Cover Effects on Citations Uncovered: Evidence from *Nature*[*]

Pietro Battiston[a], Pier Luigi Sacco[b], and Luca Stanca[c]

[a]Department of Economics and Management, University of Parma, Via J. Kennedy 6, 43125 Parma, Italy. E-mail: me@pietrobattiston.it
[b]DiSFiPEQ, University of Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy; metaLAB (at) Harvard, 42 Kirkland St, 02138 Cambridge MA USA. E-mail: pierluigi.sacco@unich.it; pierluigi_sacco@fas.harvard.edu.
[c]Department of Economics and NeuroMI, University of Milano Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy. E-mail: luca.stanca@unimib.it

April 2022
Declarations of interest: none

Abstract

We study the effects of an article featured on the cover of the journal *Nature* on citations to all articles published by its authors. Based on 30 years of bibliometric data, we find that cover articles are cited significantly more than non-cover articles, with this difference being long-lasting. However, when considering all articles (past and future) by *Nature* authors, we find that the publication of a cover article causes citations to previous articles by its authors to decline relative to citations to articles by non-cover authors.
Keywords: Bibliometric indicators, Citation flows, Research evaluation, Cover article

[*]We thank four anonymous reviewers for their comments. This research benefits from the HPC (High Performance Computing) facility of the University of Parma, Italy.

1 Introduction

In the last two decades, the availability of large citation databases has had a profound impact on how research is evaluated, financed and even discussed. Impact is the most often mentioned feature of a research product, mainly because it is easily measurable, and impact metrics can be a key driver for researchers to decide to which journal to submit articles. Journals therefore maintain a strong focus on keeping their impact metrics high, especially so relative to competing journals in the same field. Such a strong emphasis on impact reflects the difficulty of assessing the quality of research, as very large numbers of articles are published daily in increasingly differentiated and specialized fields. Quality is an elusive notion that is difficult to define or measure.

Indicators of impact are commonly used in the scientific profession for several purposes, such as tenure decisions (Heckman and Moktan, 2018), the assessment of research performance (Bordons et al., 2002; Demetrescu et al., 2019), the evaluation of scientific sources (Brown, 2007), and even the choice of issues to investigate (Monastersky, 2005). Although citations do not necessarily reflect a positive assessment of the article's quality, impact has become a commonly used proxy for quality. Despite this practice being highly controversial (PLoS Medicine Editors, 2006; Simons, 2008; Archambault and Larivière, 2009; Callaway, 2016), the lack of viable alternatives favors its adoption and diffusion. There is widespread awareness that impact itself is a multidimensional notion that is not easily captured by a single indicator, and that different measures may better serve different purposes (Bollen et al., 2009). However, citation-based indices remain a core focus throughout the scientific profession.

Despite the abundance of citation data, relatively little progress has been made in understanding the determinants of citation flows. The relevant literature is fragmentary and not commensurate in volume to the practical relevance of the issue (e.g., Price, 1976, Johnson, 1997, Ayres and Vars, 2000, Mingers, 2008, Di Vaio et al., 2012). Since other aspects of research (such as originality and methodological rigour) are comparatively difficult to measure, understanding whether impact reflects some *intrinsic* quality of research would also provide insights about the extent to which citations can be considered an appropriate measure of quality of research.

The present study belongs to the literature focusing on specific non–intrinsic factors that can affect citation flows, including reputational assets such as scientific prizes and awards, the recognition of a journal by a highly reputable scientific society, or the prestige of the academic affiliation of the author. Among the relatively few authors to identify a causal relation between non-intrinsic characteristics of research and citation success, Frandsen and Nicolaisen (2013) show that citations to *past* articles of Nobel Prizes laureates increase after the award, and Azoulay et al. (2014) describe a similar effect on citations to articles by medical researchers who are appointed as *Howard Hughes Medical Investigator*. On the other hand, awards may make scientists more self-critical about their own scientific production. Borjas and Doran (2015) find a *decrease* in productivity for mathematicians who are awarded the Fields medal. It cannot be ruled out that awards even play a demotivating role on subsequent productivity and, as a consequence, impact (Robinson et al., 2021). Publishing in journals promoted by reputable scientific societies may instead lead to a retrospective boosting of a scholar's scientific reputation, increasing visibility and relevance among peers. Battiston (2014) shows that citations to articles in the journals of the American Physical Society rise when an author publishes further articles in such journals. Medoff (2006) shows that, in the economics field, authors' affiliation to a small number of elite institutions ensures wider recognition of their articles. In a closely related work, Feenberg et al. (2017) show that research papers featured at the top of a popular newsletter tend to be relatively more downloaded and cited.

In this article we examine another type of prestige-related influence on citation flows: being featured on the *cover* of a high-impact scientific journal. The rationale for such an analysis comes from the fact that journal covers represent an explicit message of relevance to the scientific community, similarly to a prize. At the same time, they are "awarded" on a regular basis, as part of an activity — publication on research outlets — that is an integral part of any researcher's professional life. Therefore, cover articles can be studied systematically over time for larger numbers of authors than actual prizes or awards.

Our analysis focuses on the journal *Nature*, given its prestige in the scientific community and the existence of previous related studies that examine citations to articles published in this journal. Wang et al. (2017) show, based on total citation counts from the *Web of Science* database, that articles featured on the cover of *Nature* between 2008 and 2010 obtain on average more citations than articles published on *Nature* in the same period but not featured on the cover. A similar result is obtained by Kong and Wang (2020), who corroborate it by also collecting Altmetrics scores of cover and non-cover articles: the former obtain higher scores.[1] However, a causal interpretation of these findings is far from obvious. Being featured on the cover of *Nature* can increase the scientific salience and hence the citability of an article, but it is also possible that a more relevant, or highly citable, article has a higher probability of being selected for the cover. At the same time, an article can be featured on the journal cover not only for its scientific value, but also for factors such as the appeal of the topic, its current relevance for mainstream media, or simply the aesthetic quality of the associated images. These differences may apply both at the individual level of authors or articles, and at the discipline level, with scientific fields differing in their ability to attract citations and/or to produce articles considered fit for a cover. The citability of cover articles is therefore a complex topic that deserves deeper investigation, also in view of the scarce attention received so far.

Our study aims at shedding light on this issue by examining the dynamics of "cover effects" over a long time span and, most importantly, by taking into account the entire citation life of the authors of articles published in *Nature*. Our research question can be summarized as follows: are authors of articles featured on the *Nature* cover more cited because of the resulting exposure, or because they were *ex–ante* different and were therefore selected for the cover? We answer this question by implementing an event study. The paper is structured as follows. Sections 2 and 3 describe the data set and methods, respectively. Section 4 presents the results. Section 5 concludes.

2 Data

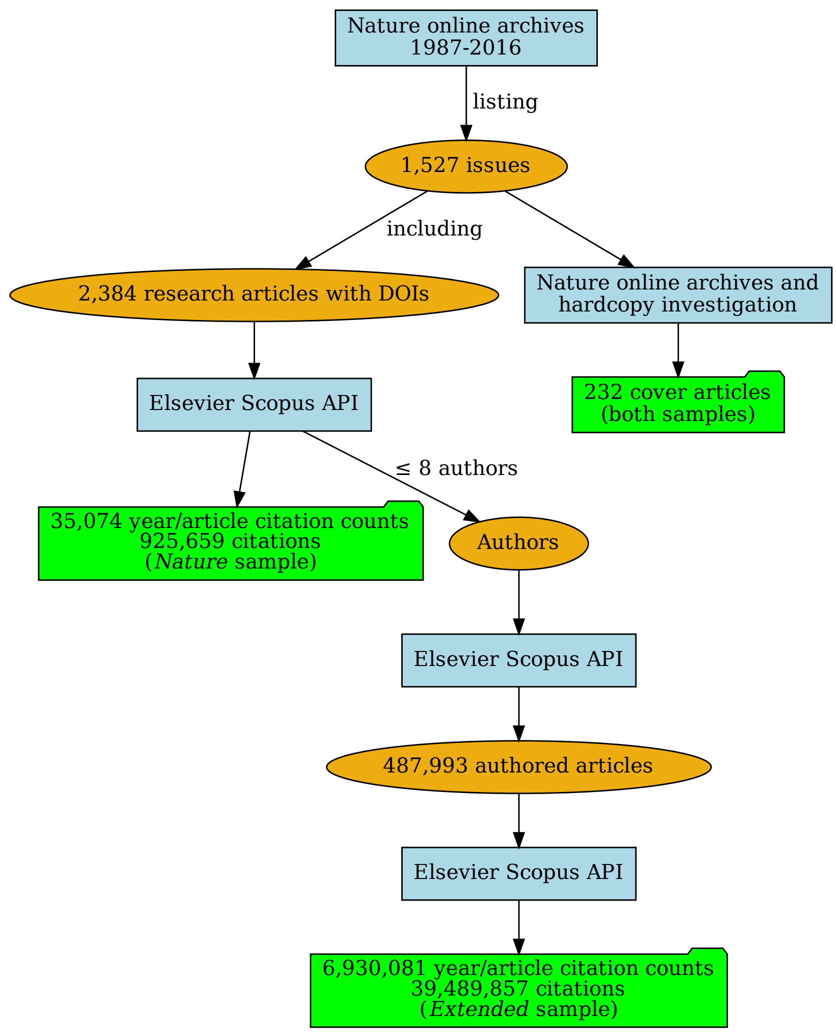
Our data set was constructed in two steps, as illustrated in Figure 1. First, we considered all research articles published in *Nature* between 1987 and 2016, with detailed information about the content of the cover of each issue,[2] by combining online *Nature* archives, bibliometric information retrieved from the *Scopus* database,[3] and the manual investigation of hardcopies. Since only part of each issue of *Nature* contains original research articles, with other

[1] Wang et al. (2015) conduct a similar analysis on *PLOS Biology* and only find a significant difference for two of the five years considered.

[2] Colored cover images were present during the entire period of time analyzed in the present study.

[3] Data were obtained via the Scopus API (`https://dev.elsevier.com`), by employing each article's doi as obtained from *Nature* archives. A small number of articles with

Figure 1: Data collection flow chart



Note: illustration of the data collection procedure. Rectangles represent data sources, ovals represent intermediate data aggregates, folders represent data ultimately employed in the econometric analysis.

sections being devoted to scientific news or commentaries, we restricted the analysis to documents appearing in the "Articles" section.[4] The resulting dataset (*Nature* sample) covers 1,527 weekly issues of the journal, that include 2,443 research articles. For each article, we observe citation flows over a period of up to 30 years (from the year of publication to 2017), 14.49 years on average. Hence, we obtain a total of 35,074 observations, as described in column 1 of Table 1, each counting the number of citations received by a given article in a given year. We excluded 6 articles for which information on authors was not available.[5]

Second, we collected from the *Scopus* database citations to all articles published by authors of the *Nature* articles mentioned above. We considered articles published in the same time range (1987-2016), and all their citations received in this time span. This allowed us to analyze the entire citation records of the authors. Indeed, this extended sample allows us to construct a measure of citation dynamics that we can employ both *before* and *after* an author publishes on *Nature*. This is central for the implementation of our event study: if we limited our analysis to the *Nature* publications, we would be unable to control for ex–ante differences between cover and non–cover articles.

We restricted this search to *Nature* articles with a limited number of authors (8 or less). This restriction was implemented for both theoretical and practical reasons. From a theoretical point of view, since we are looking for effects of a given (*Nature*) publication on a researcher's citations, it is relatively less likely that publications with a very large number of authors have a significant effect for each of them. Authors' publishing and coauthorship patterns are known to be heterogeneous across disciplines (Abramo et al., 2013), and in disciplines with large collaborations, typically producing large numbers of scientific articles, each scientific article is likely to have a relatively smaller impact on individual authors. From a practical point of view, limiting the maximum number of authors ensures that sample size remains tractable: given that, as already mentioned, large collaborations typically

duplicated or missing doi were manually matched on the basis of title and authors.

[4]While the name of this section underwent slight changes over the years, e.g., "Research Articles", it can be clearly identified during the entire period of interest.

[5]The DOIs of the six articles are 10.1038/341619a0, 10.1038/342243a0, 10.1038/364203a0, 10.1038/nature02564, 10.1038/nature03154, 10.1038/nature03156. Note that these articles might have authors listed in the *Nature* website, but they lack any association to user profiles in the *Scopus* database.

produce large numbers of publications, including in the extended sample each article by each author would result in an over-representation of such collaborations. We have experimented with different thresholds for the maximum number of authors. The results of this sensitivity analysis, presented in Appendix B, confirm that the effect of interest is attenuated as the maximum number of authors increases, but also that it is robust and its dynamics are qualitatively similar across threshold values. The threshold of 8 authors results in an *extended* sample of 487,993 scientific articles, as described in column 3 of Table 1.

Table 1: Descriptive statistics

	Nature sample		Extended sample
	Overall	$\leq$ 8 authors	$\leq$ 8 authors
Articles	2,384	1,643	487,993
Article/year combinations	35,074	26,699	6,930,081
Total citations	925,659	540,699	39,489,857
Max citations	10,879	10,879	32,935
Max citations per year	1,192	1,192	3,551
Cover articles	229	138	43,600

Note: data sources are *Nature* archives and *Scopus*. The second column describes the restricted sample of *Nature* articles (with no more than 8 authors) on which the extended sample (third column) is based. In the third column, "Cover articles" are those whose corresponding *Nature* article was featured on the cover. Six *Nature* articles with no registered authors are excluded from the count.

It is worth observing that a given article in the extended sample can be related to more than one *Nature* article: when this is the case, it appears more than once in our database, and hence it has a larger weight in the analysis. We think this is the most appropriate way to take into account such multiplicity. Consider two researchers who co-authored an article on a given journal in year 2000, and then individually published one article each on *Nature* in 2003 and 2006, respectively. When studying citations to the 2000 article, in order to consider the effects of both the 2003 and 2006 *Nature* publications, we consider the 2000 article as two separate observations.

3 Methods

Our key hypothesis focuses on the *causal* effect on citations of being featured on the cover:

> H1: If a *Nature* article is featured on the cover of the journal, this has a positive effect on subsequent citations to articles by its authors.

This hypothesis is consistent with several underlying mechanisms, such as increased visibility, salience or prestige. Cover articles could capture the attention of more readers, who in the future may be more likely to read articles by the same authors and cite them. Cover articles could also be perceived as more relevant or salient within the literature, resulting in more citations to their authors. In addition, being featured on the cover may confer increased prestige and authority, leading other researchers to read and cite more the authors' works.

A second complementary hypothesis concerns the ex-ante specificity of cover authors (selection effect):

> H2: Authors of more cited articles are more likely to be featured on the cover of a *Nature* issue.

Note that this hypothesis does not imply that the selection of the cover paper for a *Nature* issue is based *only* on the citations received by its authors. Many other factors are likely to affect the selection, including the attractiveness of the topic to a general audience and its novelty. However, it is important to understand if cover authors are *ex ante* different in terms of citations received, since this would imply that the comparison in citations between cover– and non–cover articles, as made in the existing literature (Wang et al., 2017; Kong and Wang, 2020), is affected by a selection effect.

Our analysis starts with a characterization of the key features of our citation data. Namely, we consider the following empirical specification:

$$c_{i,y} = \alpha_0 + \alpha_1 F_i + \alpha_2 y + \alpha_3 p_i + \epsilon_{i,y} \tag{1}$$

where $c_{i,y}$ is the number of citations that article i receives y years after its publication, p_i is the year of publication, F_i is a dummy variable taking value 1 if the article published in *Nature* by the authors of article i was featured on the cover, and $\epsilon_{i,y}$ is an idiosyncratic error term at the article/year level.

While the comparison of citations to cover articles with those to non-cover articles is informative, it cannot *per se* provide evidence of a causal cover effect. Articles are not randomly selected to appear on the cover. On the contrary, it can be argued that the selection is related to the quality of the article, which in turn is positively related to citation flows, so that endogeneity may be present. Hence, a simple comparison of citation flows between cover and non-cover articles does not allow to disentangle "cover effects" from selection based on (unobserved) quality. We tackle this ambiguity by analyzing the dynamics of citations, exploiting the discontinuity provided by the publication of a *Nature* paper. To this aim, we turn to the extended sample, whereby each article is associated to a *Nature* article, the publication of which is the basis for an event study.

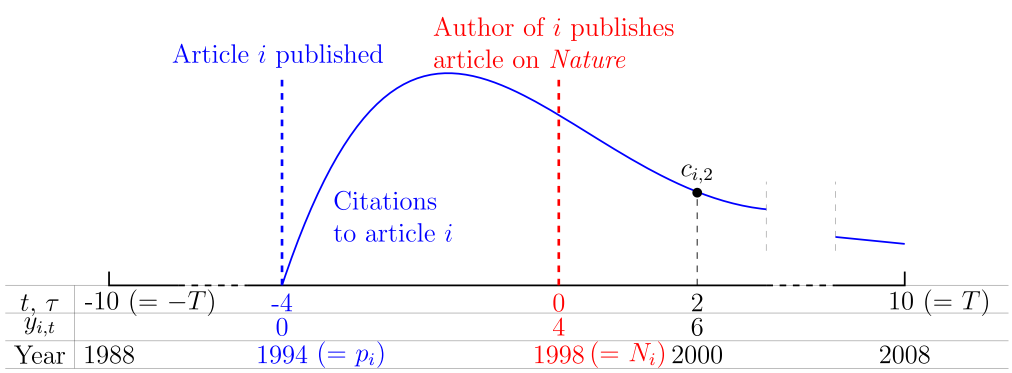
We consider citations received by articles over their life cycle and estimate the following empirical specification, obtained by augmenting Equation (1) with the interaction between the cover dummy and year–fixed effects relative to the publication of the *Nature* article:

$$c_{i,t} = \beta_0 + \sum_{\tau=-T}^{T} (\gamma_\tau + \beta_{1,\tau} F_i) Y_\tau + \beta_2 y_{i,t} + \beta_3 p_i + \epsilon_{i,t} \qquad (2)$$

where $c_{i,t}$ denotes the yearly flow of citations that article i receives t years before/after publication of the relevant *Nature* article (e.g., if the authors of article i published an article in *Nature* in 1998, $c_{i,2}$ is the number of citations to article i in 2000), Y_τ is a set of relative-time fixed effects, i.e., dummy variables equal to 1 when $t = \tau$. Year of publication (p_i) is defined in absolute terms as in Equation (1), with β_3 capturing overall time trends in citations. Notice that $y_{i,t}$ denotes years since publication of the article i itself (between 0 and 47, the length of the time span analyzed, and hence the maximum number of years over which a given article can be observed), while t and τ refer to years relative to publication of the relevant *Nature* article (from $-T$ to T), i.e., $y_{i,t} = N_i + t - p_i$, where N_i is the publication year of the relevant *Nature* article (see Figure 2 for a visual representation). In short, the offset between a reference system based on the publication of the original article and one based on the publication of the Nature article ($y_{i,t} - t$) is equal to the difference between the two publication years ($N_i - p_i$).

The coefficients γ_τ allow us to gauge the dynamics of citations to articles, as a function of years since the publication of a non–cover *Nature* article by one of their authors. The coefficients $\beta_{1,\tau}$ capture the interaction between the cover dummy variable and years since/to publication of the *Nature* article. These are the key coefficients of interest, as they measure the *additional* number of citations that an article obtains, on average, if the related *Nature* article, published τ years earlier, was featured on the cover (relative to a non-cover *Nature* article). We will be specifically interested in the estimates of $\beta_{1,\tau}$ for small positive values of τ (i.e., following the publication of the *Nature* article).

Figure 2: Timeline of event study



Note: timeline illustrating the notation of Equation (2) in the case of an article i published in 1994 by an author who then in 1998 published an article on Nature; T=10.

Equation 2 will be first estimated on the *Nature* sample in order to provide a comparison with previous results in the literature (Wang et al., 2017). This restricted data set, however, does not provide a feasible set up for a causal interpretation, given that, for *Nature* articles, we do not have citations prior to the publication date of the articles themselves. Notice that in this sample, even if an author happens to have published twice on *Nature*, the two articles are not considered as related in any way: each article's "related *Nature* article" is the article itself. In order to test hypotheses H1 and H2, we will then estimate Equation (2) in the extended sample. This allows us to observe citations flows before and after the publication of the relevant *Nature* article, thus providing an appropriate setup for an event study.[6] We

[6]One difference between our work and typical event studies (Fama et al., 1969) is that in our case research articles are at the same time the unit of analysis and the actual event:

thus exploit the discontinuity in the year of publication of the *Nature* article to obtain a causal interpretation of cover effects: the previously mentioned endogeneity issues are dealt with by focusing on changes *in the first few years* after the event. Such causal evidence cannot be obtained by looking only at articles published in *Nature*, as the choice of articles to be published on the cover might be non-random. The extended data set allows us to investigate cover effects on citations to the entire set of publications of a researcher. In addition, by analyzing "pre-*Nature*" citations flows, we are able to assess the hypothesis that cover authors are *ex ante* different from non-cover authors.

4 Results

Table 2 reports OLS coefficients for Equation (1) based on the *Nature* sample, with the specification in column (2) including an additional control for previous citations to other articles of authors. As expected, we find a positive and significant coefficient for the cover dummy variable: a *Nature* article featured on the journal's cover receives on average about 16 more citations per year than a non-cover article. Consistent with the literature (Althouse et al., 2009), more recent articles receive significantly more citations than less recent articles ($\hat{\alpha}_3$=0.817), while the number of citations to an article is negatively related to year of citation ($\hat{\alpha}_2$=-0.572), as in Aksnes (2003) and Mingers (2008). Coefficient estimates are virtually unchanged when controlling for previous citations to other articles by the same authors (column 2).

We start the presentation of cover dynamic effects with an exploratory analysis that helps to compare our findings with previous results in the literature. To this aim, Figure 3 displays the coefficients obtained by estimating Equation (2) only on the *Nature* sample (note that in this sample no observations are available for $\tau < 0$). The left panel shows that, as it is generally found in the literature, the number of citations to an article peaks 3 to 5 years after its publication and then gradually falls over time. The right panel shows that cover effects are positive and significant, reach a peak 2 to 3 years after publication, and are long-lasting: cover articles receive significantly more

this does not hinder our results, which will distinguish between the *Nature* sample – on which causal inference is not possible – and the extended sample. Our approach differs instead from the "difference-in-discontinuities" approach by Grembi et al. (2016), as we exploit a discontinuity over time and compare it cross-sectionally, rather than the opposite.

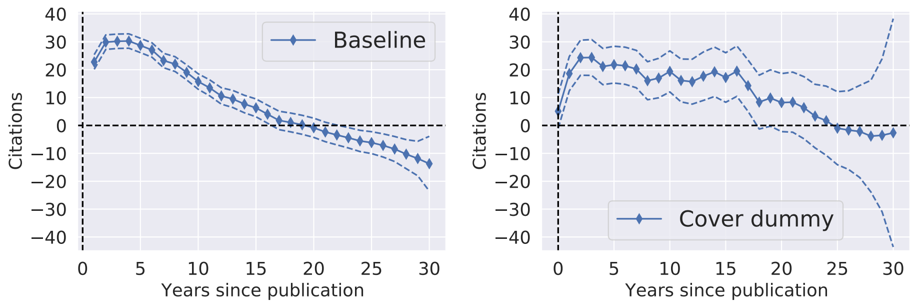
Table 2: Determinants of citations to *Nature* articles

	(1)	(2)
Intercept	-1601.376***	-1617.555***
	(67.955)	(70.038)
Cover dummy (F_i)	16.087***	16.050***
	(0.794)	(0.795)
Years since publication (y)	-0.572***	-0.572***
	(0.038)	(0.038)
Year of publication (p_i)	0.817***	0.825***
	(0.034)	(0.035)
Previous citations		-0.000
		(0.000)
Observations	35,074	35,074

Note: OLS estimates for Equation (1) based on the *Nature* sample. The dependent variable is the number of citations to each *Nature* article in each year. "Previous citations" refers to citations to other articles by authors of the *Nature* article, in the year before its publication. *p<0.1; **p<0.05; ***p<0.01

citations than non-cover articles up to 17 years after publication.

Figure 3: Citations to *Nature* articles: time profile



Note: OLS estimates for Equation (2), sample restricted to *Nature* articles: γ_τ (left) and $\beta_{1,\tau}$ (right); all articles, all years. N=35,074. Reference category: baseline citations in first year considered. Dashed lines represent 95% confidence intervals.

The results presented so far provide clear evidence of a positive difference in citations between cover and non-cover *Nature* articles, consistent with the existing literature (Wang et al., 2017). However, they do not allow us to distinguish between the causal effect of being featured on the cover and a selection effect (more cited authors are more likely candidates for the *Nature* cover). Hence, we turn to the analysis of the extended sample which, by including a time span before the *Nature* publication, allows us to distinguish between ex–ante selection and ex–post causal effects. Table 3 reports OLS estimates for Equation (1),[7] with column (2) presenting an enriched model where we also control for years elapsed since publication of the relevant *Nature* article. Similarly to the *Nature* sample, more recent articles receive more citations than earlier articles, while the number of citations to articles by *Nature* authors falls over time. More importantly, the number of citations to articles by "cover authors" is significantly higher than for non-cover authors, when considering either all articles or articles published before the *Nature* article (columns 1, 2, and 3). This is consistent with hypothesis H1, although it does not provide conclusive evidence, since it cannot rule out an

[7]Due to the very large size of the data set, estimation is obtained by taking mean citation flows for all articles related to a given *Nature* article and published in a given year; we then use a weighted OLS estimator on these clusters, where the weights are determined by the number of articles in each cluster. This explains why the number of observations is around 1 million, much lower than the number of article/year combinations reported in Figure 1 and Table 1.

ex–ante selection mechanism. The findings in column 3 are also consistent with hypothesis H2: articles by cover authors published before the *Nature* article receive more citations than articles by non-cover authors. However, when considering articles published after the *Nature* article (column 4), the number of citations to cover authors is significantly lower than for non-cover authors. This finding is at odds with hypothesis H1 and calls for further attention, as it indicates that articles whose authors *previously* published a cover article on *Nature* obtain *less* citations than authors whose *Nature* paper was not featured on the cover.

Table 3: Determinants of citations to authors of *Nature* articles

	(1)	(2)	(3)	(4)
	All	All	Before	After
Intercept	-75.783***	68.467***	-39.188***	63.817***
	(2.113)	(2.353)	(2.275)	(5.442)
Cover dummy (F_i)	0.563***	0.181***	1.328***	-0.854***
	(0.029)	(0.029)	(0.036)	(0.053)
Years since pub. (y)	-0.130***	-0.237***	-0.114***	-0.180***
	(0.001)	(0.001)	(0.001)	(0.003)
Year of publication (p_i)	0.041***	-0.031***	0.022***	-0.028***
	(0.001)	(0.001)	(0.001)	(0.003)
Years since *Nature* pub.		0.110***		
		(0.001)		
Observations	1,146,345	1,146,345	872,813	273,532

Note: estimates for Equation (1) based on the extended sample. The dependent variable is the average number of citations received each year by articles published in a given year and with a given related *Nature* article (see footnote 7). The last two columns report estimates based on articles published before and after the *Nature* article. Articles published in the same year are included in the "Before" sample. *p<0.1; **p<0.05; ***p<0.01

Estimates for Equation (2) based on the extended sample for $T = 10$, reported in Figure 4, allow us to shed light on this finding.[8] They also

[8] Note that choosing a time window wider or smaller than $T = 10$ would not affect the

provide an explanation for the differences in the estimated coefficient for F_i between the first and the second column, indicating a correlation between F_i and years since the publication of the *Nature* article. We first observe that the number of citations to non-cover authors (left panel) increases following the publication of a *Nature* article, although the difference becomes significant only after 6 years. For what concerns cover authors, it is interesting to note that they receive significantly more citations than non-cover authors even *before* publishing in *Nature* (right panel). Specifically, the estimated $\beta_{1,\tau}$ are significantly different from 0 for $\tau \in \{-10 \ldots 0\}$, providing support for hypothesis H2. The year of publication of the *Nature* article marks a turning point in citations to cover authors: while the estimated cover differential in the number of citations rises over time until publication of the *Nature* article, it falls over time thereafter. In particular, $\beta_{1,\tau} < \beta_{1,1}$, for all $\tau > 1$, with the difference being significant for $\tau > 2$. Cover authors are found to be *penalized* relative to non-cover authors in terms of citations: this provides strong evidence against H1. No significant differences in citations between cover and non-cover authors are found after about 6 years from publication of a *Nature* article.[9]

In order to interpret this finding, we consider estimates of Equation (2) for different sub-samples. Figures 5 and 6 focus on articles published before and after the *Nature* article, respectively. Citations flows to articles published before the *Nature* article are virtually identical to those for the entire sample. Instead, the cover effect is small and not significant for articles published after the *Nature* article. This indicates that the turning point observed for the entire sample is mainly explained by the change in citation flows to articles published before the *Nature* article. *Ceteris paribus*, the publication of a cover article would appear to *crowd out* citations to previous articles by its authors.[10] Other interpretations of this finding are possible, as will

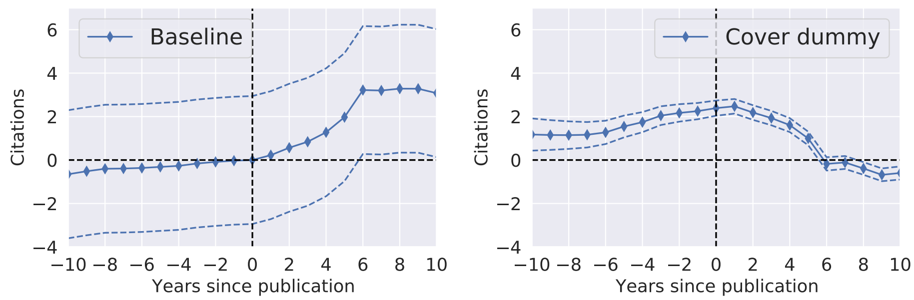
---

estimated coefficients. Estimation with $T = 5$, for example, results in the same estimates for coefficients $\beta_{1,-5}$ to $\beta_{1,5}$ as obtained with $T = 10$. To see why, consider that these coefficients refer to variables interacted with mutually exclusive time–fixed effects and all employ the same reference category (baseline citations in year 0). For the same reason, considering asymmetric time windows (e.g., from -10 to +5) would not change the results.

[9] The results are qualitatively unchanged when controlling for publication year p_i interacted with elapsed years y or with publication year of the *Nature* paper. In particular, fixed effects (baseline plots) change significantly, given the obvious correlation with the new control, but the interactions with the cover dummy are virtually unchanged.

[10] It should be noted that publishing a cover article results in a reduction of citations to other articles *only* when compared to non-cover authors. The overall effect on citations

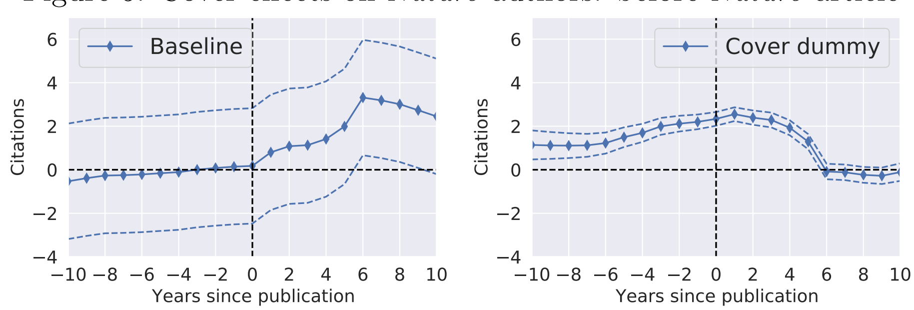
Figure 4: Cover effects on *Nature* authors: all articles



Note: estimates from Equation (2) on the extended sample: all articles, all years ($T = 10$). See Figure 3 for details.

be discussed in Section 5; we anticipate, however, that the effect cannot be related to a change in the publication habits of an author after publishing on *Nature*, since it refers to citations to *previous* publications.

Figure 5: Cover effects on *Nature* authors: before *Nature* article

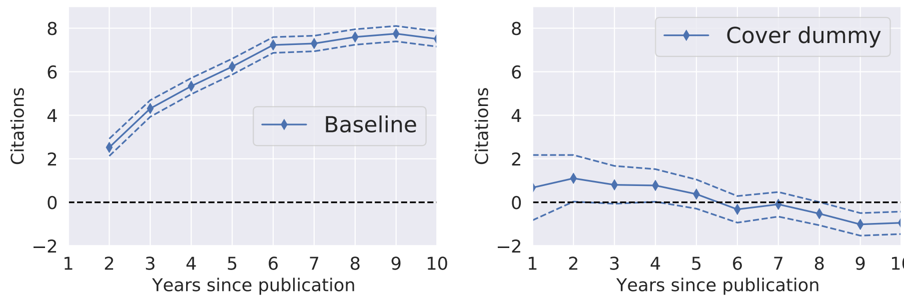


Note: Equivalent of Figure 4 restricted to articles published after the relative *Nature* article.

In order to assess the robustness of our findings, we provide in Appendix A a set of placebo tests reproducing our analyses above with the cover dummy F_i randomly assigned to one article within each issue. No significant change is observed for estimated baseline coefficients (left panels), while the cover effect (right panels) vanishes, with the relative coefficients never being significant.

to articles by cover authors is positive, although not significant.

Figure 6: Cover effects on *Nature* authors: after *Nature* article



Note: Equivalent of Figure 4 restricted to articles published after the relative *Nature* article.

As an additional robustness check, we estimated Equation (2) by introducing additional control variables (see Figure 11 in Appendix B): the results are qualitatively unchanged.

5 Conclusions

We analyze the bibliographic effects of being featured on a journal's cover, by focusing on authors who have published in *Nature*. When considering only articles published in *Nature*, we find that cover articles receive significantly more citations than non-cover articles, consistent with the existing literature (Wang et al., 2017), and this difference is long-lasting. This remains true when controlling for the number of citations received by the authors before publication of the *Nature* article. However, when considering citations to all articles (past and future) by authors who have published in *Nature*, we find that authors whose articles have been featured on the cover receive *ex ante* more citations than non-cover authors, supporting the hypothesis that being featured on the cover is subject to selection effects. More importantly, the hypothesis that the cover articles give a boost to citations receveived by an author is refuted: a significant downturn is observed in citations received by cover authors, relative to non–cover authors, in the years following the *Nature* publication.

These results, apparently in contrast with the expectation that the publication of a cover paper should increase the authors' visibility in the scientific community, are compatible with two different explanations. The first is a crowding–out effect: when an author publishes a *Nature* cover article, this article tends to be highly cited and can overshadow other articles by the same author. As a consequence, other articles obtain relatively less citations in the following years.

A second interpretation can be suggested by considering the left and right plots of Figure 4 together: the *ex-ante* relative advantage of cover authors may be offset by the fact that non-cover authors also published an article in *Nature*. For instance, it might be the case that only high–impact researchers are featured on the cover of *Nature*, but publishing an article in *Nature* (be it on the cover or not) is one way to become a high–impact researcher – and hence close the gap in citations relative to those who already are (Sekara et al., 2018). In other words, we would not be observing a negative effect of publishing on the cover, but rather a positive effect of publishing on *Nature*, which is instead absent for already prominent scholars. These two explanations are not mutually exclusive, and could both contribute to explain our results.

Overall, our analysis casts doubts on the view that being featured on the cover of a prestigious scientific journal results in a boost in the number of citations received by an author. Indeed, being featured on the cover is found to *signal* scientific prominence, rather than enhance it. However, our analysis does not provide indications about the underlying mechanisms, that deserve further investigation in future research. More generally, the identification and quantification of determinants of citation patterns that are unrelated to the intrinsic characteristics of cited works is a promising endeavor, as it can contribute to improve our understanding — and possibly the construction — of citation–based impact metrics.

References

Abramo, G., C. A. D'Angelo, and G. Murgia (2013). The collaboration behaviors of scientists in Italy: A field level analysis. *Journal of Informetrics 7*(2), 442–454.

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation 12*(3), 159–170.

Althouse, B. M., J. D. West, C. T. Bergstrom, and T. Bergstrom (2009). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science and Technology 60*(1), 27–34.

Archambault, É. and V. Larivière (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics 79*(3), 635–649.

Ayres, I. and F. E. Vars (2000). Determinants of Citations to Articles in Elite Law Reviews. *The Journal of Legal Studies 29*(S1), 427–450.

Azoulay, P., T. Stuart, and Y. Wang (2014). Matthew: Effect or Fable? *Management Science 60*(1), 92–109.

Battiston, P. (2014). Citations are Forever: Modeling Constrained Network Formation. Working Paper 2014/19, LEM Paper Series, Sant'Anna School of Advanced Studies.

Bollen, J., H. Van de Sompel, A. Hagberg, and R. Chute (2009). A Principal Component Analysis of 39 Scientific Impact Measures. *PloS one 4*(6), e6022.

Bordons, M., M. Fernández, and I. Gómez (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics 53*(2), 195–206.

Borjas, G. J. and K. B. Doran (2015). Prizes and Productivity: How Winning the Fields Medal Affects Scientific Output. *Journal of Human Resources 50*(3), 728–758.

Brown, H. (2007). How impact factors changed medical publishing – and science. *BMJ 334*(7593), 561–564.

Callaway, E. (2016). Beat it, impact factor! Publishing elite turns against controversial metric. *Nature News 535*(7611), 210.

Demetrescu, C., F. Lupia, A. Mendicelli, A. Ribichini, F. Scarcello, and M. Schaerf (2019). On the shapley value and its application to the italian vqr research assessment exercise. *Journal of Informetrics 13*(1), 87–104.

Di Vaio, G., D. Waldenström, and J. Weisdorf (2012). Citation success: Evidence from economic history journal publications. *Explorations in Economic History 49*(1), 92–104.

Fama, E. F., L. Fisher, M. C. Jensen, and R. Roll (1969). The adjustment of stock prices to new information. *International Economic Review 10*(1), 1–21.

Feenberg, D., I. Ganguli, P. Gaule, and J. Gruber (2017). It's good to be first: Order bias in reading and citing NBER working papers. *Review of Economics and Statistics 99*(1), 32–39.

Frandsen, T. F. and J. Nicolaisen (2013). The ripple effect: Citation chain reactions of a nobel prize. *Journal of the American Society for Information Science and Technology 64*(3), 437–447.

Grembi, V., T. Nannicini, and U. Troiano (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, 1–30.

Heckman, J. J. and S. Moktan (2018). Publishing and Promotion in Economics: The Tyranny of the Top Five. Technical Report 25093, National Bureau of Economic Research.

Johnson, D. (1997). Getting Noticed in Economics: The Determinants of Academic Citations. *The American Economist 41*(1), 43–52.

Kong, L. and D. Wang (2020). Comparison of citations and attention of cover and non-cover papers. *Journal of Informetrics 14*(4), 101095.

Medoff, M. H. (2006). Evidence of a Harvard and Chicago Matthew effect. *Journal of Economic Methodology 13*(4), 485–506.

Mingers, J. (2008). Exploring the dynamics of journal citations: Modelling with s-curves. *Journal of the Operational Research Society 59*(8), 1013–1025.

Monastersky, R. (2005). The Number That's Devouring Science. *Chronicle of Higher Education 52*(8), 14.

PLoS Medicine Editors (2006). The impact factor game. *PloS Medicine* (3(6)), e291.

Price, D. d. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the Association for Information Science and Technology 27*(5), 292–306.

Robinson, C. D., J. Gallus, M. G. Lee, and T. Rogers (2021). The demotivating effect (and unintended message) of awards. *Organizational Behavior and Human Decision Processes 163*, 51–64.

Sekara, V., P. Deville, S. E. Ahnert, A.-L. Barabási, R. Sinatra, and S. Lehmann (2018). The chaperone effect in scientific publishing. *Proceedings of the National Academy of Sciences 115*(50), 12603–12607.

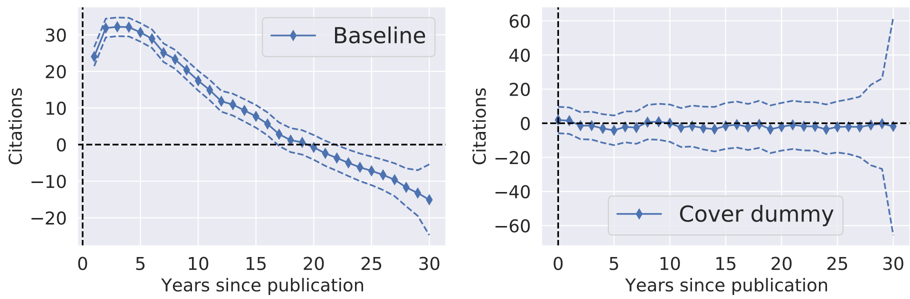
Simons, K. (2008). The Misused Impact Factor. *Science 322*(5899), 165–165.

Wang, G., J. Gregory, X. Cheng, and Y. Yao (2017). Cover stories: An emerging aesthetic of prestige science. *Public Understanding of Science 26*(8), 925–936.

Wang, X., C. Liu, and W. Mao (2015). Does a paper being featured on the cover of a journal guarantee more attention and greater impact? *Scientometrics 102*(2), 1815–1821.

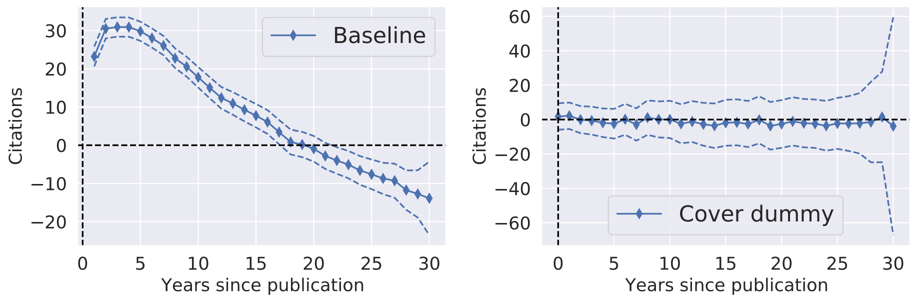
A Placebo tests

Figure 7: Nature articles - placebo test (baseline)



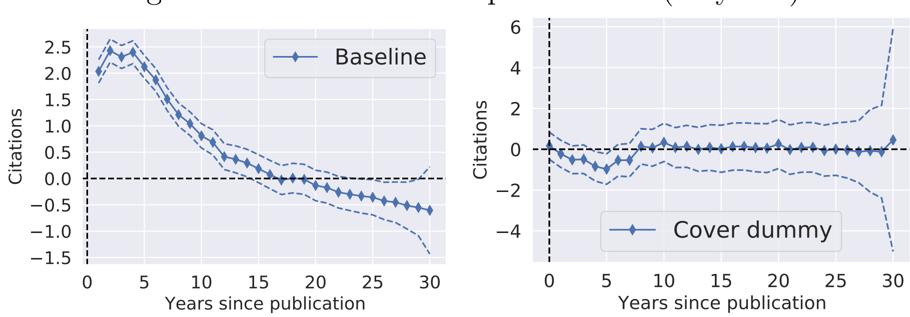
Note: Equivalent of Figure 3 with the "cover" dummy randomly assigned to one article within each issue.

Figure 8: Nature articles - placebo test (non-self)



Note: Equivalent of Figure 7 excluding self-citations.

Figure 9: Nature articles - placebo test (only self)

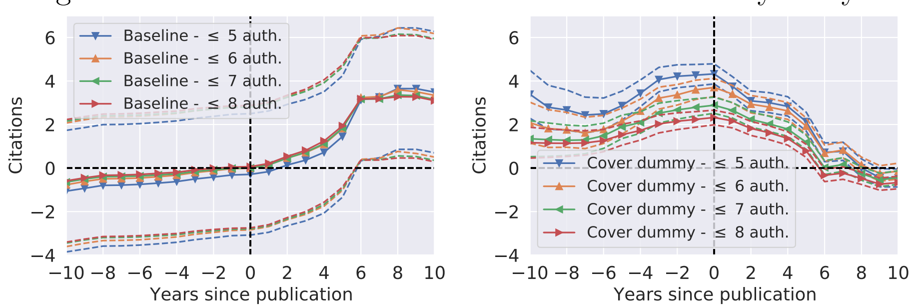


Note: Equivalent of Figure 7 restricting to self-citations.

B Sensitivity tests

Figure 10 displays the equivalent of Figure 4 when changing the maximum number of authors in the *Nature* article (which is set to 8 in the main analysis). We can observe that there is little qualitative change in the dynamic effect of the cover dummy variable, albeit the effect does decrease in magnitude as the number of authors increases.

Figure 10: Cover effects on *Nature* authors: sensitivity analysis

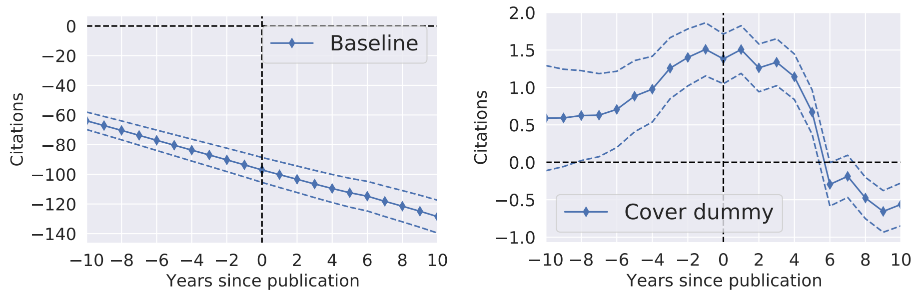


Note: Equivalent of Figure 4 when restricting the sample to *Nature* articles with less than the given number of authors.

Figure 11 displays the results obtained by re–estimating Equation (2) while also controlling for the number of authors of each article, and for the interaction between the year of publication of the given article and the year of publication of its related *Nature* article. As shown in the right panel, the coefficients for the cover dummies are virtually unchanged. Note that the values for the baseline coefficients (left panel) have no direct interpretation, as they refer to the comparison with a reference category which corresponds to an article published in year 0 (that is, $p_i = 0$) by an author who also published on *Nature* in year 0 (that is, $N_i = 0$).

Similar results are obtained by controlling for the interaction between the year of publication of the given article and the year of publication of the related *Nature* article, without including the number of authors. The pattern exhibited by cover interaction dummies is also robust to whether each article is attributed the same weight (as in the main analysis) or a weight that is the inverse of the number of articles related to a given *Nature* article (i.e., weights are normalized so that each *Nature* article is given the same

Figure 11: Estimation with additional controls



Note: estimates from Equation (2), additionally controlling for number of authors of an article, and interaction of its publication year with the related *Nature* article's publication year: γ_τ (left) and $\beta_{1,\tau}$ (right); all articles, all years ($T = 10$). See Figure 3 for details — notice however the different scale.

importance). Finally, qualitatively similar results are obtained by replacing the citation count $c_{i,t}$ with $\log(c_{i,t} + 1)$.