an open access journal

RESEARCH ARTICLE

Inferring the causal effect of journals on citations

V. A. Traag

Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands

Keywords: Bayesian model, causal inference, citations, journal effects, science of science

ABSTRACT

Articles in high-impact journals are, on average, more frequently cited. But are they cited more often because those articles are somehow more "citable"? Or are they cited more often simply because they are published in a high-impact journal? Although some evidence suggests the latter, the causal relationship is not clear. We here compare citations of preprints to citations of the published version to uncover the causal mechanism. We build on an earlier model of citation dynamics to infer the causal effect of journals on citations. We find that high-impact journals select articles that tend to attract more citations. At the same time, we find that high-impact journals augment the citation rate of published articles. Our results yield a deeper understanding of the role of journals in the research system. The use of journal metrics in research evaluation has been increasingly criticized in recent years and article-level citations are sometimes suggested as an alternative. Our results show that removing impact factors from evaluation does not negate the influence of journals. This insight has important implications for changing practices of research evaluation.

Citation: Traag, V. A. (2021). Inferring the causal effect of journals on citations. *Quantitative Science Studies*, *2*(2), 496–504. https://doi.org/10.1162/qss_a_00128

DOI:
https://doi.org/10.1162/qss_a_00128

Peer Review:
https://publons.com/publon/10.1162/qss_a_00128

Supporting Information:
https://doi.org/10.1162/qss_a_00128

Received: 6 November 2020
Accepted: 10 January 2021

Corresponding Author:
V. A. Traag
v.a.traag@cwts.leidenuniv.nl

Handling Editor:
Staša Milojević

1. INTRODUCTION

Journals play a central role in scholarly communication, yet their role is also contested. The journal impact factor in particular has been criticized on several accounts (Larivière & Sugimoto, 2019). The main critique is its pervasive use in the context of research evaluation, for example in tenure decisions (McKiernan, Schimanski et al., 2019). Scientists shape their research with impact factors in mind (Müller & de Rijcke, 2017; Rushforth & de Rijcke, 2015). In a meeting in San Francisco in 2012, cell biologists called for a ban on the impact factor from research evaluation, and conjoined the "San Francisco Declaration on Research Assessment"[1] (DORA). A group of researchers and editors called for publishing entire citation distributions instead of impact factors, to counter inappropriate use (Larivière, Kiermer et al., 2016). More recently, a group of editors and researchers came together and called for "rethinking impact factors" (Wouters, Sugimoto et al., 2019).

At the same time, journal impact is one of the most clear predictors of future citations (Abramo, D'Angelo, & Felici, 2019; Callaham, 2002; Levitt & Thelwall, 2011; Stegehuis, Litvak, & Waltman, 2015). The question is why. Possibly, high-impact journals select articles that somehow tend to be cited frequently. Another possibility is that articles are cited more frequently *because* they are published in a high-impact journal, not because they tend to be cited frequently *per se*. Neither citations of an article nor the journal in which it is published needs to

[1] https://sfdora.org

The MIT Press

be representative of "quality." Here, we simply study whether citations of an article are influenced by the journal in which it is published, not their relationship to "quality."

Answering this question is not straightforward. In rare cases, publications appear in multiple journals, and researchers found that the version in a higher impact journal was more frequently cited than its twin in a lower impact journal (Cantrill, 2016; Larivière & Gingras, 2010; Perneger, 2010). However, duplicate publications are quite special, limiting the generalizability of this observation. Some other earlier work claimed that citations were not affected by the journal (Seglen, 1994).

We answer this question by comparing citations of preprints with citations of the published version. The number of citations C may be influenced by both the latent citation rate ϕ and the journal J in which the article is published (Figure 1). Possibly, high-impact journals perform a stringent peer review of articles, selecting only articles with a high latent citation rate, so that ϕ influences the journal J. The latent citation rate itself may be influenced by many factors and characteristics (Onodera & Yoshikane, 2015) and motivations for citing the paper (Bornmann & Daniel, 2008). These factors are not limited to the characteristics of the paper itself, but may also include author reputation (Petersen, Fortunato et al., 2014) or institutional reputation (Medoff, 2006). Regardless of which factors influence the latent citation rate, the number of citations of the preprint before it is published in a journal C' is unaffected by where it will be published and is affected only by the latent citation rate ϕ. We rely on this insight to estimate the causal effect of the journal on citations $\Pr(C \mid \mathrm{do}(J))$. The identification of the causal effect is possible because of the so-called "effect restoration" (Kuroki & Pearl, 2014), provided we can estimate $\Pr(C' \mid \phi)$. We construct a parametric model that provides exactly such an estimate.

2. METHODOLOGY

We gathered information about 1,341,016 preprints from arXiv, and identified the published version for 727,186 preprints (54%; see Supplementary Material for more details). We extracted citations of both the preprint version and the published version from references in Scopus. Preprint dates, publication dates, and citation dates are all extracted from Crossref, using a daily granularity. We used the major subject headings of arXiv as field definitions. The impact of journals is calculated as the average number of citations received in the first 5 years after publication for all research articles and reviews in Scopus. We perform our analysis per year (2000–2016) and field, as the journal effect may vary per year and field. Moreover, we restrict

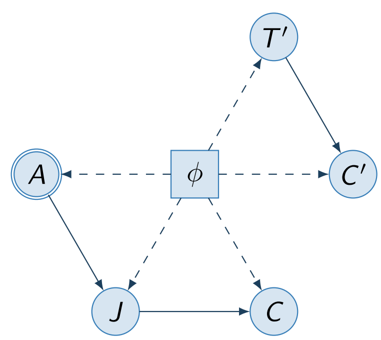

Figure 1. Simple causal model of the confounding effect of the latent citation rate ϕ of an article being published in a journal J and the citations it accrues C. In contrast, citations of preprints C' are affected by the latent citation rate ϕ only. The selection bias on arXiv preprints A does not bias the causal effect of J on C once ϕ is controlled for. The time before publication T' affects preprint citations C' and complicates the analysis.

our analysis to journals that have at least 20 articles that were published at least 30 days after appearing as a preprint on arXiv (Figure S1). Clearly, our data has a selection bias (Bareinboim & Pearl, 2012) on papers being submitted to arXiv or not (A). However, we can show that this does not affect our estimate of the causal effect $\Pr(C \mid \mathrm{do}(J))$ (see Supplementary Material).

Time complicates our analysis. The time T' before a preprint was published, the *preprint duration*, will clearly affect the number of prepublication citations C', while the total time since publication T will affect the postpublication citations C. Preprints with a higher latent citation rate may perhaps be more quickly published, thus affecting T'. To tackle this problem, we model the full temporal dynamics of both pre- and postpublication citations.

Citation dynamics are influenced by a wide range of factors, such as a rich-get-richer effect and a clear temporal decay (Fortunato, Bergstrom et al., 2018), but was captured reasonably well by a recent model by Wang, Song, and Barabási (2013). We build on that model and include a parameter that modulates the citation rate based on where the article is published. We assume that the number of citations $c_i(t)$ that article i receives at time t is distributed as

$$c_i(t) \sim \mathrm{Poisson}[\lambda_i(t) f_i(t)(m + C_i(t-1))], \tag{1}$$

with effective citation rate $\lambda_i(t)$ and $C_i(t) = \sum_{\tau=0}^{t} c_i(\tau)$ the cumulative number of citations, and m a parameter affecting the initial citation accumulation. The temporal decay of the accumulation of citations is captured by $f_i(t)$, which is modeled by an exponential distribution, with inverse rate β_i. We assume that preprint i attracts citations at an effective rate of ϕ_i, where ϕ_i is the latent citation rate of article i. The published version attracts citations at an effective rate of $\phi_i\theta_{J_i}$, where θ_{J_i} is the *journal citation multiplier* for journal J_i in which article i is published. We equate θ_j with the causal effect on citations of publishing in journal j, which is identical for all articles published in journal j, regardless of the characteristics of those papers. We call $C_i' = C_i(T_i')$ the prepublication citations and $C_i = C_i(T_i) - C_i(T_i')$ the postpublication citations. The expected number of long-term citations is about

$$m\left(e^{\phi_i \theta_{J_i}} - 1\right), \tag{2}$$

assuming prepublication citations are negligible (see Supplementary Material).

The selection of articles by peer review is assumed to lead to a distribution of latent citation rates for journal j,

$$\phi_i \sim \mathrm{LogNormal}\left(\Phi_j, \epsilon_j\right). \tag{3}$$

If Φ_j is high, journal j will tend to publish articles of higher latent citation rates ϕ_i. The median latent citation rate of journal j is e^{Φ_j}. Effectively, this is a Bayesian hierarchical model, and we specify informed prior distributions based on earlier results (Wang et al., 2013) (see Supplementary Material for full details and analysis of the model). We illustrate the model in Figure 2.

3. RESULTS

The numbers of pre- and postpublication citations are not clearly related (Figure 3, panel A). The numbers of prepublication citations also do not clearly relate to journal impact (Figure 3, panel B). The relation between preprint duration and the number of prepublication citations is also not clear (Figure 3, panel C). The ratio of postpublication citations and prepublication citations is higher for high-impact journals (Figure 3, panel D). Articles in high-impact journals accumulate more postpublication citations relative to prepublication citations compared to articles that have appeared in lower impact journals. These results are possibly obfuscated by two counteracting effects: Higher latent citation rates lead to higher prepublication citations,

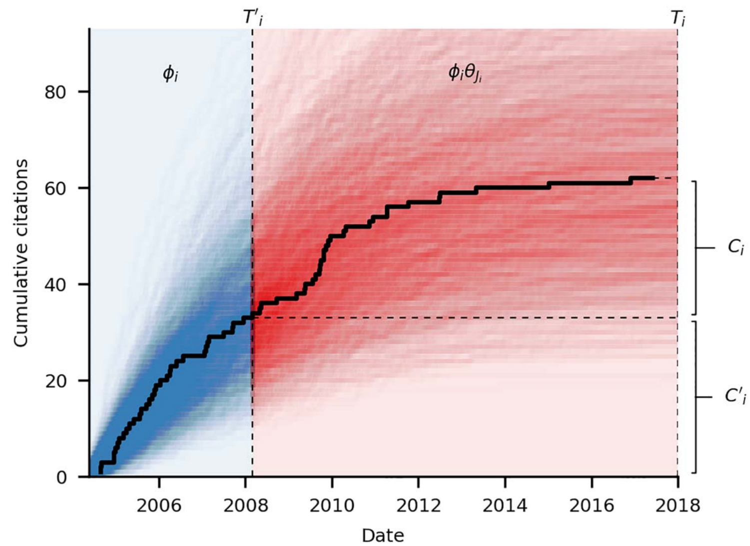

Figure 2. Illustration of citation dynamics. This example, `astro-ph/0405353`, was first submitted to arXiv in 2004 and was published in *Journal of Cosmology and Astroparticle Physics* almost 4 years later ($T_i' = 1{,}385$). It was cited 33 times before it was published ($C_i' = 33$), and 29 times after it was published ($C_i = 29$). We assume citations are attracted at a rate of ϕ_i before it was published and at a rate of $\phi_i\theta_{J_i}$ after it was published. The thick solid line represents the empirically observed number of citations. The thin lines in the background represent samples from the posterior predictive distribution of our model.

but perhaps also to shorter preprint durations, reducing the time to attract prepublication citations. The model that we constructed is intended to address this issue.

We here report results from our model for the five largest fields and the publication year 2016. Other fields and years show qualitatively similar results (see Figures S2 and S4). Our model presents a good fit of both pre- and postpublication citations (Figure S5).

The journal citation multiplier is consistently higher than 1 (Figure 4, panel A). Publishing in journals, compared to being available on arXiv only, multiplies the citation rate substantially, as expected. For example, *Nature* shows a multiplier of 6.0–9.9 (95% CI) for papers published in 2016 in the subject of Condensed Matter and *Science* shows a multiplier of 7.5–12.0 (95% CI) for such papers. Using the median estimates and the approximation in Eq. 2, this implies that a

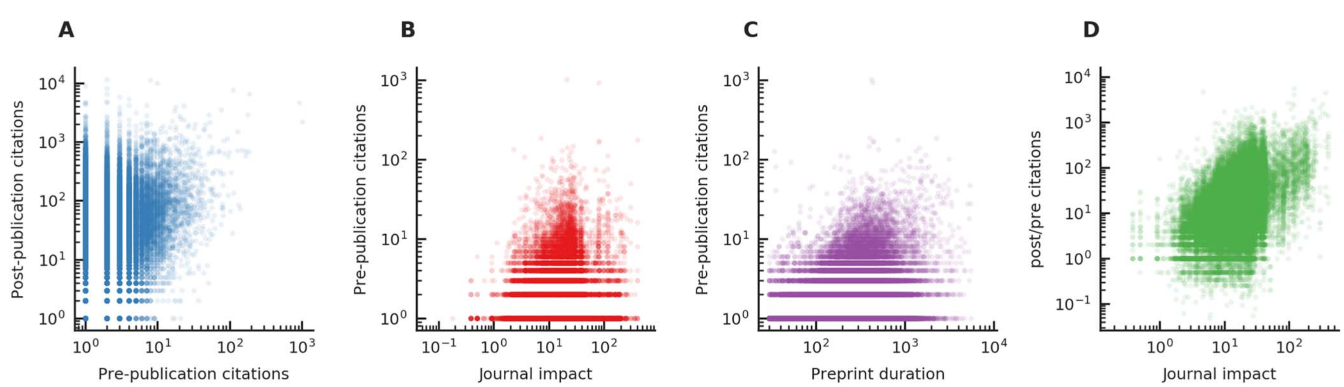

Figure 3. Impact versus pre- and postpublication citations.

Condensed Matter article published in *Nature* in 2016 that obtained about 200 citations would not have obtained even 10 citations had it been available on arXiv only. Had it been published in *Science* instead, it would have obtained almost 350 citations. This is only an illustration: Both parameter estimates and the citation dynamics themselves exhibit considerable uncertainty (see Supplementary Material).

Most relevant to our question, higher impact journals tend to show higher citation multipliers. The correlation between the (logarithm of) the journal impact and the (logarithm of) the median journal citation multiplier θ_j is on average 0.45 for each combination of field and year. It ranges from 0.063 for High Energy Physics in 2002 to 0.79 for Astrophysics in 2012. Interestingly, the correlation grows stronger for High Energy Physics and Astrophysics over time, hovering around 0.6–0.7 for recent years (Figure S3).

At the same time, the median latent citation rate e^{Φ_j} is also clearly increasing with journal impact (Figure 4, panel B). For example, the U.S.-based *Physical Review Letters* has a relatively high journal impact and shows a latent citation rate of 0.15–0.17 (95% CI) for Condensed Matter in 2016. Its lower impact European counterpart *Europhysics Letters* shows a latent citation rate of 0.013–0.027 (95% CI) in the same field and year. Overall, the correlation between the (logarithm of) the journal impact and Φ_j is on average 0.54 for each combination of field and year. For High Energy Physics in 2002 the correlation is 0.72, while for Astrophysics in 2012 the correlation is 0.050. The highest correlation of 0.85 is observed for Astrophysics in 2006. This correlation grows weaker for High Energy Physics and Astrophysics over time (Figure S3). The median effective citation rate of a journal is $e^{\Phi_j}\theta_j$, which aligns closely with the observed journal impact (Figure S6).

The latent citation rates also vary within journals, and are controlled by ϵ_j. Journals with a higher ϵ_j tend to publish articles with a larger variety of latent citation rates. For example, *Europhysics Letters* shows an ϵ_j of 0.7–1.1 (95% CI), while *Science* shows an ϵ_j of 0.2–0.3 (95% CI), resulting in a broader distribution of ϕ_i for *Europhysics Letters* than *Science*. In general, high-impact journals show more narrow distributions of latent citation rates than lower impact journals (Figure 4, panel C).

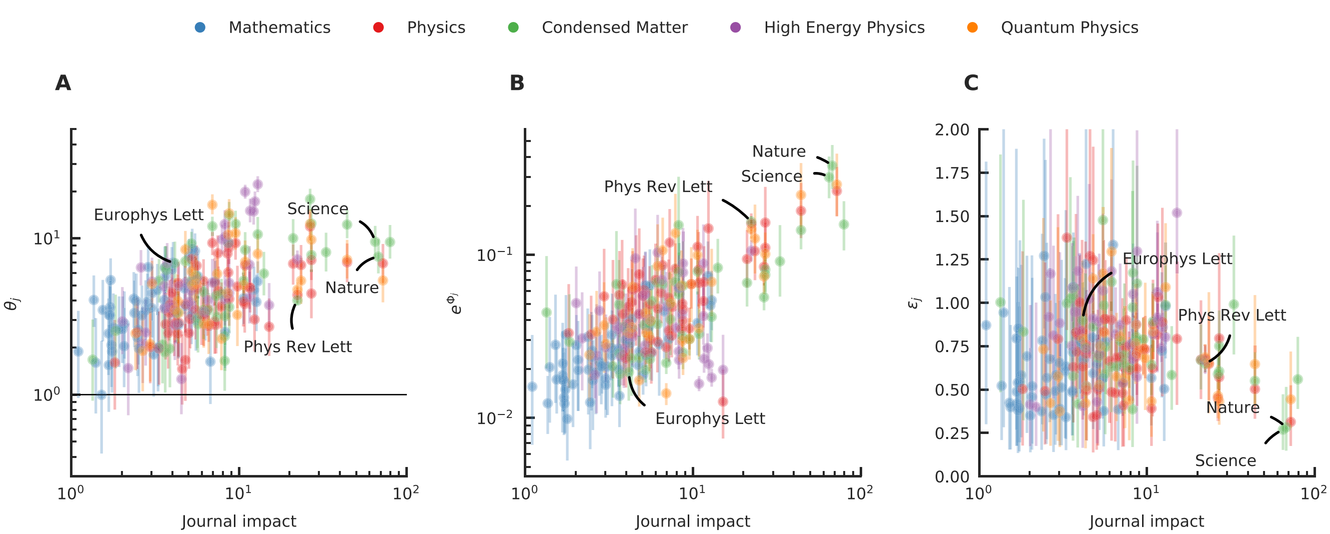

Figure 4. Posterior results for model of citation dynamics for five largest fields and publication year 2016. Error bars represent the average 95% credible interval. Highlighted journals indicate results in the field of Condensed Matter.

4. DISCUSSION

Why articles in high-impact journals attract more citations is a fundamental question. We have provided clear evidence that articles in high-impact journals are highly cited because of two effects. On the one hand, articles that attract more citations are more likely to be published in high-impact journals. On the other hand, articles in high-impact journals will be cited even more frequently because of the publication venue. This amplifies the cumulative advantage effect for citations (Price, 1976).

A recent publication (Kim, Portenoy et al., 2020) took a similar approach and compared citations of preprints with citations of the published version. Using a more rudimentary model they obtained similar results and also find an influence of the journal on citations, although they do not address the causal mechanism. They also find that preprints with more citations are more likely to be published, but do not analyze in what journals they are published.

Several mechanisms may play a role in the causal effect of journals on citations. High-impact journals tend to have a higher circulation (Peritz, 1995), and reach a wider audience. In addition, researchers may prefer to cite an article from a high-impact journal over an article from a low-impact journal, even if both articles would be equally fitting. Both mechanisms are consistent with our results and earlier results (Cantrill, 2016; Kim et al., 2020; Larivière & Gingras, 2010; Perneger, 2010). Distinguishing between these two causal mechanisms is difficult (Davis, 2010) and should be investigated further.

An alternative explanation may be that published preprints are more highly cited because the preprints were improved by high-quality peer review in high-impact journals. We deem this an unlikely scenario. Differences between the preprint and the published version are textually minor (Klein, Broadwell et al., 2016). Those modifications can of course be substantively important. Peer review may substantially improve and strengthen a manuscript. Nonetheless, we think it is unlikely to alter a paper's core contribution so as to affect its citation rate considerably.

Our analysis is limited to mostly physics and mathematics because of our reliance on arXiv. We expect to see similar effects in the medical sciences and the social sciences, in line with earlier results (Cantrill, 2016; Larivière & Gingras, 2010; Perneger, 2010). It would be interesting to replicate our analysis on younger preprint repositories, such as bioRxiv or SocArxiv, once they have had more time to accumulate citations. Another limitation is that we considered references from published articles only. It would be interesting to include also the references of preprints. This presumably increases the number of prepublication citations (Larivière, Sugimoto et al., 2014), which may decrease the overall inferred journal causal effect.

In our model we assumed that the effect of publishing in a journal is identical for all articles published in that journal. However, the effect of publishing in a journal may possibly vary for different articles. For example, articles from well-known authors may be cited frequently regardless of the exact journal in which they are published, while articles from more junior authors may benefit more from publishing in high-impact journals. Teasing out these different effects is not straightforward, but presents an interesting avenue for future research.

The latent citation rate itself may be influenced by many factors and characteristics of the paper (Onodera & Yoshikane, 2015) and motivations for citing the paper (Bornmann & Daniel, 2008). Overall, our results suggest that the characteristics (X_1, X_2, …) that drive citations (C) overlap or correlate with factors that drive journal (J) peer review (Figure 5). For example, novelty, relevance, and scientific breadth (X_2 to X_4) may affect both journal evaluation and citations directly, while methodological aspects affect journal evaluation (X_1) and authors' reputation (X_5) only affects citations. Because the journal also affects citations, methodological

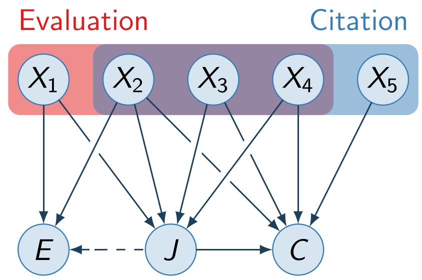

Figure 5. Causal model of factors and characteristics X_1, X_2, …, journals J, citations C, and evaluation E.

aspects would have an indirect effect on citations in this example. What factors drive journal evaluation and what factors drive citations is not clear and should be further investigated.

We hypothesize that a subset of factors that are used in journal evaluation are also used in postpublication research evaluation, such as the UK REF (Traag & Waltman, 2019). This means that research evaluation (E) tends to correlate with journals (J) because of underlying common factors (Figure 5). Even if factors that influence research evaluation do not influence citations directly, they will still correlate because of the mediating effect of the journal. For example, if methodological aspects (X_1) affect research evaluation (E), it would correlate with citations (C) only because methodological aspects affect the journal (J). If our hypothesis holds true, citations would be indicative of the evaluation of articles only *because* they were published in a particular journal. In that case, citations should not be normalized based on the journal in which they are published, as was attempted by Zitt, Ramanana-Rahary, and Bassecoulard (2005). Doing so would effectively control for the journal, thereby blocking these causal pathways. Indeed, Adams, Gurney, and Jackson (2008) find that journal-normalized citations do not correlate with evaluation. Similarly, Eyre-Walker and Stoletzki (2013) report an absence of various correlations with evaluations when controlling for the journal. These results provide some evidence for our hypothesis. Journal metrics might even be a more appropriate indicator than citations to individual articles, as was suggested by Waltman and Traag (2020), although our results neither affirm nor refute this possibility.

Possibly, evaluation itself is also affected directly by the journal in which an article is published, and depending on the context, perhaps also by its citations. Indeed, the proposed causal diagram only captures part of a larger web of entanglement.

The use of citations and journals in research evaluation is often debated. Removing the use of journal metrics from research evaluation, as for example advocated by DORA, may decrease the pressure on authors to publish in high-impact journals. The use of article-level citations for evaluation could be condoned by DORA, but the use of journal metrics could not. Even if journal metrics were to be removed from research evaluation, journals would continue to play a role in research evaluation, albeit indirectly. Evaluating researchers based on citations then may still reward authors who publish in high-impact journals. This may effectively exert selective pressures that drive the evolution of the research system (Smaldino & McElreath, 2016). Simply removing impact factors from research evaluation therefore does not negate the influence of journals.

ACKNOWLEDGMENTS

I thank Rodrigo Costas, Ludo Waltman, Jesper Schneider, and other colleagues from CWTS. I gratefully acknowledge use of the Shark cluster of the LUMC for computation time.

COMPETING INTERESTS

The author has no competing interests.

DATA AVAILABILITY

All data necessary to reproduce the results in this analysis is available from Traag (2020a) and all source code is available from Traag (2020b).

REFERENCES

Abramo, G., D'Angelo, C. A., & Felici, G. (2019). Predicting publication long-term impact through a combination of early citations and journal impact factor. *Journal of Informetrics*, *13*, 32–49. **DOI:** https://doi.org/10.1016/j.joi.2018.11.003

Adams, J., Gurney, K., & Jackson, L. (2008). Calibrating the zoom—A test of Zitt's hypothesis. *Scientometrics*, *75*, 81–95. **DOI:** https://doi.org/10.1007/s11192-007-1832-7

Bareinboim, E., & Pearl, J. (2012). Controlling selection bias in causal inference. In *International Conference on Artificial Intelligence and Statistics,* vol. 22, (pp. 100–108).

Bornmann, L., & Daniel, H. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*, 45–80. **DOI:** https://doi.org/10.1108/00220410810844150

Callaham, M. (2002). Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA*, *287*, 2847. **DOI:** https://doi.org/10.1001/jama.287.21.2847, **PMID:** 12038930

Cantrill, S. (2016). Imperfect impact. *Chemical connections* blog. January 23. https://stuartcantrill.com/2016/01/23/imperfect-impact/

Davis, P. (2010). Impact factors – A self-fulfilling prophecy? *The Scholarly Kitchen* blog, June 9. https://scholarlykitchen.sspnet.org/2010/06/09/impact-factors-a-self-fulfilling-prophecy/.

Eyre-Walker, A., & Stoletzki, N. (2013). The assessment of science: The relative merits of post-publication review, the impact factor, and the number of citations. *PLOS Biology*, *11*, e1001675. **DOI:** https://doi.org/10.1371/journal.pbio.1001675, **PMID:** 24115908, **PMCID:** PMC3792863

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., … Barabási, A.-L. (2018). Science of science. *Science*, *359*. **DOI:** https://doi.org/10.1126/science.aao0185, **PMID:** 29496846, **PMCID:** PMC5949209

Kim, L., Portenoy, J. H., West, J. D., & Stovel, K. W. (2020). Scientific journals still matter in the era of academic search engines and preprint archives. *Journal of the Association for Information Science and Technology*, *71*, 1218–1226. **DOI:** https://doi.org/10.1002/asi.24326

Klein, M., Broadwell, P., Farb, S. E., & Grappone, T. (2016). Comparing published scientific journal articles to their pre-print versions. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries - JCDL '16* (pp. 153–162). New York: ACM Press. **DOI:** https://doi.org/10.1145/2910896.2910909

Kuroki, M., & Pearl, J. (2014). Measurement bias and effect restoration in causal inference. *Biometrika*, *101*, 423–437. **DOI:** https://doi.org/10.1093/biomet/ast066

Larivière, V., & Gingras, Y. (2010). The impact factor's Matthew Effect: A natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, *61*, 424–427. **DOI:** https://doi.org/10.1002/asi.21232

Larivière, V., Kiermer, V., MacCallum, C. J., McNutt, M., Patterson, M., … Curry, S. (2016). A simple proposal for the publication of journal citation distributions. *bioRxiv*, 062109. **DOI:** https://doi.org/10.1101/062109

Larivière, V., & Sugimoto, C. R. (2019). The journal impact factor: A brief history, critique, and discussion of adverse effects. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 3–24). Cham: Springer. **DOI:** https://doi.org/10.1007/978-3-030-02511-3_1

Larivière, V., Sugimoto, C. R., Macaluso, B., Milojević, S., Cronin, B., & Thelwall, M. (2014). arXiv E-prints and the journal of record: An analysis of roles and relationships. *Journal of the Association for Information Science and Technology*, *65*, 1157–1169. **DOI:** https://doi.org/10.1002/asi.23044

Levitt, J. M., & Thelwall, M. (2011). A combined bibliometric indicator to predict article impact. *Information Processing and Management*, *47*, 300–308. **DOI:** https://doi.org/10.1016/j.ipm.2010.09.005

McKiernan, E. C., Schimanski, L. A., Muñoz Nieves, C., Matthias, L., Niles, M. T., & Alperin, J. P. (2019). Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. *eLife*, *8*, e47388. **DOI:** https://doi.org/10.7554/eLife.47338, **PMID:** 31364991, **PMCID:** PMC6668985

Medoff, M. H. (2006). Evidence of a Harvard and Chicago Matthew Effect. *Journal of Economic Methodology*, *13*, 485–506. **DOI:** https://doi.org/10.1080/13501780601049079

Müller, R., & de Rijcke, S. (2017). Thinking with indicators. Exploring the epistemic impacts of academic performance indicators in the life sciences. *Research Evaluation*, *26*, 157–168. **DOI:** https://doi.org/10.1093/reseval/rvx023

Onodera, N., & Yoshikane, F. (2015). Factors affecting citation rates of research articles. *Journal of the Association for Information Science and Technology*, *66*, 739–764. **DOI:** https://doi.org/10.1002/asi.23209

Peritz, B. C. (1995). On the association between journal circulation and impact factor. *Journal of Information Science*, *21*, 63–67. **DOI:** https://doi.org/10.1177/016555159502100108

Perneger, T. V. (2010). Citation analysis of identical consensus statements revealed journal-related bias. *Journal of Clinical Epidemiology*, *63*, 660–664. **DOI:** https://doi.org/10.1016/j.jclinepi.2009.09.012, **PMID:** 20097531

Petersen, A. M., Fortunato, S., Pan, R. K., Kaski, K., Penner, O., … Pammolli, F. (2014). Reputation and impact in academic careers. *Proceedings of the National Academy of Sciences*, *111*, 15316–15321. **DOI:** https://doi.org/10.1073/pnas.1323111111, **PMID:** 25288774, **PMCID:** PMC4217436

Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, *27*, 292–306. **DOI:** https://doi.org/10.1002/asi.4630270505

Rushforth, A., & de Rijcke, S. (2015). Accounting for impact? The Journal Impact Factor and the making of biomedical research in the Netherlands. *Minerva*, *53*, 117–139. **DOI:** https://doi.org/10.1007/s11024-015-9274-5, **PMID:** 26097258, **PMCID:** PMC4469321

Seglen, P. O. (1994). Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science*, *45*, 1–11. **DOI:** https://doi.org/10.1002/(SICI)1097-4571(199401)45:1<1::AID-ASI1>3.0.CO;2-Y

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, *3*, 160384. **DOI:** https://doi.org/10.1098/rsos.160384, **PMID:** 27703703, **PMCID:** PMC5043322

Stegehuis, C., Litvak, N., & Waltman, L. (2015). Predicting the long-term citation impact of recent publications. *Journal of Informetrics*, *9*, 642–657. **DOI:** https://doi.org/10.1016/j.joi.2015.06.005

Traag, V. A. (2020a). Replication data: Inferring the causal effect of journals on citations. *Zenodo*. https://zenodo.org/record/3582974

Traag, V. A. (2020b), Replication source code: Inferring the causal effect of journals on citations. *Zenodo*. https://zenodo.org/record/3583012

Traag, V. A., & Waltman, L. (2019). Systematic analysis of agreement between metrics and peer review in the UK REF. *Palgrave Communications*, *5*, 29. **DOI:** https://doi.org/10.1057/s41599-019-0233-x

Waltman, L., & Traag, V. A. (2020). Use of the journal impact factor for assessing individual articles need not be wrong. *F1000Research*, *9*. **DOI:** https://doi.org/10.12688/f1000research.23418.1

Wang, D., Song, C., & Barabási, A.-L. (2013). Quantifying long-term scientific impact. *Science*, *342*, 127–132. **DOI:** https://doi.org/10.1126/science.1237825, **PMID:** 24092745

Wouters, P., Sugimoto, C. R., Larivière, V., McVeigh, M. E., Pulverer, B., … Waltman, L. (2019). Rethinking impact factors: Better ways to judge a journal. *Nature*, *569*, 621–623. **DOI:** https://doi.org/10.1038/d41586-019-01643-3, **PMID:** 31138924

Zitt, M., Ramanana-Rahary, S., & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics*, *63*, 373–401. **DOI:** https://doi.org/10.1007/s11192-005-0218-y