAL researchers. (C) Proportion of articles published by individual DEAL institutions in Elsevier journals. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5\*IQR. Points denote individual institutions, with added horizontal jitter for visibility. (D) Number of Elsevier versus Non-Elsevier articles published for each individual DEAL institution (totals aggregated for years 2012-2020). Point size is scaled by the total number of articles published by an institution. Institutions that published >10,000 articles in total are labelled.

Figure 2A shows the change in the number of articles published by DEAL researchers between 2012 and 2020 (limited to articles with DEAL first and last authors). The total number of articles published per year increased during this period, from 38,849 articles in 2012 to 51,510 articles in 2020. Comparatively, the number of articles published in Elsevier journals increased from 9,401 articles in 2012 to 11,651 articles in 2016, and subsequently decreased to 10,623 articles in 2020. In terms of Elsevier’s market share of DEAL articles (Figure 2B), the years 2013-2015 show a trend of relatively small year-on-year (YOY) market share gains (<0.5% per year), with Elsevier’s market share reaching a peak of 25.3% in 2015. Subsequent years were characterised by a trend of larger market share losses, resulting in a final market share of 20.6% in

<sup>1</sup>As an example, if DEAL researchers published 1000 articles in Year One, of which 200 were in Elsevier journals, and 1500 articles in Year Two, of which 240 were in Elsevier journals, then the change in proportion from Year One to Year Two is calculated as  $(240/1500) - (200/1000)$ , equal to -0.04 and interpreted as a market share loss of 4%.

2020. Whilst the general trend appears to occur independently of the timing of Elsevier access restrictions in mid-2018, the largest YOY loss occurred in 2020 ( $>1.5\%$  loss of market share), covering the period 18-30 months following the access restrictions. It is important to note that these results reflect the numbers and proportions of articles *published* in journals, but do not necessarily reflect article *submission* dynamics: articles take many months to proceed through peer-review and publication processes (and these processes are generally faster in STM fields versus social sciences/arts/humanities/economics fields; Björk and Solomon, 2013), and acceptance/rejection rates may not have remained static or proportional over time.

Figure 2C shows changes in the proportion of articles produced by each individual DEAL institution published in Elsevier journals. Patterns broadly reflect those shown in Figures 2A and 2B, with the proportion of Elsevier articles remaining relatively static from 2012-2016 and subsequently declining. Interestingly, a small number of DEAL institutions appear to publish 100% of their articles in Elsevier journals - however, upon inspection we find that these are all institutions with extremely low publication volumes ( $<10$  articles in a given year). Figure 2D shows the total number of Elsevier versus non-Elsevier articles published by DEAL institutions, aggregated over the entire time period of 2012-2020. [To do: expand on this section.]

**Did publishing patterns differ at institutions whose contracts with Elsevier expired in different years?** Our dataset of DEAL institutions consists of three groups of institutions whose contracts with Elsevier expired at different time points: one group at the end of 2016, one at the end of 2017 and one at the end of 2018. In Figure 3 we explore different in publishing patterns in these different groups, with the hypothesis that those whose contracts expired earlier will show different effects to those whose contracts expired more recently. Our results, however, do not clearly support this hypothesis: all three groups show similar dynamics with an overall gain in Elsevier’s market share between 2012 and 2015, and an overall loss between 2016 and 2020, although the magnitude and exact pattern differs between each group. A reason for the relatively homogenous behaviour may be that although the contracts expired at different timepoints, the time at which access to Elsevier was restricted was relatively similar across all institutions; those whose contracts expired at the end of 2016 and 2017 lost access in July-2018 (not including a brief 6-week period at the beginning of 2017), whilst those whose contracts expired at the end of 2018 lost access at that same time.

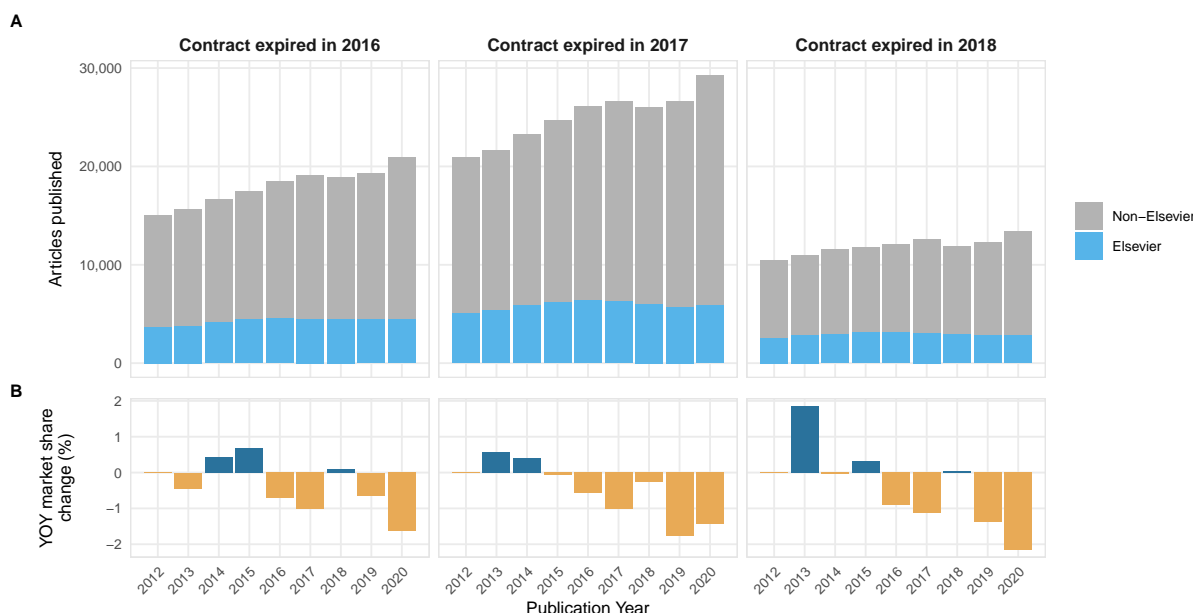


Figure 3: Changes in publishing behaviour of DEAL researchers, 2012-2020, dependent on year of contract expiration with Elsevier. (A) Total number of articles published by DEAL researchers in Elsevier and non-Elsevier journals. (B) Year-over-year (YOY) change in Elsevier’s market share of articles published by DEAL researchers.



**How did publishing behaviour vary with respect to research disciplines?** A complicating factor in our dataset is that we have included articles covering multiple research disciplines. A recent analysis of the effect of the PAR agreements made between DEAL and Springer Nature, and DEAL and Wiley (Haucap et al., 2021), chose to focus on a single discipline, Chemistry, with the justification that:

*“Manuscript turnaround times differ substantially between different fields of science and are rather long in some disciplines such as economics (see, e.g., Ellison, 2002). Hence, the vast majority of articles published in economics journals in 2019 and 2020 will have been submitted before the DEAL agreements were announced. Therefore, our analysis focuses on the field of chemistry which has much faster turnaround times so that we can expect the DEAL agreements to already have at least some impact.”*

Our analysis covers a longer time period than that of Haucap et al. (2021), and we assess effects covering the entire time period in which negotiations with Elsevier began in 2016, the time at which access was restricted in mid-2018, and 30-months thereafter until the end of 2020. We therefore feel justified in including a broader range of disciplines. Nonetheless, we have also analysed changes at the level of individual disciplines (i.e. Dimensions Fields of Research) (Figure 4). For visualisation purposes we limited results here to the top-10 disciplines by publishing volume, which tend to focus primarily on STM disciplines; full results for all 22 disciplines can be found in the supplementary material [To do: Add supplementary figure].

Overall, the patterns display a higher degree of fluctuation than our analysis of all disciplines together (likely due to smaller sample sizes), and a general tendency towards a reduction in market share for Elsevier over time. However, no discipline-specific patterns that can be clearly attributed to the restriction of access to Elsevier journals are evident.

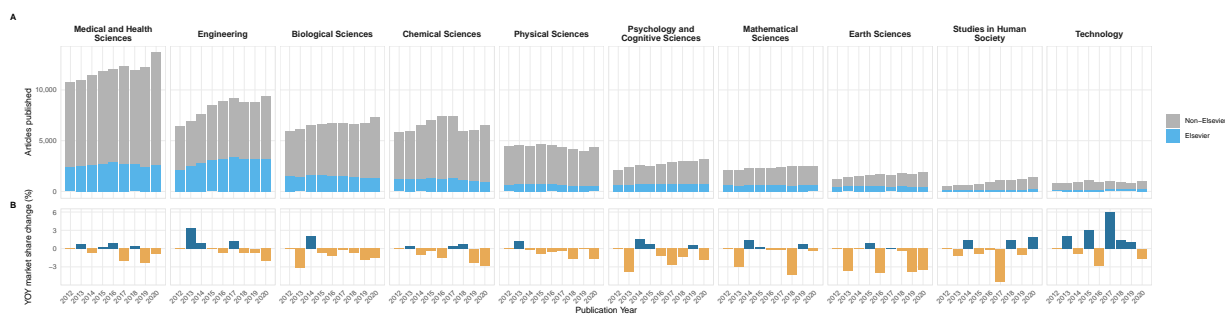


Figure 4: Changes in publishing behaviour of DEAL researchers, 2012-2020, dependent on research discipline (top-10 by publishing volume are shown). (A) Total number of articles published by DEAL researchers in Elsevier and non-Elsevier journals. (B) Year-over-year (YOY) change in Elsevier’s market share of articles published by DEAL researchers.

**How did publishing behaviour vary with respect to collaboration patterns?** Another potential confounder of our basic results in earlier sections relates to that of collaboration behaviour. An article written solely by researchers at DEAL institutions may have a different publication strategy compared to one written in an internationally-collaborative project, where international colleagues may be less knowledgeable of Elsevier access restrictions, or less disturbed in their daily research activities. Figure 5 shows changes in publishing behaviour with respect to collaboration status. We classify articles into 3 distinct collaboration classes: (1) “DEAL collaboration”, referring to articles where all authors of the article are based exclusively at DEAL institutions, (2) “National collaboration”, referring to articles where some authors are based at DEAL institutions and others at non-DEAL institutions within Germany, and (3) “International collaboration”, referring to articles where some authors are based at DEAL institutions and others at institutions outside of Germany. Here we observe some interesting patterns: with respect to DEAL-only collaborations, the total number of articles published has grown relatively slowly between 2012 and 2020, yet the number of articles published in Elsevier journals fell from 6,673 articles in 2016 to 5,384 articles in 2020. By comparison, the total number of national and international collaborative articles has proportionally grown more rapidly

between 2012 and 2020, but the number of articles published in Elsevier journals has remained relatively constant. In terms of market share... [To do: expand this section]

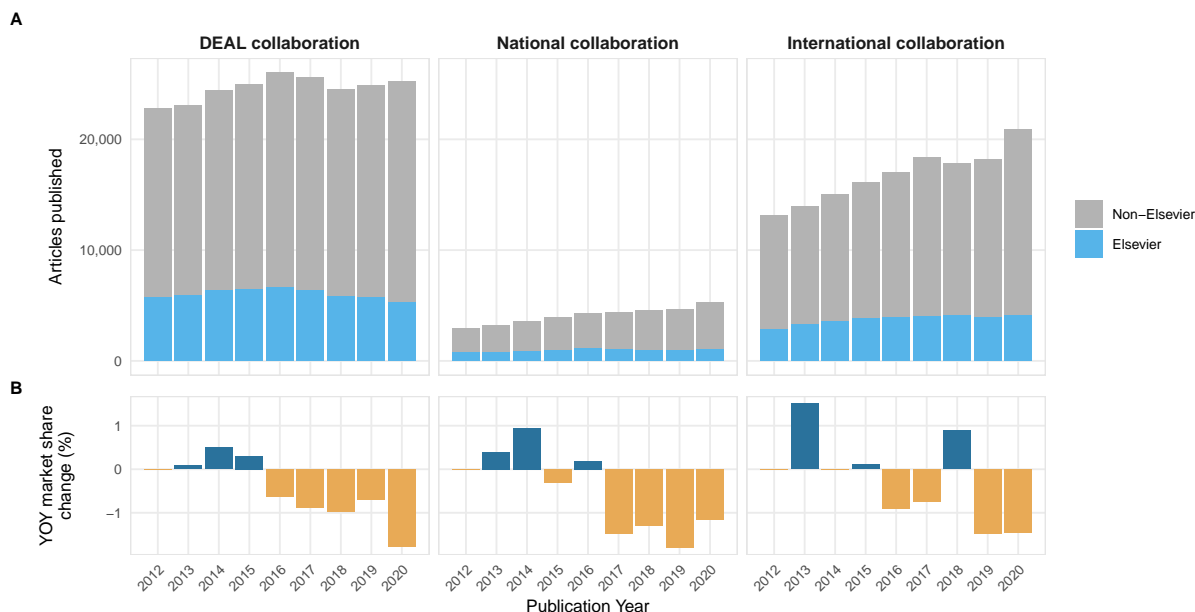


Figure 5: Changes in publishing behaviour of DEAL researchers, 2012-2020, dependent on collaboration status. 'DEAL collaboration' refers to articles where all authors of the article are based exclusively at DEAL institutions, 'National collaboration' refers to articles where some authors are based at DEAL institutions and others at non-DEAL institutions within Germany, and 'International collaboration' refers to articles where some authors are based at DEAL institutions and others at institutions outside of Germany. (A) Total number of articles published by DEAL researchers in Elsevier and non-Elsevier journals. (B) Year-over-year (YOY) change in Elsevier's market share of articles published by DEAL researchers.

**How did publishing behaviour vary with respect to open access status?** Between 2010 and 2018 the proportion of articles authored by German researchers that were made OA increased dramatically, from 27% in 2010 to 52% in 2018 (Hobert et al., 2020). We also aimed to determine whether the restriction of access to Elsevier journals had an effect on OA publishing behaviour of DEAL researchers in Elsevier journals, with the hypothesis that they would be more likely to publish their articles under OA licenses to ensure that their colleagues are able to access them. Articles were classified into OA categories following the same schema used by Unpaywall (more information on the classification schema is available [here](#)). An important point for the analysis of OA shares is that our dataset measures OA availability at the time of measurement (in our case, February 2021), and so OA shares do not necessarily reflect the OA status of an article at the time of its publication.

Results in Figure 6 broadly agree with the findings of Hobert et al. (2020), showing an overall increase in the volume of OA articles published by DEAL researchers between 2012 and 2020, and a decrease in the volume of closed articles. In particular, we observe dramatic growth of hybrid OA from 2018 (3,067 articles) to 2020 (14,376 articles), which may be driven, at least in part, by new publishing agreements made with publishers including Springer Nature and Wiley (Haucap et al., 2021).

With respect to OA publishing behaviour of DEAL researchers in Elsevier journals, we observe a decrease in the number of closed-access articles published from 2015 (7,807 articles) to 2020 (6,712 articles), and increases in the number of Gold (1,193 articles in 2015 to 1,497 articles in 2020) and Hybrid OA (609 articles in 2015 to 1,188 articles in 2020) publications over the same time period. In terms of market share, the most prominent feature is that of YOY losses for Elsevier in the Hybrid OA market share of >5% in 2019 and 2020; these losses, however, appear to be driven more by the surge of Hybrid OA publishing in other venues



rather than any reduction in the volume of Hybrid OA published by Elsevier.

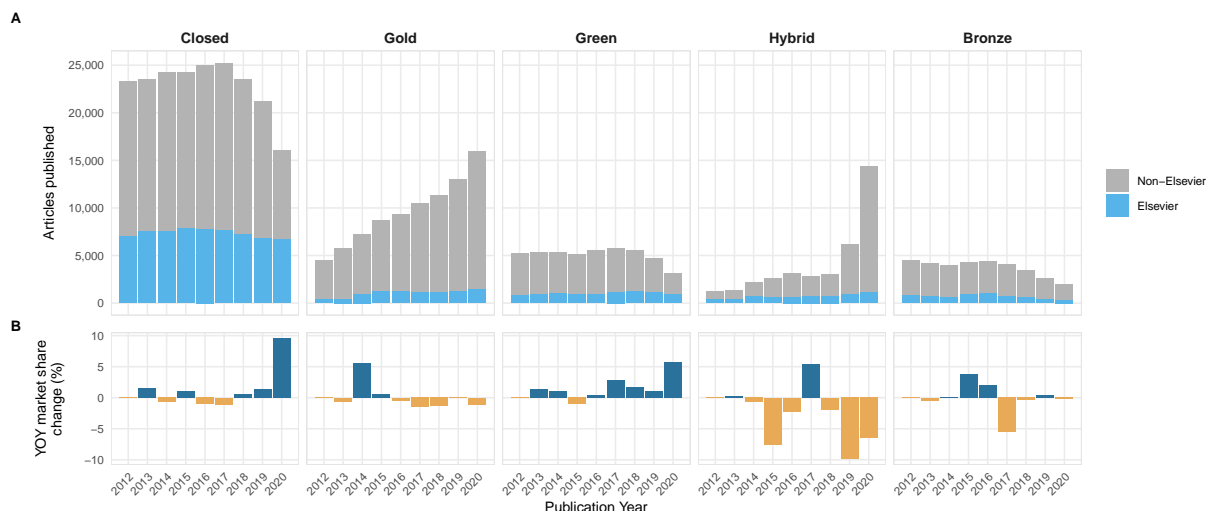


Figure 6: Changes in publishing behaviour of DEAL researchers, 2012-2020, dependent on OA status. (A) Total number of articles published by DEAL researchers in Elsevier and non-Elsevier journals. (B) Year-over-year (YOY) change in Elsevier's market share of articles published by DEAL researchers.

### Citing behaviour of DEAL researchers

A number of studies have found a citation advantage of open-access publications over their closed-access counterparts (c.f. [Piwowar et al., 2018](#)), with the implication that having access to articles makes them easier to read, download and ultimately more likely to be cited. If this were true, we would expect that restricting access to a set of articles would have the opposite effect, i.e. reduce their ability to be cited. We therefore aimed to investigate the effect of Elsevier access restrictions on DEAL researchers citing behaviour, using a set of 16,919,143 references (DOI-DOI links) from articles authored by DEAL researchers (limited to articles where the first and last author was at a DEAL institution, and articles that contained more than a single reference).

A complicating factor in the analysis of citing behaviour is that there exists time variation in both the year of publication of an article, and in the year of the articles they cite, i.e. articles may cite other articles published in any prior year. An overview of citing dynamics for our dataset of DEAL articles is displayed in Figure 7, where we show the mean number of references per article made to Elsevier or non-Elsevier articles, as a function of citation year (citation year refers to the difference in years between the publication date of the citing article and the publication date of the cited article). Although we look at the problem in an inverse way to the majority of citation studies (i.e. we analyse outgoing rather than incoming citations), our results display similarly typical dynamics (see, e.g. [Parolo et al., 2015](#)): articles cite relatively few articles published in the same year (presumably, as citing articles must first be written and proceed through the lengthy peer-review process), the number of cited articles peaks in citation years 2-3, and subsequently slowly declines over the following years.

For the purposes of the following analyses, we make the assumption that citing behaviour in response to Elsevier access restrictions is most likely to change for recent citations, which we define as citations with a citation age less than or equal to 2 years. This assumption is based on the fact that access restrictions affected new journal issues at all DEAL institutions, whilst only a subset of DEAL institutions also lost access to their back-catalogue of articles ([Vogel, 2017a](#)). We may therefore reasonably expect that authors who have had access restricted to Elsevier journals post-2018 either still have access to the older back-catalogues, or may have saved older articles to local storage media (e.g. in reference management software) during the time when access was still available. For assessing changes in citing behaviour, we define three related metrics of measurement in similar way to our analysis of publishing behaviour: (1) the proportion of articles published

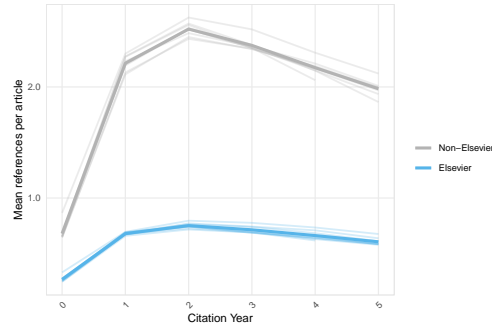


Figure 7: Referencing dynamics for articles published by DEAL researchers. Lines show the evolution of the mean number of references made to Elsevier or non-Elsevier articles, as a function of citation year. Thick line represents the overall mean, thin lines represent articles from individual publication years (labelled).

by DEAL researchers that cited *any* article in an Elsevier journal, (2) the total number of references made to articles in Elsevier versus non-Elsevier journals from articles published by DEAL researchers, and (3) the annual change in the absolute proportion of references made to articles in Elsevier journals from articles published by DEAL researchers (which can be thought of as equivalent to Elsevier’s “market share” of references).

Overall results for articles with publication years between 2012 and 2020 are shown in Figure 8, as well as citing behaviour for individual institutions. All results are based on citations with a maximum citation age of two years. 8A shows the proportion of articles published in a given year, that cite at least a single Elsevier article. Overall, the proportion of articles citing an Elsevier article has remained relatively constant, with values fluctuating between approximately 57% and 58%. 8B shows the total number of references made to Elsevier versus non-Elsevier articles in the same time period. The total number of references has grown from 276,127 in 2012 to 392,128 in 2020, although we note a plateau on the order of ~330,000 references between 2016 and 2019. [To do: Expand this section]

**Did citing patterns differ at institutions whose contracts with Elsevier expired in different years?** See Figure 9

**How did publishing behaviour vary with respect to research disciplines?** See Figure 10

**How did citing behaviour vary with respect to collaboration patterns?** See Figure 11

**How did citing behaviour vary with respect to open access status?** See Figure See Figure 12. Note: OA status here refers to the OA status of the **cited** article, not the **citing** article. Still need to think about this...

## Discussion and Conclusions

- General summary of results
- Compare results with those expected from surveys
- How have researchers mitigated the effect of Elsevier restrictions?
  - SciHub - proportion of downloads for Elsevier articles from Germany increased rapidly at end of 2017 shortly before access was cut off (Figure 13; data from [Strecker, 2018](#))
  - ILLs - would be future work to survey/interview librarians about effects - some information available in Bibsam surveys - no increase in ILL requests after 9 months. See also <https://arxiv.org/ftp/arxiv/papers/2009/2009.04287.pdf>

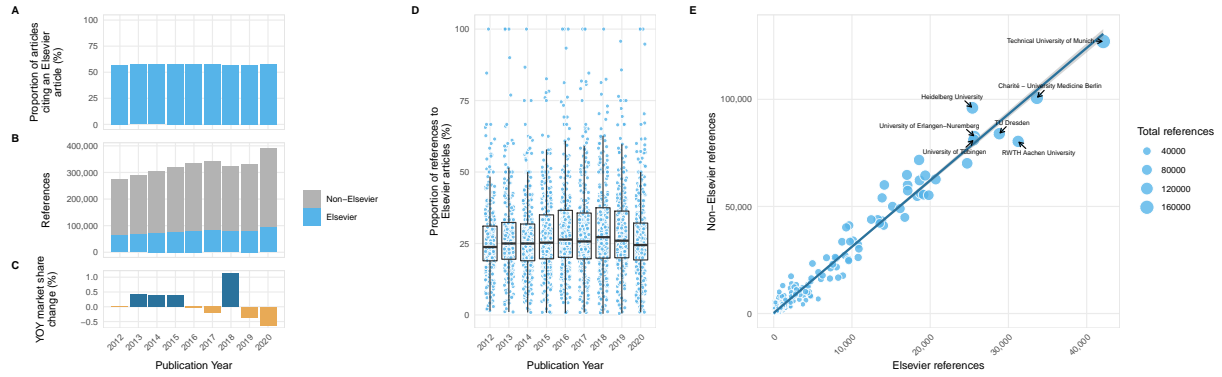


Figure 8: Citing behaviour of DEAL researchers, 2012-2020. All data is based on references from articles published by DEAL researchers to articles with a citation age of 2 years or less. (A) Proportion of articles published by DEAL researchers that cite at least one single Elsevier article. (B) Total number of references in articles published DEAL researchers to articles published in Elsevier versus non-Elsevier journals. (C) Year-on-year (YOY) change in Elsevier’s market share of references in articles published by DEAL researchers. (D) Proportion of references in articles published by individual DEAL institutions to articles in Elsevier journals. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to  $1.5 \times \text{IQR}$ . Points denote individual institutions, with added horizontal jitter for visibility. (E) Number of references to articles in Elsevier versus Non-Elsevier journals for articles published by individual DEAL institutions (totals aggregated for years 2012-2020). Point size is scaled by the total number of references from an institution. Institutions referencing  $>100,000$  articles are labelled.

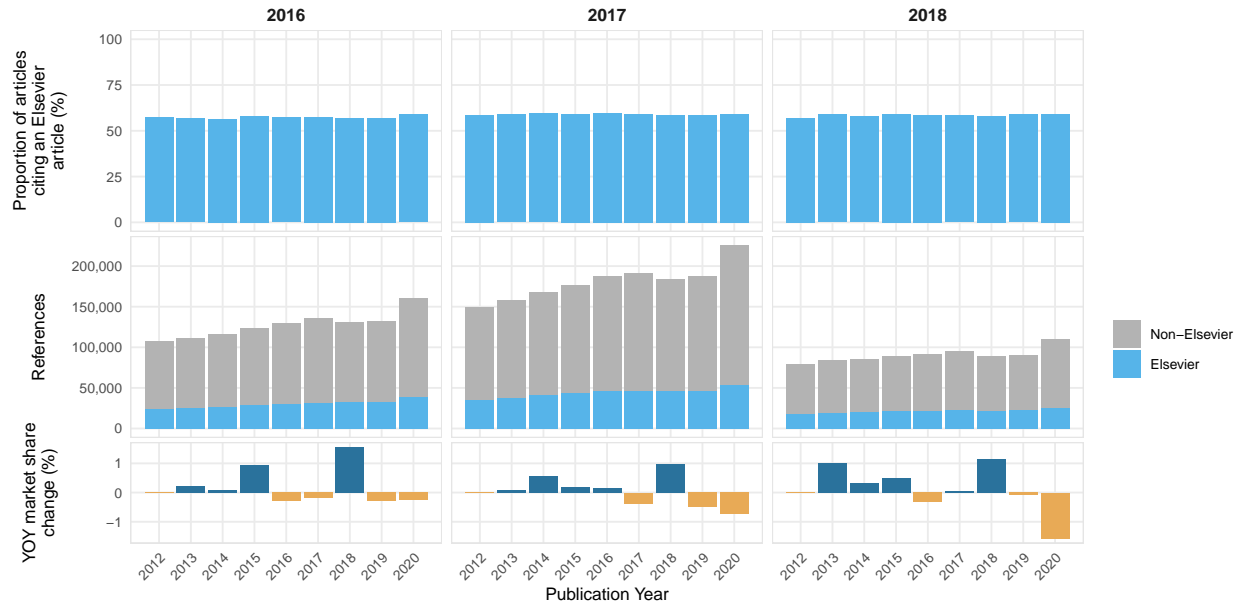


Figure 9: Citing behaviour of DEAL researchers, 2012-2020, dependent on year of contract expiration with Elsevier. All data is based on references from articles published by DEAL researchers to articles with a citation age of 2 years or less. (A) Proportion of articles published by DEAL researchers that cite at least one single Elsevier article. (B) Total number of references in articles published DEAL researchers to articles published in Elsevier versus non-Elsevier journals. (C) Year-on-year (YOY) change in Elsevier’s market share of references in articles published by DEAL researchers.

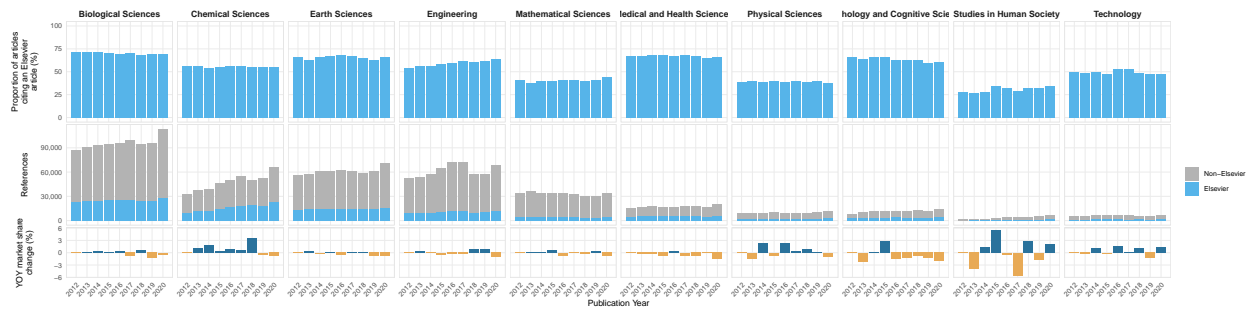


Figure 10: Citing behaviour of DEAL researchers, 2012-2020, dependent on research discipline (top-10 disciplines by publication volume are shown). All data is based on references from articles published by DEAL researchers to articles with a citation age of 2 years or less. (A) Proportion of articles published by DEAL researchers that cite at least one single Elsevier article. (B) Total number of references in articles published DEAL researchers to articles published in Elsevier versus non-Elsevier journals. (C) Year-on-year (YOY) change in Elsevier's market share of references in articles published by DEAL researchers.

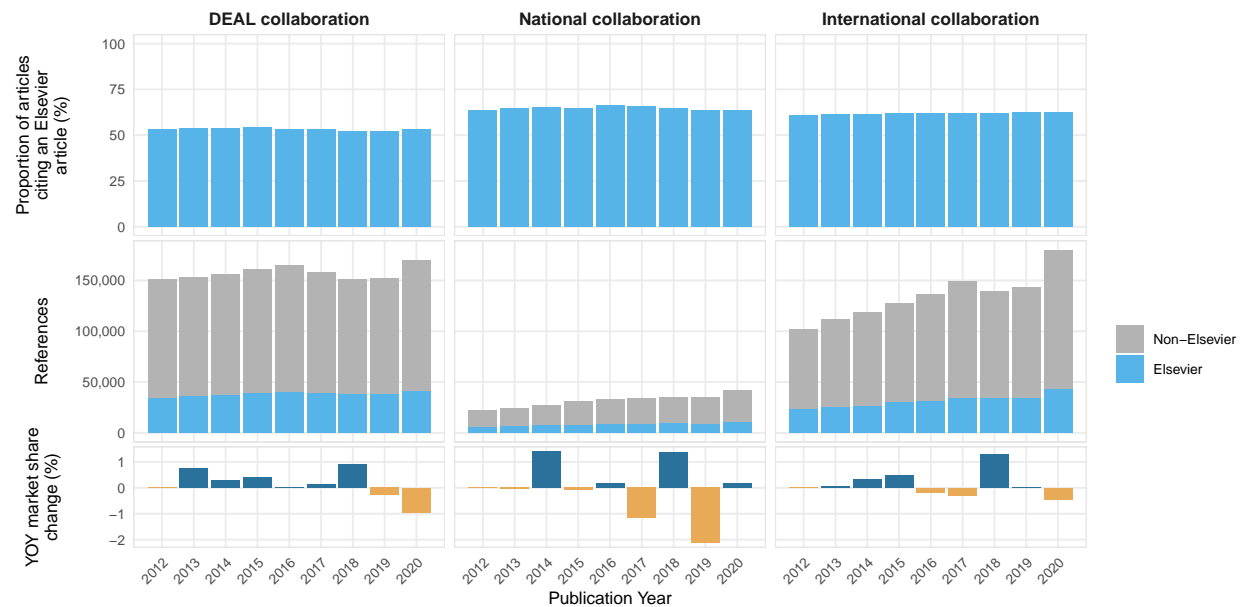


Figure 11: Citing behaviour of DEAL researchers, 2012-2020, dependent on collaboration status. 'DEAL collaboration' refers to articles where all authors of the article are based exclusively at DEAL institutions, 'National collaboration' refers to articles where some authors are based at DEAL institutions and others at non-DEAL institutions within Germany, and 'International collaboration' refers to articles where some authors are based at DEAL institutions and others at institutions outside of Germany. All data is based on references from articles published by DEAL researchers to articles with a citation age of 2 years or less. (A) Proportion of articles published by DEAL researchers that cite at least one single Elsevier article. (B) Total number of references in articles published DEAL researchers to articles published in Elsevier versus non-Elsevier journals. (C) Year-on-year (YOY) change in Elsevier's market share of references in articles published by DEAL researchers.

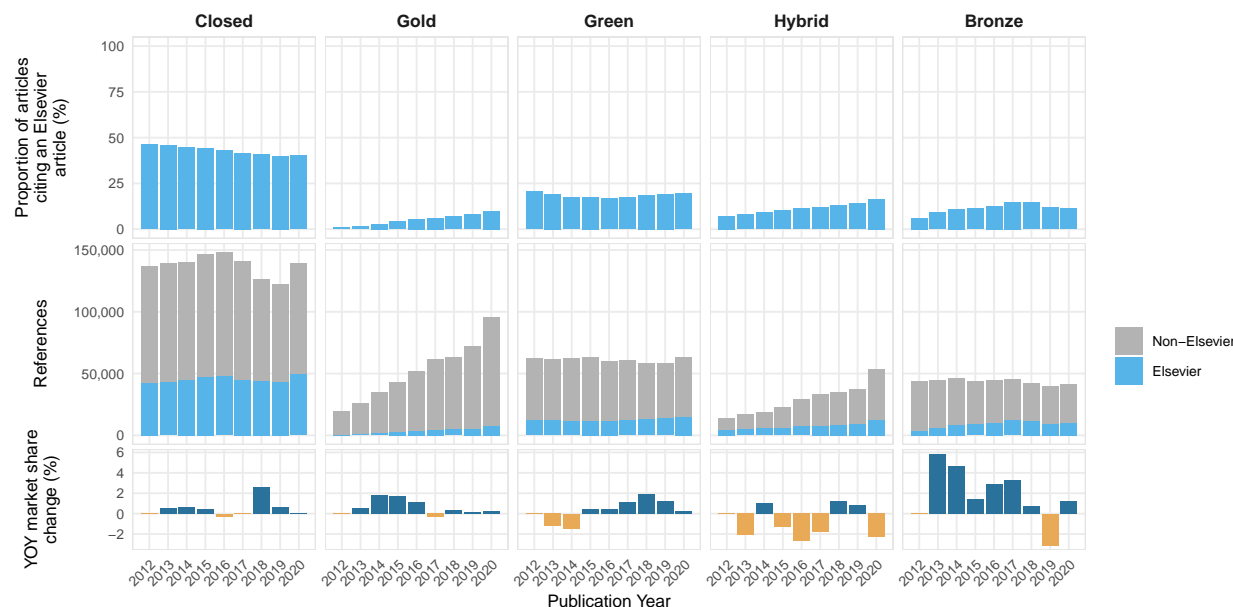


Figure 12: Citing behaviour of DEAL researchers, 2012-2020, dependent on OA status of the cited article. All data is based on references from articles published by DEAL researchers to articles with a citation age of 2 years or less. (A) Proportion of articles published by DEAL researchers that cite at least one single Elsevier article. (B) Total number of references in articles published DEAL researchers to articles published in Elsevier versus non-Elsevier journals. (C) Year-on-year (YOY) change in Elsevier’s market share of references in articles published by DEAL researchers.

- Co-authors
- Other “dark” methods #icanhazpdf
- Some citations made to papers that academics do not read in full
- Recommendations for future negotiations

**Limitations** Our study has a number of limitations, which may be built upon in future studies:

- Data sources
- Publication timelines
- Short time period

### Future directions

- Verify results with surveys and knowledge from librarians
- Measure changes over a longer time period

### Other things to check

- Consistency of wording - e.g. “references” versus “citations”.

### References

- Auto-generate from .bib file

### Acknowledgments

[To do: Add acknowledgements (OASE/OAUNI funding, Dimensions access)]

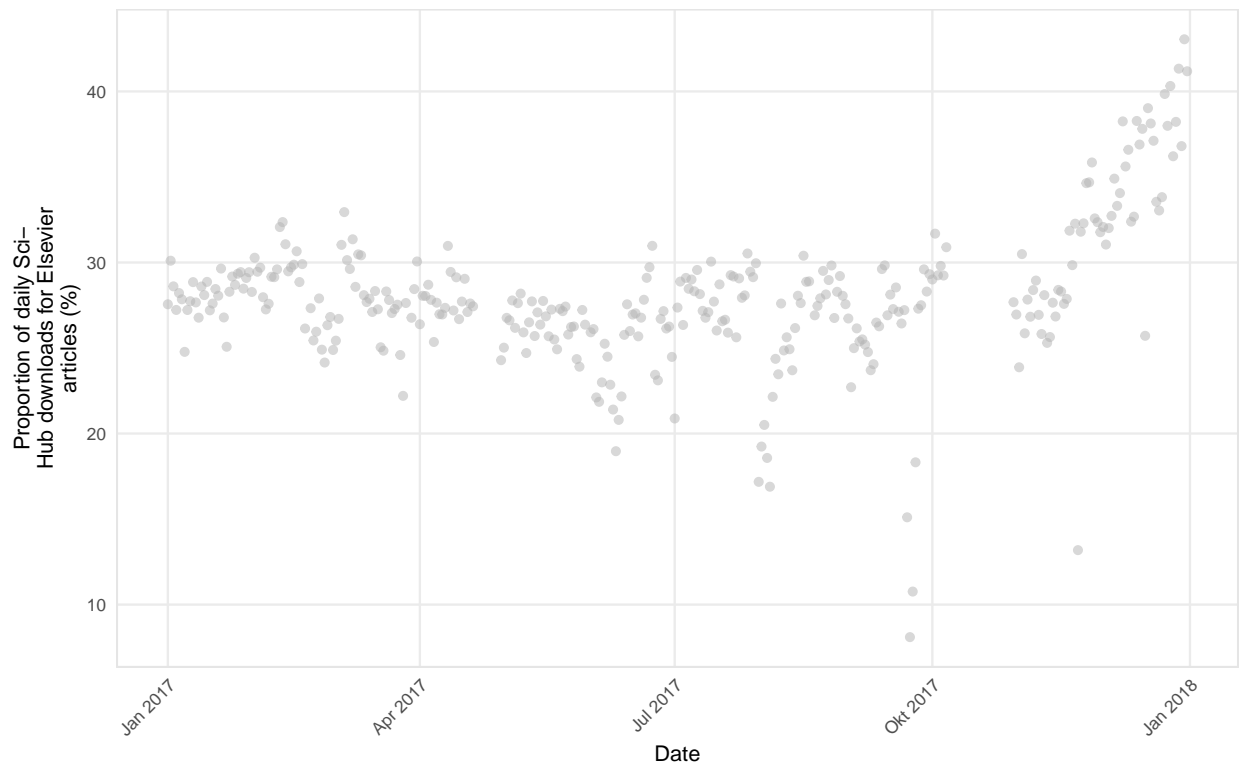


Figure 13: Proportion of daily Sci-Hub downloads from Germany in 2017 for Elsevier articles. Data from [Strecker, \(2018\)](#).