

Scientific Journals' Twitter Accounts and Impact Factor: A Causal Analysis Based on a Synthetic Control

Andreas Nishikawa-Pacher^{1,*}

¹TU Wien Bibliothek, Vienna, Austria

Abstract

Does a journal's actively curated Twitter account cause a rise in Journal Impact Factor (JIF)? The present paper analyses this causal assumption by looking at the JIF trajectories of various journals that set up a Twitter account in 2015. Having identified one suitable journal (namely, *International Sociology*), a counterfactual unit was generated with the synthetic control method, a control unit that resembles the same journal but one that would not have registered a Twitter account in 2015. That control unit is synthesised via dozens of other journals that exhibit similar characteristics like the treated unit. The paper finds that the actual *International Sociology* enjoyed higher JIFs than its counterfactual control unit after 2015. The scale of this effect, however, is negligible and the finding is statistically insignificant. The main result of this paper, therefore, is that the null hypothesis claiming no-effect-of-Twitter-upon-JIF cannot be rejected.

Keywords

Causal inference, Impact Factor, Twitter, Synthetic Control, Altmetrics, Science Communication

1. Introduction

In disseminating research findings, scientific journals compete for attention and innovativeness that is (imperfectly) measured with citation counts. Greater visibility can be achieved through community engagement, such as when journals curate their own blogs, newsletters, podcasts, or social media channels [1]. As the dominant platform for informal discussions among academics is Twitter, a wider propagation of research is associated with journals setting up their own Twitter accounts [2]. Some scientometric studies have indeed shown that Tweets about scientific articles are correlated slightly but positively with higher article-level citation counts and therefore with greater journal-level, citation-based indicators – among them, most prominently, the journal impact factor, or JIF, which is an oft-used indicator of the average citation count an item published in a journal garners in the two years after its publication [3]. Given that researchers may find inspiration through serendipitous encounters on platforms such as Twitter [4], one may expect that a journal's dedicated Twitter account would cause an increase in its JIF in the

BIR 2023: 13th International Workshop on Bibliometric-enhanced Information Retrieval at ECIR 2023, April 2, 2023

*Corresponding author.

✉ andreas.pacher@tuwien.ac.at (A. Nishikawa-Pacher)

🌐 <https://oair.org> (A. Nishikawa-Pacher)

🆔 0000-0001-5149-6294 (A. Nishikawa-Pacher)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

years thereafter in a way that it would not have achieved without the Twitter account. But this has so far only been researched with *correlational* study designs; could one not examine this assumption as a *causal* hypothesis?

One difficulty with validating this assumption lies in the fast-growing landscape of scholarly publications, the increasing number of references per paper, and therefore an accompanying ‘inflation’ of the JIF [5, 6]. A generic journal’s JIF thereby increases almost by default, obtaining more citations simply because there are more journals with more publications and more references in their bibliographies. It is thus difficult to pin down whether an increase in JIF is due to the journal having set up a Twitter account, or whether the increase would have happened anyway without the Twitter account. The causal assumption is counterfactual and thus not directly observable; one cannot set up a randomized experiment in which a given journal would not have established a Twitter account in a year in which it actually did, and to calculate the treatment effect, i.e., the effect on the JIF of a journal setting up a Twitter account, by comparing it to the non-treatment effect, i.e., the JIF-trajectory of the same journal not setting up a Twitter account.

An alternative to a randomized controlled experiment would be to use quasi-experimental designs. For instance, one could use a statistical matching approach by comparing journals that bear similar values of selected covariates – such as by matching journals that have set up a Twitter account in a certain year with journals that have no Twitter accounts based on similar citation counts. However, scientific journals are so heterogeneous in terms of citation counts, discipline, topical niche, number of articles, number of authors, editorial prestige, use of special issues etc. that the approach of matching risks too much of an extrapolation [7].

Given that the untreated units are often unsuitable control groups for a journal, there is another quasi-experimental approach that would not be apt for the present question, namely the difference-in-differences approach [8, ch. 5.1-5.2]. It would observe the trajectory of tweeting journals’ JIFs with that of similar, but non-tweeting journals’ JIF across years, assuming that the two sets of units may not necessarily start at the same level of JIF, but that all other factors affecting each journal’s JIF develop in similar ways. This ‘parallel trend’-assumption, however, seems unrealistic as citation counts fluctuate too greatly – for instance, due to editorial board changes, due to the setting up of new journals in a given field, due to the size of the subdiscipline it is dedicated to, due to the salience of the issues the journal publishes about in society at large, due to open access and open peer review policies, and so on. Journals exhibit completely different trajectories with regards to their citation life cycle. The difference-in-differences design would nevertheless give all the untreated units an equal weight.

The approach of a synthetic control avoids some of the problems [9, ch. 10]. Given a unit of observation (such as a journal that established a Twitter account in 2015), the synthetic control method does not require a comparison unit (such as a journal without a Twitter account) that exhibits otherwise highly similar covariates, and it does not give the same weight across every single comparison unit. Instead, it combines multiple units of comparison and thereby generates an artificial, synthetic control unit which “is selected as the weighted average of all potential comparison units that best resembles the characteristics of the case of interest” [10, p. 496]. Given a journal with a Twitter account, the absence of an otherwise highly similar (but non-tweeting) journal does not pose a hindrance for analysing the causal hypothesis anymore. The synthetic control method takes X_1 , or specific parameters of the treated unit (the journal

with the Twitter account since 2015) before the intervention (before the setting up of a Twitter account), and compares them with X_0 , or with the same parameters of synthesised comparison units given specific weights of the variables, W . By taking the formula of $X_1 - X_0W$, the approach selects the synthetic control that carries the minimum difference, or W^* . The pre-intervention synthetic control should thus closely reproduce the values of the parameters of the treated unit. As a result, the difference in the post-intervention outcomes between the treated unit and the synthetic control, or $Y_1 - Y_0W^*$, allows one to quantify the causal effect of a given intervention (the setting up of a Twitter account) on the outcome of interest (the JIF). Another problem that is tackled by the synthetic control method is that given an extended period in a longitudinal dataset, the approach controls for unmeasured factors that may influence the outcome variable [10, p. 498]; the heterogeneity of the individual units does not pose a problem anymore, as it assumes that the treated unit and the synthetic control behave similarly and are therefore alike for a stretched period of pre-intervention time.

Meta-scientific studies have already used synthetic controls so as to model effects of Nobel prize bestowals on citations [11]; to gauge whether open data boost journal impacts [12]; and to assess how scientific mobility can affect one's scholarly output [13].

The present paper adds to this strand of causality-assessing meta-scientific research by tackling the question whether the creation of an (actively curated) Twitter account by a scholarly journal leads to a higher JIF in the years thereafter. The following section outlines the synthetic control method. The results section finds an almost negligible positive effect, but one that lacks statistical significance. The final section discusses possible reasons for the null result (such as the limited sample comprising only three disciplines and just one treatment journal) and points to possible future research venues.

2. Methods

2.1. Data Collection and Sample

The sample from which both the treated and the control units are to be taken consist of all journals that are indexed in Web of Science's 2021 Social Science Citation Index (SSCI) in the three categories Sociology, Political Science and International Relations. Their Twitter accounts were collected manually. The Twitter API was then used to find out the respective year of the Twitter accounts' establishment. Their respective two-year JIFs ranging from 2005 to 2020 were likewise collected for all these journals.

Possible journals to be chosen as the treated units were those that fulfilled three criteria. First, they had to be located in the middle quartiles (Q2 and Q3) of the respective ranking (ordered by JIF) in the 2021 version of the SSCI (the most recent one at the time of data collection). The journals in the top quartile (Q1) and in the bottom one (Q4) were not considered due to the risk that they could exhibit extreme values regarding the predictors (citation counts, number of Tweets about their articles and the JIF) so that no synthetic equivalent could be generated.

Second, the journals had to have established their Twitter accounts in the year 2015. That year was chosen because it ensures a long enough pre-intervention period spanning 15 years, thereby controlling for possible unobserved confounders, and a long enough post-intervention period so as to estimate an effect of the Twitter account registration.

The third criterion was that the Twitter account pertaining to the journal should be an actively curated one, indicated by an average of at least one Tweet a week since the inception of the Twitter account. As an approximation, therefore, only journals that had issued at least 364 Tweets so far (52 weeks multiplied by 7 years) would be used as treated units.

The donor pool, or the “reservoir of potential comparison units” [10, p. 497] out of which the control group would be synthesised, consisted of all the other journals indexed in those SSCI categories, given that they had not created any Twitter accounts yet, or only did so no earlier than 2021.

Using the International Standard Serial Number (ISSN) of the journals, the scholarly metadata service *CrossRef's* API was accessed to obtain the Digital Object Identifiers (DOIs) of all papers published in the sample journals between 2005 and 2021. For each DOI, the citation count was fetched using the *OpenCitations* API, and the Tweeter count via the *Altmetric* API.

A total of 20 journals matched the profile for the first two treatment criteria – that is, they were located in the middle quartiles of the JIF-based rankings and they created their Twitter accounts in 2015. However, 15 of these journals had missing values across 2005 to 2021 (such as due to their later indexing in Web of Science). To regain balanced data, or a longitudinal dataset with all values available for each unit in every year [10, p. 497], the 15 journals were removed. Only the following five journals were left: *Acta Sociologica*, *American Politics Research*, *British Journal of Sociology*, *International Sociology*, and *Political Studies*. However, the Twitter accounts of both *Acta Sociologica* and *American Politics Research* did not reach the threshold of 364 Tweets; the former one exhibited 318 Tweets, the latter just 21 Tweets in seven years. Thus, only three journals – *British Journal of Sociology*, *International Sociology*, and *Political Studies* – fulfilled all three criteria to count as potential treatment units.

The donor pool – journals in the same three SSCI categories that did not have Twitter accounts as of 2020 – originally harboured 93 journals, but only 43 of them had complete (i.e., balanced) data. The analysis thus remains with three journals in the possible treatment group and with 43 journals in the donor pool.

2.2. Data Analysis: Generating the Synthetic Control

To generate the synthetic control group, three predictors were used that were known to affect the JIF: first, the average count of Tweepers per article (as of December 2021); second, the average count of citations for each article (as of December 2021); and finally, the JIF as calculated by Web of Science (note that the predictors “may include pre-intervention values of the outcome variable” [14, p. 497]). With the R library *RSynth*, the use of the predictors was optimized from 2005 until the year of treatment in 2015, and a time plot was generated from 2005 to 2020.

Only one of the three potential treatments journals received an apt synthetic control, namely *International Sociology*, a journal located in Q3 that established its Twitter account in June 2015, having issued 3033 Tweets ever since. The two other journals’ synthetic controls had pre-intervention values that deviated too much from the actual journal’s predictors.

Table 1 shows the mean values of the predictors prior to 2015 for the journal and its synthetic equivalent, with the final column suggesting that the synthetic control provides a much better comparative unit than the mean values of the whole sample.

In line with the literature that found article-level citation counts to be better predictors of a

	<i>Int Soc</i> (actual)	<i>Int Soc</i> (synthetic)	Sample Mean
Average Tweeter count per article	0.545	0.545	1.270
Average citation count per article	8.400	8.400	23.029
Journal Impact Factor	0.715	0.715	0.944

Table 1

The average values of the predictors in the pre-intervention period for *International Sociology*, its synthetic control, and the whole sample.

	Weight
Average Tweeter count per article	0.389
Average citation count per article	0.566
Journal Impact Factor	0.036

Table 2

The synthetic weights of the three predictors used to generate the control group for *International Sociology*.

journal's JIF than Twitter activities [15], the weighting of the covariates put greater value on the citation-based measures than to the Tweeter counts, as visible in Table 2. To be more technical, the values in that table are determined by the R package based on an algorithmic search "for the set of weights that generate the best fitting convex combination of the control units. In other words, the predictor weight matrix V is chosen among all positive definite diagonal matrices such that MSPE is minimized for the pre-intervention period" [16].

A low pre-intervention mean squared prediction errors (MSPE) of 0.026 (in the unit of the target variable – that is, in terms of the JIF) likewise indicates that the synthetic control matches the actual journal closely in the pre-intervention period.

The control of *International Sociology* was synthesised with the help of the *Korean Journal of Defense Analysis* (31%), *Space Policy* (9.1%), the *American Journal of Sociology* (8.9%), *Politická Ekonomie* (8.2%), *Society* (6.7%), *Political Science* (5.2%), *Parliamentary Affairs* (2%) and 35 other journals whose respective weight remained below 2%.

The following section will present the synthetic control effect of the treatment (the setting up of the journal's Twitter account in 2015) upon the JIF, and infer the significance of the finding.

3. Results

3.1. The Effect of Twitter on the Impact Factor

The adoption of a Twitter account exhibits a minuscule positive effect on the increase of the JIF in the five years thereafter, albeit there are two caveats to this finding: first, the scale of this effect is negligible, and second, the finding is statistically insignificant. In other words, to obtain an answer to the question whether an active Twitter account generally causes an

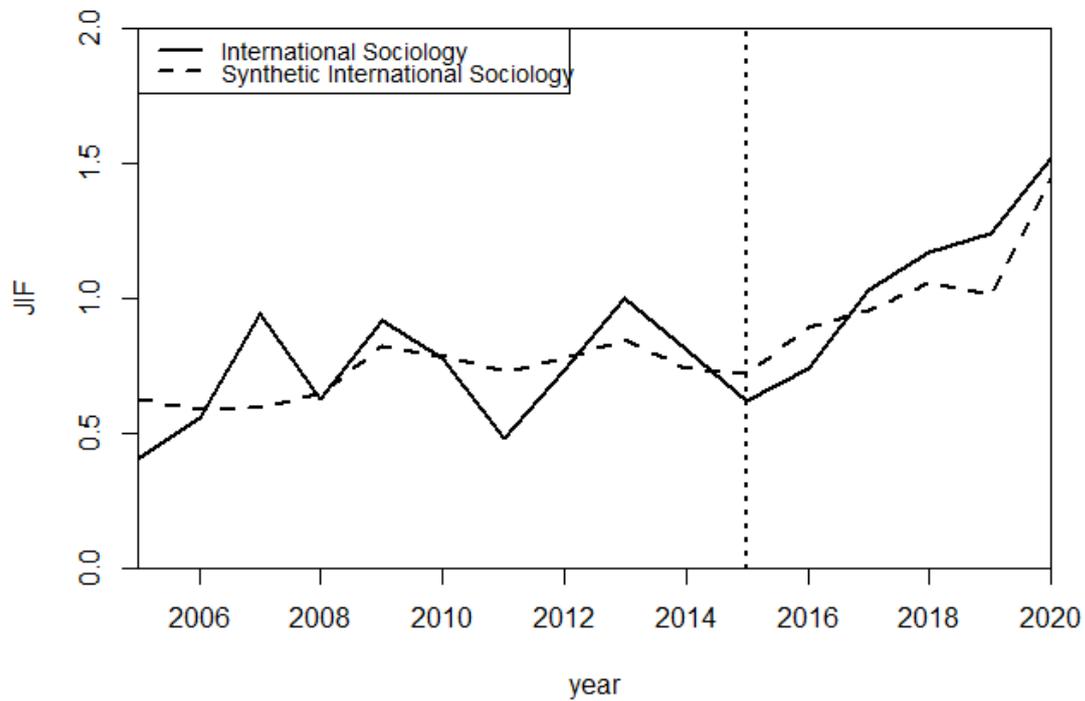


Figure 1: Path plot for the journal *International Sociology* and its synthetic control. The dotted vertical line in 2015 signifies the treatment (the establishment of the journal’s Twitter account).

increase in a journal’s JIF in the five years thereafter, a different dataset would be needed to find a generalisable response. While the synthetic control method may be suitable in answering the question at hand, it does not seem to be so when one uses a sample of just the middle-ranking journals in Sociology, Political Science and International Relations whose Twitter accounts were established in 2015.

One may nevertheless have a look at the results based on the data at hand. The path plot in Figure 1 shows that *International Sociology* enjoyed higher JIFs than the control group after 2015 (in the post-intervention period) than the counterfactual journal (the synthetic control) that did not establish a Twitter account, with the exception of the first post-treatment year (in 2016) where the journal underperformed relatively to its synthetic control.

Figure 1 also shows that there is a certain pre-intervention divergence between the two units. While the mean values of the predictors exhibited a close fit when seen across the whole pre-intervention period (see Table 1 above), the year-by-year JIFs deviate from each other with a mean annual gap of 0.07 (median: 0.08) prior to the intervention in 2015. A merely graphical analysis would thus indicate that the fit between the treated unit and the synthetic control was not as close as previously suggested by Table 1, unless one resorts to mean values summarised

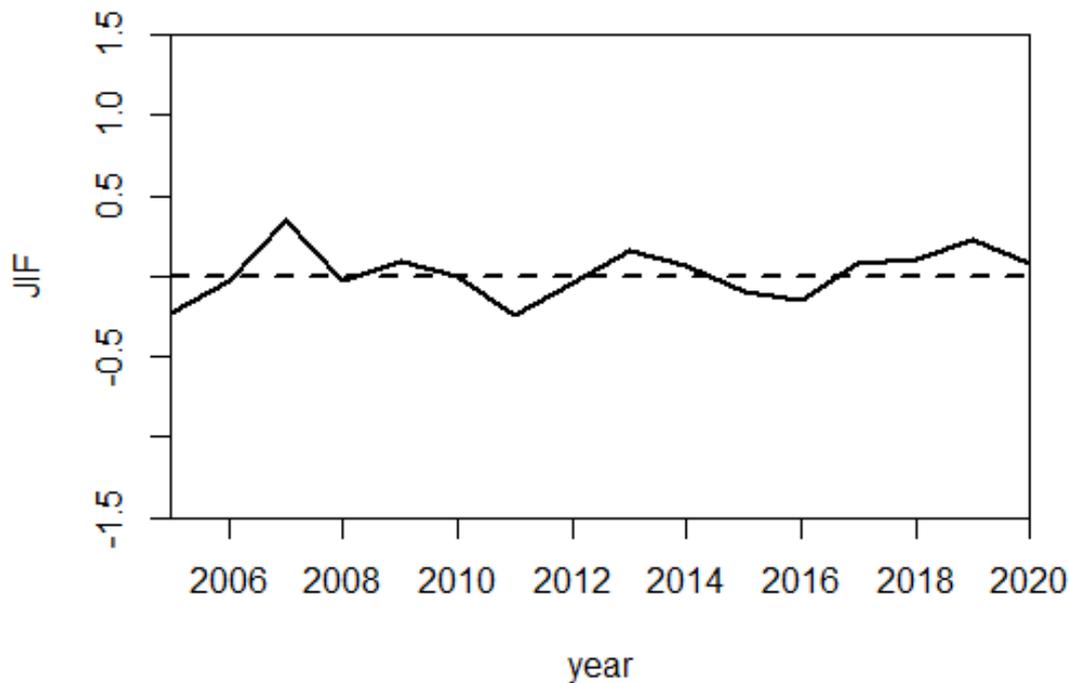


Figure 2: Gap plot showing the deviations between the JIF of the treated unit (the journal *International Sociology*) with its synthetic control's.

across the whole pre-intervention period.

Assuming nevertheless that the pre-intervention synthetic control reconstructs the actual *International Sociology* closely, one may hold that the effect of the establishment of an actively curated Twitter account in 2015 upon the JIF in the five years thereafter is rather minuscule. After the treatment in 2015, the average gap in JIF in a year was reduced to a mere 0.01 (median: 0.01); see Figure 2 for a gap plot showing how *International Sociology*'s JIF differed from its synthetic control year by year. A clear causal effect would have shown how “the two lines diverge substantially” [10, p. 503], but this is not the case here.

3.2. Placebo Tests: The Significance of the Effect

What is the statistical significance of this minuscule effect? The usual inferential techniques (such as the *t*-test etc.) are suitable for large samples, but not necessarily for comparative cases with a low number of units as in the present case where only one unit received the treatment. Inference in the synthetic control method is thus exercised via so-called placebo tests [14, 17]. These placebo tests counterfactually assign the treatment in the same period to each untreated

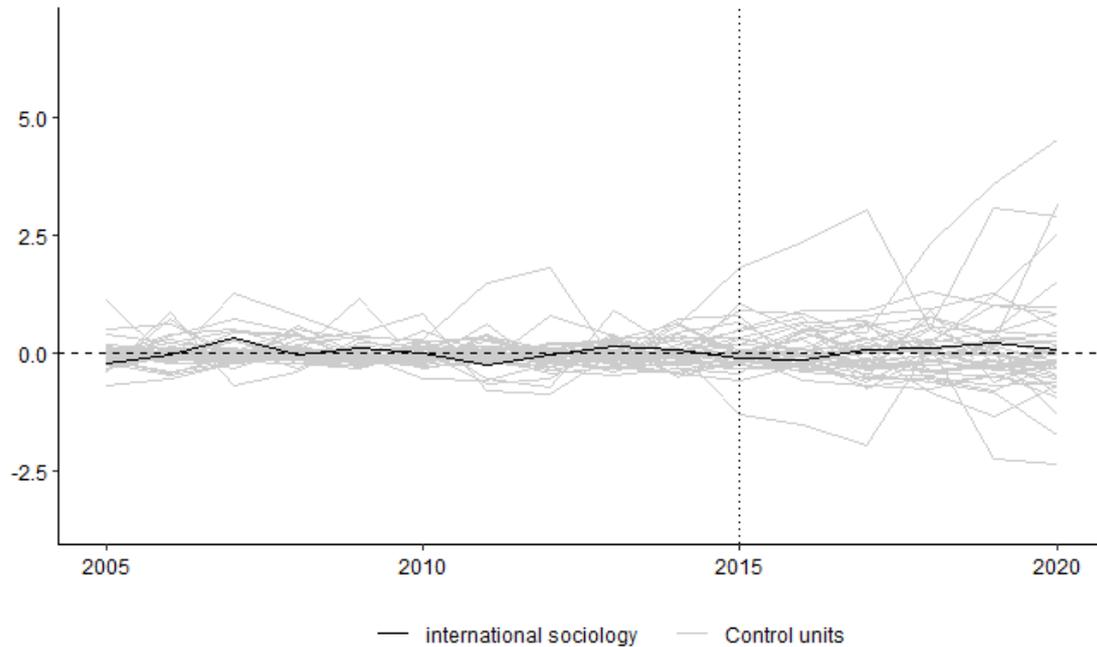


Figure 3: The result of the placebo tests.

unit, with every one of them obtaining their own synthetic controls (excluding the actually treated unit). And “[i]f the distribution of placebo effects yields many effects as large as the main estimate, then it is likely that the estimated effect was observed by chance” [18, p. 836]; that is, “the graphical analysis would suggest that the treatment effect is different from zero if t_{post}^{obs} ‘unusually’ large relative to the distribution of [the placebos]” [17, p. 8].

The placebo tests in Figure 3, however, does not show any unusual treatment effect. On the contrary, the treated unit’s line is positioned at the very center of the distribution of placebos. A large number of other lines exhibit more pronounced trajectories – even though “if another [journal] is chosen at random, identified as treated, and the synthetic control is applied, no effect should be visible for these” [19, p. 296]. Figure 3 seems to suggest the opposite, as if not-having-established-a-Twitter-account leads to more volatile JIF changes than the treatment.

Note, however, that “[t]he placebo effects may be quite large if those units were not matched well in the pretreatment period” [18, p. 837]. It is also possible that the post-treatment period was chosen to be too brief for a useful analysis. As a further means of inference, therefore, one may go beyond a graphical analysis and look at the distribution of the ratios of the MSPE in the post- and the pre-intervention period. If the intervention of the Twitter account registration was significant with regards to its effect upon the journal’s JIF, then *International Sociology*’s ratio would be much greater than the ratios of the placebos.

Figure 4 lists the post/pre-intervention MSPE ratios for the treated journal and the placebos. Again, *International Sociology* does not stand out at all.

The p -value is 0.91, indicating that even a random assignment of the intervention would

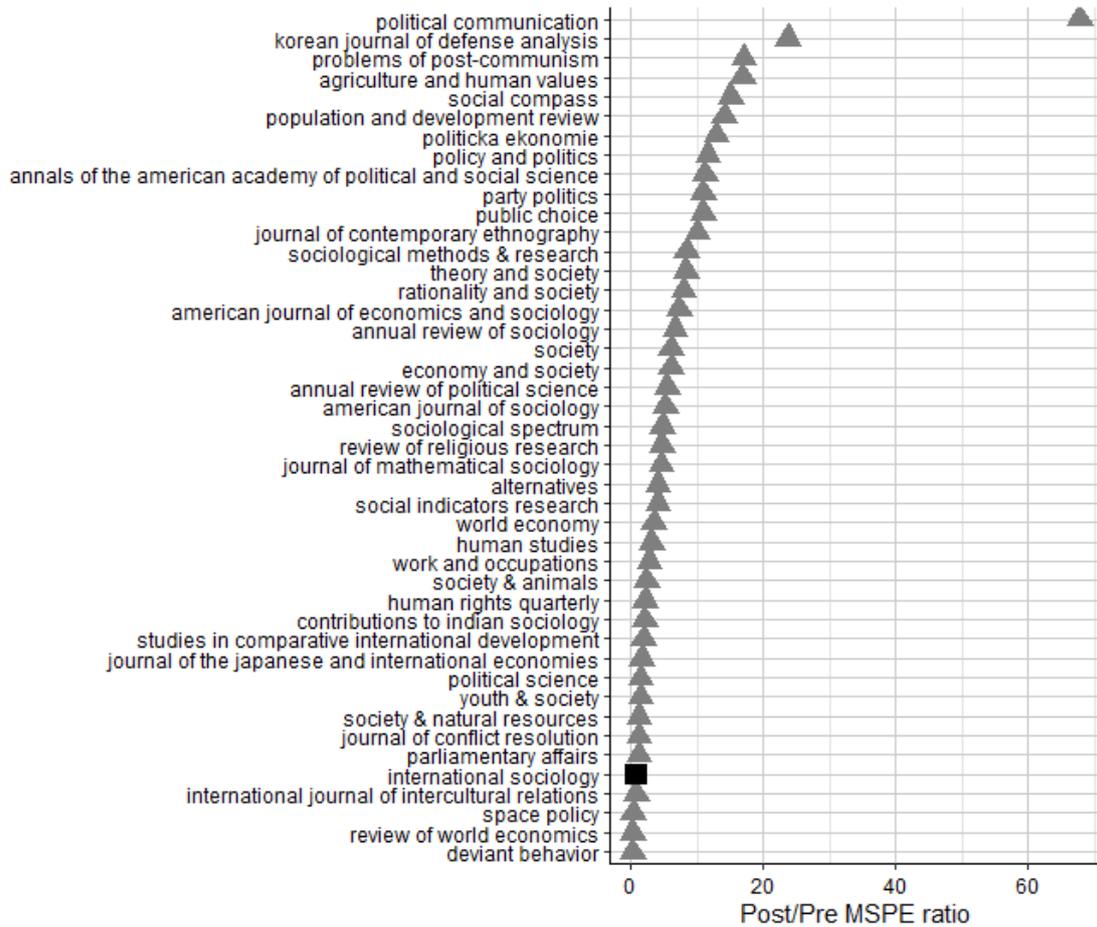


Figure 4: The ratio of mean squared prediction errors between the post- and the pre-intervention period for all placebo tests.

render it highly probable that one obtained the same post/pre-intervention MSPE ratio as *International Sociology's* [10, p. 500]. In other words, the null hypothesis, or the assumption that the establishment of an actively curated Twitter account had absolutely no effect upon the JIF at all, still holds.

4. Discussion

The effect of a scholarly journal setting up an active Twitter account amounts to a minuscule increase in its JIF (of an average of just 0.01 a year). However, this effect is statistically insignificant; in other words, despite having found a negligible effect in a one-unit sample pertaining to the journal *International Sociology*, the general causal relation between a Twitter account and the JIF cannot be determined with sufficient clarity for a broader population of scholarly journals with the method used in this paper – that is, not when one uses the synthetic control

method with regards to middle-ranked Sociology (and related) journals, not when one places the intervention in 2015, and not when one generates the synthetic control group via journals of the same categories with the three predictors of an average citation count per publication, an average Tweeter count per publication, and the JIF. The main result of this paper, therefore, is that the null hypothesis claiming no-effect-of-Twitter-upon-JIF cannot be rejected.

Perhaps one reason for the failure to identify a statistically significant causal effect is that the (2-year) JIF itself is immensely volatile. Reasons behind citations are complex and manifold, and chance encounters of serendipity push the scientific system to such unpredictable bibliometric links that a statistical analysis with a limited number of clear-cut quantitative parameters cannot capture the dynamics sufficiently. In addition, as Twitter is a relatively new tool for scholarly journals, perhaps the pre- and post-intervention periods for generating synthetic controls and for analyzing their effects were too brief for a meaningful estimation. Finally, drawing from Q2 and Q3 journals from just three social scientific disciplines may have been too narrow of a sample, indicated by the fact that only one single suitable unit with an apt control group was identified; a larger sample should be considered, taking into account discipline-specific differences with regards to a community's use of Twitter. Even if the question whether a journal's community outreach causally affects its scientific impact may be worthwhile to investigate, a limited data sample hinders the generalizability of the result found.

Further research could thus use a different sample of journals, one that could cover all kinds of scientific disciplines. The dataset of Twitter accounts from almost 3500 Web of Science-indexed journals may offer a good starting point [20], heightening the probability that multiple treatment units with suitable control groups can be found. In addition, the focal year of the Twitter account registration could be located prior to 2015 so as to allow for a more extended observation of the treatment effects. Finally, the predictors might draw from less volatile variables than the one used here; perhaps citation-based metrics spanning more than two years (like the five-year rather than two-year JIF, or median citation counts that span four years) may offer suitable parameters.

References

- [1] N. Erskine, S. Hendricks, The Use of Twitter by Medical Journals: Systematic Review of the Literature, *Journal of Medical Internet Research* 23 (2017) e26378. doi:10.2196/26378.
- [2] J. L. Ortega, The presence of academic journals on Twitter and its relationship with dissemination (tweets) and research impact (citations), *Aslib Journal of Information Management* 69 (2017) 674–687. doi:10.1108/AJIM-02-2017-0055.
- [3] H. Hughes, A. Hughes, C. G. Murphy, The Use of Twitter by Trauma and Orthopaedic Surgery Journals: Twitter Activity, Impact Factor, and Alternative Metrics, *Cureus* (2018). doi:10.7759/cureus.1931.
- [4] R. Ferreira Araujo, Communities of attention networks: introducing qualitative and conversational perspectives for altmetrics, *Scientometrics* 124 (2020) 1793–1809. doi:10.1007/s11192-020-03566-7.
- [5] B. M. Althouse, J. D. West, C. T. Bergstrom, T. Bergstrom, Differences in impact factor

- across fields and over time, *Journal of the American Society for Information Science and Technology* 60 (2009) 27–34. doi:10.1002/asi.20936.
- [6] B. D. Neff, J. D. Olden, Not So Fast: Inflation in Impact Factors Contributes to Apparent Improvements in Journal Quality, *BioScience* 60 (2010) 455–459. doi:10.1525/bio.2010.60.6.9.
- [7] M. Kellogg, M. Mogstad, G. A. Pouliot, A. Torgovitsky, Combining Matching and Synthetic Control to Tradeoff Biases From Extrapolation and Interpolation, *Journal of the American Statistical Association* 116 (2021) 1804–1816. doi:10.1080/01621459.2021.1979562.
- [8] J. D. Angrist, J.-S. Pischke, *Mostly Harmless Econometrics*, Princeton University Press, Princeton, New Jersey, 2009.
- [9] S. Cunningham, *Causal Inference*, Yale University Press, New Haven, Connecticut, 2021.
- [10] A. Abadie, A. Diamond, J. Hainmueller, Comparative Politics and the Synthetic Control Method, *American Journal of Political Science* 59 (2015) 495–510. doi:10.1111/ajps.12116.
- [11] R. Farys, T. Wolbring, Matched control groups for modeling events in citation data: An illustration of nobel prize effects in citation networks, *Journal of the Association for Information Science and Technology* 68 (2017) 2201–2210. doi:10.1002/asi.23802.
- [12] L. Zhang, L. Ma, Does open data boost journal impact: evidence from Chinese economics, *Scientometrics* 126 (2021) 3393–3419. doi:10.1007/s11192-021-03897-z.
- [13] G. Aykac, The value of an overseas research trip, *Scientometrics* 126 (2021) 7097–7122. doi:10.1007/s11192-021-04052-4.
- [14] A. Abadie, A. Diamond, J. Hainmueller, Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program, *Journal of the American Statistical Association* 105 (2010) 493–505. doi:10.1198/jasa.2009.ap08746.
- [15] V. T. Warren, B. Patel, C. J. Boyd, Analyzing the relationship between Altmetric score and literature citations in the Implantology literature, *Clinical Implant Dentistry and Related Research* 22 (2020) 54–58. doi:10.1111/cid.12876.
- [16] E. Dunford, *tidysynth: A Tidy Implementation of the Synthetic Control Method*, 2021. URL: <https://CRAN.R-project.org/package=tidysynth>.
- [17] B. Ferman, C. Pinto, *Placebo Tests for Synthetic Controls*, Technical Report 78079, University Library of Munich, Germany, 2017. URL: <https://ideas.repec.org/p/pramprapa/78079.html>.
- [18] S. Galiani, B. Quistorff, The Synth_runner Package: Utilities to Automate Synthetic Control Estimation Using Synth, *The Stata Journal* 17 (2017) 834–849. doi:10.1177/1536867X1801700404.
- [19] A. Besuyen, T. Coupé, K. K. Das, Effectiveness of foreign exchange interventions: evidence from New Zealand, *New Zealand Economic Papers* 55 (2021) 289–309. doi:10.1080/00779954.2020.1871063.
- [20] A. Nishikawa-Pacher, *The Twitter accounts of scientific journals: a dataset* 36 (2023) 1. doi:10.1629/uksg.593.