DEPARTMENT OF ECONOMICS AND FINANCE

SCHOOL OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

Do Replications Receive Fewer Citations? A Counterfactual Approach

NOTE: This paper is a revision of WP No. 15/2024

Tom Coupé
Thomas Logchies
W. Robert Reed

WORKING PAPER

No. 18/2024

Department of Economics and Finance
UC Business School
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 18/2024

Do Replications Receive Fewer Citations? A Counterfactual Approach

Tom Coupé[1†]
Thomas Logchies[1]
W. Robert Reed[1]

December 2024

Abstract: There is a widespread belief that replication studies are less cited than original research. This study introduces three counterfactual approaches for measuring the citations that an author or a journal would have received had they produced a non-replication study. Two of the measures are designed to measure citation incentives from the perspective of authors. One measure focuses on the perspective of journals. We collect data on 428 replications in economics published between 1958 and 2021 and assess whether these are cited less frequently than their matched counterfactuals. We obtain a wide range of estimates. Our preferred estimates use the ratio of citations of a replication to the citations of its matched counterfactuals. Using this measure, we estimate citation penalties as large as 51% and citation benefits as great as 227%. Most replications receive fewer citations than their matched counterfactuals, but a sizable portion, and sometimes even a majority, receive more. Finally, there is some evidence that replications that do not fully support the original study have more favorable citation rates than those that confirm the original study. While our analysis does not produce an unambiguous answer to whether replications receive less citations than their counterfactuals, it does revise the widely held, one-sided view that replications receive fewer citations.

Keywords: Replications, Citations, Incentives, Academic Publishing

JEL Codes: A10, A14, B41, C80

Data Availability: All the data and code required to reproduce the analyses in this study are available at:
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PTEUCR .

[1] Department of Economics and Finance, University of Canterbury, NEW ZEALAND

† Corresponding author: Tom Coupé. Email: tom.coupe@canterbury.ac.nz

1. Introduction

There is widespread concern in many fields of science that there exists a "reproducibility crisis" (Ioannidis, 2005a, 2005b; Prinz et al., 2011; Begley & Ellis, 2012; Open Science Collaboration, 2015; Duvendack et al., 2015, Camerer et al., 2016, 2018; Errington et al., 2021; Van Noorden, 2023). Fuelling this concern is the failure of many key studies to have their findings successfully replicated. This failure to replicate has highlighted problematic research practices such as insufficient sample sizes, selective reporting of results, and high levels of flexibility in data analysis, which may lead to statistically significant but irreproducible results.

As a result of this state of affairs, there have been increased calls for replication studies across a wide range of scientific disciplines (Burman et al., 2010; Maxwell et al., 2015; Heirene, 2020; Köhler & Cortina, 2019; Dennis & Valacich, 2014). A major motivation in these efforts is to identify findings that are reliable. Beyond that, replications have also been proposed as a means to deter questionable research practices: The prospect of post-publication scrutiny gives researchers a strong pre-publication motivation to ensure the credibility and robustness of their findings (Duvendack et al., 2017).

Despite these calls, studies are rarely replicated. Only about 0.13% of studies in education are replicated (Makel & Plucker, 2014), 0.45% in criminology (Pridemore et al., 2018), 0.25% in second language research (Marsden et al., 2018), 1.07% in psychology (Makel et al., 2012), 0.41% in special education research (Lemons et al., 2016), and 0.1% in economics (Mueller-Langer et al., 2019). This scarcity is often attributed to a lack of incentives for both authors and journals, where the imperative for citations—which influence both the impact factor of journals and the career trajectories of researchers—plays a critical role (Wilhite & Fong, 2012; Krauss, 2007; Chang & McAleer, 2016; Seeber et al., 2019; Biagioli, 2016; Lockwood, 2020).

Given the link between citations and the incentives, it is important to know whether there exists a citation incentive or disincentive associated with replications vis-à-vis other kinds of academic research. As demonstrated in TABLE 1, there is a widespread perception that replications are cited less than "original research". The evidence for this appears to be based upon a comparison of the citations of replications with the citations of the papers they replicate.

Table 1
Papers that Claim a Replication Citation Penalty

Paper	***Quote***
Feigenbaum and Levy (1993, p. 218)	"Given the lower expected citation return to positive and negative replication studies versus original work, a citation maximizing researcher should produce replications only when they are of significantly lower cost (primarily in terms of time) to complete."
Koole and Lakens (2012, p611)	"Currently, replications receive few citations, because academics most often cite research that breaks new grounds. When replications are rarely cited, researchers have little to gain from conducting and publishing replication research."
Zimmerman (2015, p.3)	"Replications are less citable."
Duvendack et al (2015, p.170)	"Journal editors are concerned with per-page citations, and replications are thought to perform less well than original studies."
Byington and Felps (2017, p.154)	"As such, Hollenbeck advises authors that "straight replication, which is valuable but not urgent in the eyes of most editors… will not be viewed as generating new knowledge nor will it generate a large number of citations—which of course, is a problem for the editor" (2008: 23).""

However, this is not the correct comparison if one wants to compare the incentives of authors and journals to publish replications versus original research.[1] To do that, one needs to

[1] This comparison can be of interest, however, from an open science perspective as a study that cites the original but not the replication does not fully contextualize the original paper's findings. Some have therefore argued for a "co-citation requirement", suggesting that journals mandate papers citing the original study to also cite its replication (Coffman and Niederle, 2015; Koole and Lakens, 2012).

compare the citations of replications with the citations that researchers and journals would have received had they conducted/published non-replication research. Accordingly, we develop three counterfactual measures of citations for replications. We then calculate these measures for 428 replications that we have collected from economics.

Our analysis produces a range of outcomes: At one extreme, we find a "citation penalty" where replication studies are cited at approximately half the rate of non-replications. At the other extreme, we find a substantial "citation benefit" where replications are cited over three times as frequently as their non-replication counterparts. These mixed results challenge the one-sided view that replications are less cited than other academic research. By quantifying the actual citation impact of replication versus non-replication studies, this research supports researchers and journals in making informed decisions in allocating resources for replication efforts.

This paper proceeds as follows. Section 2 presents three counterfactual approaches for comparing the citations of replications with the citations of other research that researchers and journals could publish instead. Section 3 describes the data used for this study and confirms the widely held belief that replications receive far fewer citations than the papers they replicate. Section 4 presents the counterfactual measures for comparing citation rates that are relevant for authors. Section 5 does the same from the perspective of journals. Included in that section is an analysis of the relationship between the relative citation rate of replications and replication outcomes (confirmed, mixed/unclear, disconfirmed). Section 6 concludes.

2. How should one compare citations of replications and original papers?

This study compares the citations of replications with the citations of alternative research that would have been produced if the author had written a non-replication study, or if the journal had decided to publish a non-replication study. In other words, we want to identify the citations associated with the "counterfactual" had the author/journal decided to instead allocate their efforts towards a non-replication activity.

The following thought experiment will guide our construction of counterfactuals. Imagine an experiment involving many authors. Each writes one replication and then spends the same amount of time to write one original paper. They then simultaneously submit the papers to journals. After publication, the researcher tracks citations to the two types of papers and computes whether there was a difference in their rates of citation. While we cannot implement this experiment in reality, our counterfactual measures are designed to capture key elements of this thought experiment.

Our first counterfactual records the citations received by all papers that were published in the same year as the replication study and that cited the original study. By focusing on papers that cited the original study, we attempt to match the subject area of the constructed counterfactual with that of the replication study. This is important because different subject fields can have different citation rates (Lillquist & Green, 2010; Bornmann & Wohlrabe, 2019). We call this approach the "same original counterfactual". In comparing the citations of replications and their "same original counterfactuals", we note that the distributions of citations tend to be skewed. Accordingly, we report the median as well as the mean of the respective citation counts.

While the above approach tries to keep the subject area fixed, an alternative approach focuses on authors. It may be that the authors of "same original counterfactuals" are more (or less) established scholars than the authors of replication studies. This would affect any comparison of citations across the two sets of authors. To address this issue, we next focus on non-replication papers published in the same year by the same authors who published the replication studies. We call this the "same author counterfactual".

So far, we have focused on counterfactuals that mainly pertain to authors. We next look at the decision to publish a study from the perspective of journals. Assuming journals only care about citations, our third counterfactual compares the citations of replications to the citations of other papers published in the same journal issue as the replication paper. We call this the "same issue counterfactual"[2].

Our three "counterfactual" measures are designed to best represent what would have happened if an author had decided to write a non-replication study rather than a replication study, or if a journal had decided to publish a non-replication study rather than a replication study. Citations are an important component of the incentives facing both authors and journals. Our findings can help align citation perceptions with actual citation patterns that feed into authors' and journals' decisions to conduct or publish replications versus other types of academic research.

3. Data and Preliminary Analysis

The replications included in this study come from two websites that collect data on replications in economics: The Replication Network and ReplicationWiki (Höffler 2017). Since we are interested in analysing citations, we restrict the sample to replications published before 2022 so that we have at least 2.5 years of citations for each paper. Scopus identification numbers (EIDs) were obtained for each replication paper and each original paper. We were able to find Scopus EIDs for 428 original-replication pairs, and these form the basis of the analysis in this paper.

A Comparison of Citations for Replications and the Original Papers They Replicated. Our analysis begins by comparing the citation counts of replications with those of the original

[2] If a volume has several issues, we use the specific issue in which the replication was published. If there is only one volume and no issue, we use the volume.

papers they replicate—a comparison that has sometimes been misinterpreted as evidence that replications garner fewer citations than originals. TABLE 2 reports annual citations for the 428 replications in our sample and the original papers they replicated. Citations are tracked from the year the replication study was published. All citation counts are divided by the number of years since publication, so that the values in the table represent annual citation rates.

The mean value of annual citations for the replications is 3.19 per year. The corresponding number for the replicated originals is 35.11 per year. The median citation rate for replications is 1.38 per year compared to 13.69 citations per year for the originals. Collectively, replications are cited at approximately one-tenth the rate of the papers they replicated. One reason may be that researchers are generally unaware of the replications. Another reason may be due to the Matthew effect, where papers get cited more just because they are already cited a lot (Birkmeier and Wohlrabe, 2014).

TABLE 2
A Comparison of Citations of Replications and the Studies They Replicated

	Replication Citations (1)	Original Citations (2)	Ratio (3)
Mean	3.19	35.11	0.35
Median	1.38	13.69	0.11
Ratio < 1 (%)	----	----	97.4%
Min	0	0.02	0.00
Max	79.42	681	65.00
Observations	428	428	428

NOTE: The table compares citations of 428 replications in economics with the citations of the papers they replicated. Columns (1) and (2) report total citations counted from the time each replication was published to mid-year 2024. Column (3) reports the ratio of citations of the replication paper to citations of the replicated paper. "Ratio < 1 (%)" reports the percent of replications having fewer citations than the papers they replicated.

Unlike the first and second columns, which treat replications and originals as independent, the third column (*Ratio*) matches each replication with the paper it replicated. For each matched, replication-original pair, *Ratio* calculates the replication paper's citations divided by the citations of the paper it replicated.[3] This provides a more direct comparison of citations of papers that are closely related in terms of topic and subfield.[4]

Column (3) reports that, on average, a replication receives approximately one-third (0.35) the citations received by the paper it replicated. The median *Ratio* value is 0.11. In other words, approximately half of all replication papers receive less (more) than one-ninth as many citations as the original papers (counting citations from the time of the publication of the replication). "Ratio < 1 (%)" shows that 97.4% of the replications in our sample received fewer citations than the originals they replicated.

These findings are consistent with the general perception that replications receive fewer citations than other academic papers.[5] However, they are not particularly relevant to the incentives for researchers and journals to publish replications. The authors of replications are generally very different from the authors of the studies they replicate. Mueller-Langer et al. (2019) find that replicated papers are typically high-impact studies authored by renowned researchers and institutions. Therefore, their citation counts are unlikely to be the appropriate counterfactuals for authors deciding to write a replication, or journals deciding to publish one.

---

[3] Column (3) is not the ratio of Columns (1) and (2). Since the replication with average cites is not necessarily linked to the original with average cites, the average of the ratio is different from the ratio of the averages.

[4] For example, Choi and Zhao (2021) has received 4 citations since 2021. They replicated Carhart (1997) which received over 2724 citations since 2021. The ratio is thus 0.001(4/2724). At the other extreme, Mehlum et al. (2006) has been cited 1509 since 2006. They replicated Sachs and Warner (1997) which have been cited 502 times since 2006. The ratio for this replication-original pair is thus about 3.

[5] Rather than comparing citations of articles, one could also compare the journals in which replications and originals are published. In 44% of the cases, the replication is published in the same journal as the original. To measure the impact of a journal, we use two journal rank measures from Scopus, the Scimago Journal Rank and the Source-Normalized Impact per Paper. Higher numbers imply higher impact for both measures. Using the most recent journal impact values, we find that on average (median), the Scimago Journal Rank for the journals in which the replications were published is 6.38 (2.13) against 12.02 (8.34) for the journals in which the originals were published. For the Source-Normalized Impact per Paper, the numbers are 2.87 (1.7) for replications and 4.69 (4.12) for the originals. Hence, if not published in the same journals, the replications tend to be published in journals with a lower impact than the originals.

To do that, we need to compare the citations of replications with the appropriate "counterfactuals." This is what we do in the next section.

4. Results: Authors' Perspective

A Comparison of Citations for Replications and "Same Original Counterfactuals". FIGURE 1 presents the distribution of 395 citations for the replications and the "same original counterfactuals". The latter are the set of papers that were published in the same year as the replication study and that cited the original study. The idea is that the citations received by these papers represent the citations the author of the replication could have received if they had produced a non-replication study in the same subject area.

There are 395 replications rather than 428 because some of the matched originals were not cited by any other papers in the year of publication of the replication. As multiple papers often cited a given original, we take either the mean or median of the citations of these papers. FIGURE 1 shows the distribution of citations for the "same original counterfactuals" when we use the mean.

Compared to the distribution of citations for the replications, the distribution for the counterfactuals has both higher mean and higher median values, though smaller outliers. The latter is due to the fact that the figure calculates citations for counterfactuals by taking the mean citation over all papers in the matched set of counterfactuals. This serves to moderate the effect of individual outliers among the counterfactuals. As a robustness check, we winsorize outliers among the replications. Because this only marginally impacts our results, we present and discuss the unwinsorized results in the text and report the winsorized results in an Online Appendix.[6]

[6] Cf. TABLES 3A, 4A, and 6A in the Online Appendix. For example, the "Ratio < 1 (%)" values in Columns (4) and (5) of TABLE 3 are 71.9% and 50.6%. The corresponding winsorized values are 73% and 51%. In TABLE 4, the "Ratio < 1 (%)" values in Columns (4) and (5) are 55.1% and 49.9%, compared to the winsorized analogues

Following on FIGURE 1, TABLE 3 provides a more in-depth comparison of the citations of replications and their "same original counterfactuals". Column (1) reproduces the analysis of Column (1) of TABLE 2 for the smaller set of 395 replications. Columns (2) and (3) quantify citations from the "same original counterfactuals". The difference between Column (2) and Column (3) is that when more than one "same original counterfactual" is matched with a given replication, Column (2) uses the mean to represent their citations. Column (3) uses the median.

Based on Columns (2), we would conclude that there exists a citation penalty associated with producing a replication study versus a non-replication alternative. Mean annual citations for replications in Column (1) is 3.39. Mean annual citations for the "same original counterfactuals" in Column (2) is 4.3. The corresponding median values are 1.50 and 3.38, respectively.[7] Using these measures, we would calculate citation penalties of approximately 21% (3.39 versus 4.3) and 56% (1.50 versus 3.38).

As was evident from FIGURE 1, the distributions of citations are generally highly skewed, so using the average for each set of "same original counterfactuals" can be misleading. For this reason, Column (3) uses the median of the matched, "same original counterfactuals". If one is concerned with the distortionary impact of outliers, medians may provide a better indicator of the "typical" counterfactual for each replication.

of 57% and 51%. And in TABLE 6, the "Ratio < 1 (%)" values in Columns (4) and (5) are 73.8% and 58.4% compared to 77% and 60% in TABLE 6A.

[7] The fact that the mean of the matched counterfactuals produces fewer positive results for replications compared to the median of matched counterfactuals suggests that originals are more likely to have positive outliers. In other words, the typical original is not that different from the typical replication, but outliers of originals have a chance to be citation "hits". Replications are less likely to have big citation "hits".

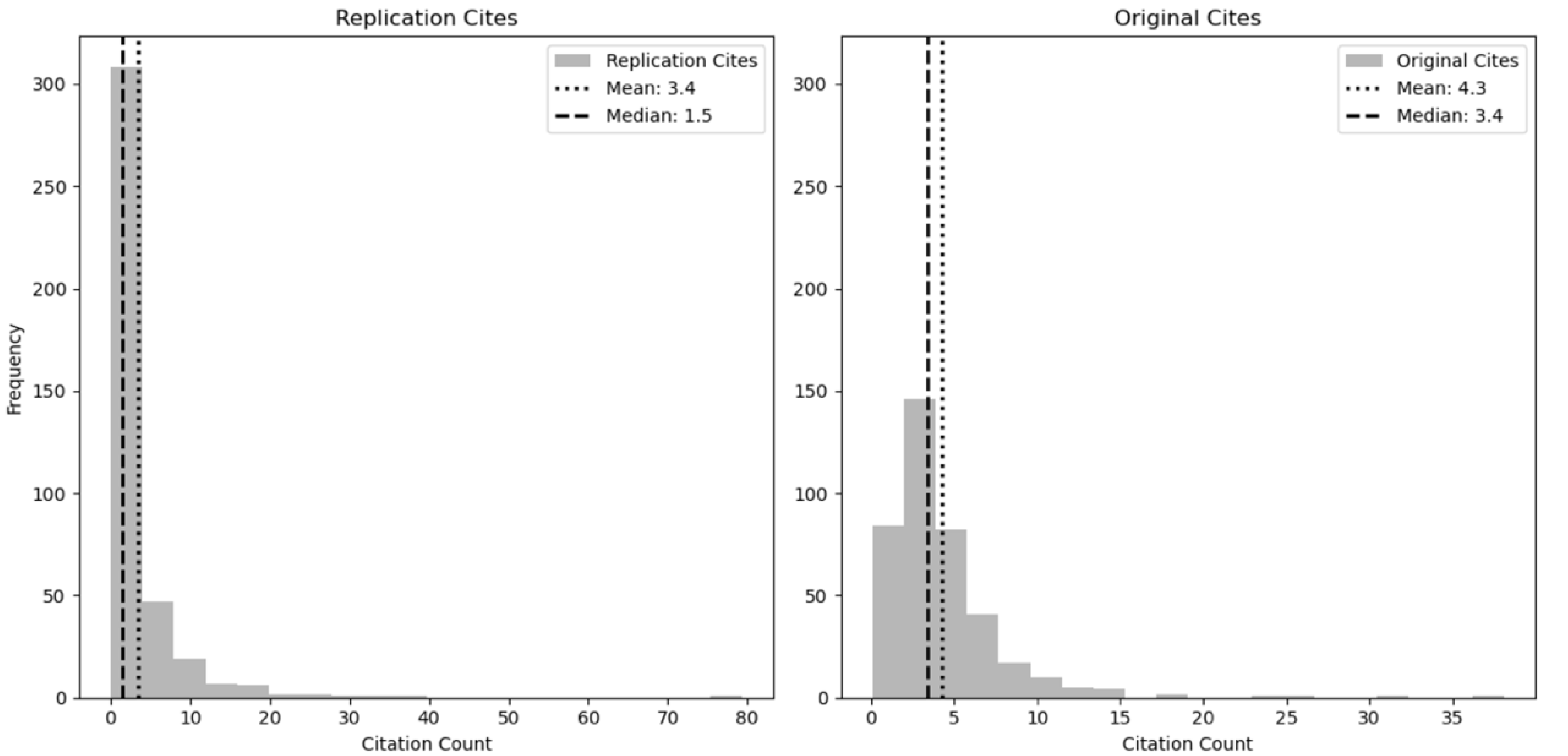

FIGURE 1
A Comparison of the Distribution of Citations: Replications versus Same Original Counterfactuals

TABLE 3

A Comparison of Citations: Same Original Counterfactual

	Citations			**Ratio of Citations**	
	Replication (1)	**Same Original (Mean) (2)**	**Same Original (Median) (3)**	**Same Original (Mean) (4)**	**Same Original (Median) (5)**
Mean	3.39	4.3	2.17	1.12	2.19
Median	1.50	3.38	1.71	0.52	0.95
Ratio < 1 (%)	----	----	----	71.9%	50.6%
Min	0.00	0.08	0.08	0.00	0.00
Max	79.42	38.13	24.47	23.09	35.93
Observations	395	395	395	395	395

NOTE: This table compares citations of replications with the citations of papers that cited the paper that was replicated in the same year the replication was published. Since each replication is matched to several counterfactuals, we explore both the distribution when using the mean counterfactual for each replication (Column 2), and the median counterfactual (Column 3). Column (4) calculates the ratio of the citations of the replication to the mean citations of its matched counterfactuals, and Column (5) does the same thing using the median of the matched counterfactuals. This sample excludes the 6 replications whose counterfactual had zero citations and the 27 replications who had original papers that were not cited by other papers in the year of the publication of the replication. "Ratio < 1 (%)" reports the percent of replications having fewer citations than the papers they were matched with.

Using the median-based counterfactual citations from Column (3) changes the picture. If one now compares the mean of the counterfactual citations in Column (3) to the mean of the replication citations in Column (1), there is now a citation benefit to producing replications. The average annual citation count for replications is more than 50% more than the average citation count for the "same original counterfactuals" (3.39 versus 2.17).

If one compares the median of the distributions of citations from Columns (1) and (3), one returns to the conclusion that replications experience a citation penalty. However, the size of the penalty is much smaller than the values based on Column (2) counterfactuals. Comparing median values in Columns (1) and (3) results in a citation penalty of only 12% (1.50 versus 1.71), versus the citation penalty of 56% based on Column (2).

A shortcoming of our analysis so far is that the distributions of citations for the replications and their counterfactuals are independent. That is, while the 395 sets of counterfactuals are constructed based on the original papers the replications replicated, no effort is made to match any given replication with the counterfactuals based on its original. Columns (4) and (5) address that shortcoming.

Each replication is associated with its matched with the counterfactuals based on its original and a ratio is constructed where the numerator is the citations of the replications, and the denominator is either the mean or median citations of the matched counterfactuals. A ratio smaller than 1 indicates that a given replication receives fewer citations than its corresponding counterfactuals. Column (4) uses the mean of the set of matched counterfactuals to represent the citations the author could have received by producing a non-replication study. Column (5) uses the median of the set of matched counterfactuals.

The first row of Columns (4) and (5) reports the average of the ratios of citations received by replications over the citations received by their matched counterfactuals, where the latter are measured by the mean and median of the set of matched counterfactuals, respectively.

By these measures, replications receive more citations than their "same original counterfactuals". We calculate citation benefits of producing a replication of 12% (Column 4) and 119% (Column 5).

If we instead compare the medians of the respective distributions of ratios, we calculate a citation penalty. By the second row of Columns (4) and (5), replications receive a citation penalty of 48% (Column 4) and 5% (Column 5). The third row ("Ratio < 1 (%)") indicates that a majority (71.9% and 50.6%, respectively) receive fewer citations than their matched counterfactuals.

Note that Columns (4) and (5) sometimes presented a very different pictures of relative citation rates than simple comparisons of Columns (2) and (3) with Column (1). The difference arises because those previous calculations did not directly match replications with their corresponding counterfactuals. For this reason, our preferred measures of relative citations will use the *Ratios* represented by Columns (4) and (5).

To summarize up to this point, our analysis of whether authors receive a citation penalty for producing replications has produced mixed results. Using the preferred estimates of Columns (4) and (5). our calculations range from a citation penalty for replications of 48% (i.e., replications receive 48% fewer citations than their "same original counterfactuals"), to a citation benefit of 119% (replications receive over twice as many citations than their counterfactuals). While most replications receive fewer citations than their matched counterfactuals, a substantial minority receive as many or more (28.1% and 49.4%).

<u>A Comparison of Citations for Replications and "Same Author Counterfactuals"</u>. TABLE 3 identified counterfactuals as all papers published in the same year as the replication study that cited the original paper that was replicated. By citing the same papers that the replications replicated, these papers identified themselves as having the same research interest(s) as the replication study. Thus, the citations they receive are indicative of what the

author of the replication might have received if they had written a non-replication paper on the same subject.

A shortcoming of this approach is that it fails to hold constant the citation productivity of authors. That is, perhaps the authors of the counterfactuals were generally more or less citation productive than the authors of the replications. Accordingly, we next pursue an alternative strategy of identifying counterfactuals using non-replication studies published by the replication authors in the same year they published their replications ("same author counterfactuals"). We then track the citations of these non-replication papers.

FIGURE 2 provides a visual comparison of the two sets of citation distributions. The left panel displays the distribution of replication citations. The right panel does the same for the distribution of "same author counterfactuals". Where an author publishes multiple non-replication papers, the figure uses the mean of the corresponding citations.

A total of 497 replications are represented in the figures. We note that some replications were dropped because the author of the replication did not publish any non-replication papers in the same year.[8] On the other hand, replications with multiple co-authors are included more than once if the co-authors separately published non-replication papers in the same year. The two distributions are generally similar. The distribution of replication citations has a greater mean (4.0 versus 3.8), but a smaller median (1.6 versus 2.1). It also has larger outliers. This visual comparison is examined in greater detail in TABLE 4.

[8] One could argue that the replication authors who are included in this sample are likely to be the more prolific authors as they are more likely to have several papers in a given year. Indeed, to be included in this sample one needs both a replication and non-replication paper(s) in the same year. If there's a Matthew effect, their replications might have a bigger impact than the replications of less prolific authors [who are less likely to be included in the same-author sample]. The Online Appendix (cf. TABLE A) provides evidence that on average, included authors published more highly cited replications, but they also replicated more highly cited originals. Importantly, the median Ratio of citations of replications to replicated originals was very similar, 0.13 for the included sample and 0.11 for the not-included sample (cf. Columns 5 and 6 of TABLE A). Unfortunately, we cannot directly test for differences in same-author effects given the absence of counterfactuals for those authors who didn't have another publication in the year of replication.

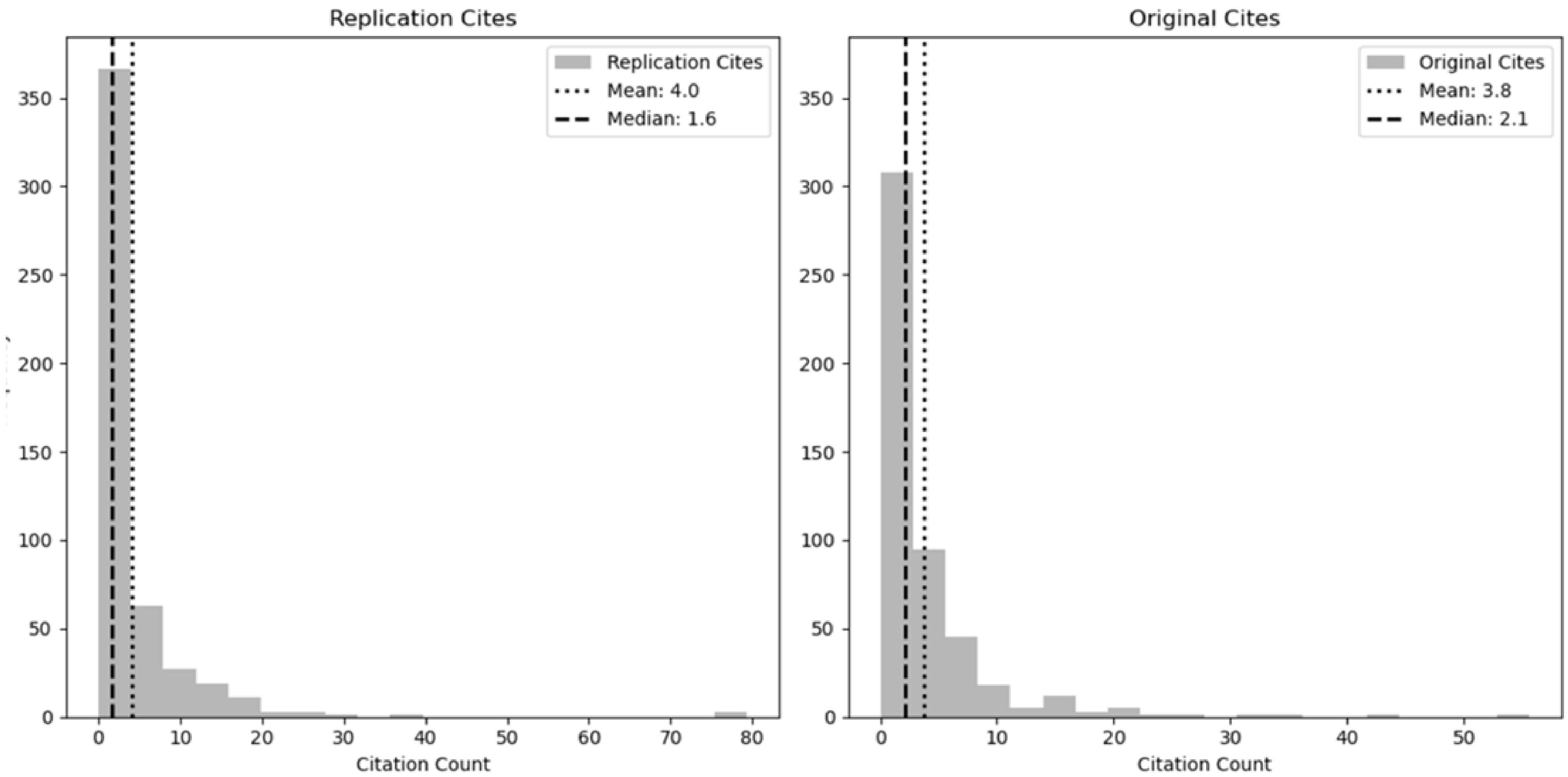

FIGURE 2
A Comparison of the Distribution of Citations: Replications versus Same Author Counterfactuals

TABLE 4
A Comparison of Citations: Same Author Counterfactual

	Citations			Ratio of Citations	
	Replication (1)	**Same Author (Mean) (2)**	**Same Author (Median) (3)**	**Same Author (Mean) (4)**	**Same Author (Median) (5)**
Mean	4.03	3.76	3.3	2.32	3.27
Median	1.64	2.14	1.83	0.78	1
Ratio < 1 (%)	----	----	----	55.1%	49.9%
Min	0.00	0.03	0.03	0.00	0.00
Max	79.42	55.57	55.57	101	127
Observations	497	497	497	497	497

NOTE: This table compares citations of replication papers with citations of non-replication papers by the same author published in the same year. As some authors published multiple non-replication papers, we report both mean and median citations of the counterfactuals (Columns 2 and 3). Column (4) calculates the ratio of the citations of the replication to the mean citation of its matched counterfactuals, and Column (5) does the same thing using the median of the matched counterfactuals. This sample excludes the 13 authors whose counterfactual had zero citations. Authors of replications who had no non-replication papers that were published in the year of the publication of the replication are excluded. "Ratio < 1 (%)" reports the percent of replications having fewer citations than the counterfactual papers they were matched with.

As in previous tables, Column (1) reports details about the distribution of replication citations. Columns (2) and (3) are similar to their analogues in TABLE 3 except that the set of counterfactuals consists of non-replication papers published by the replication author in the same year as the replication ("same author counterfactuals"). In line with TABLE 3, Column (2) uses the mean of matched counterfactuals' citations when there is more than one counterfactual for a given replication. Column (3) uses the median. As most replication authors in our sample do not publish multiple non-replication studies in the same year, Columns (2) and (3) produce similar values.

A comparison of Column (1) with Columns (2) and (3) reveals that replications tend to receive roughly the same number of citations whether the study is a replication or original research. If we compare the means of the distributions (first row), replications receive more citations (4.03 versus 3.76 and 3.3). If we compare medians (second row), they receive fewer (1.64 versus 2.14 and 1.83). As noted before, these comparisons do not attempt to match replications to their respective counterfactuals. To do that, we use the ratio of replication citations to their matched, counterfactual citations. These are reported in Columns (4) and (5).

Interestingly, as in TABLE 3, when we compare the means of the two distributions of *Ratios*, replications are measured to have a citation advantage – a large citation advantage. The citation benefits are 132% (Column 4) and 227% (Column 5). When we compare the medians, we calculate a 22% citation penalty for replications Column (4) and no citation penalty according to Column (5). Approximately half of all replications (55.1% and 49.9%) receive fewer citations than the counterfactual papers with which they are matched (cf. "Ratio < 1 (%)").

Our analysis of citations from "same author counterfactuals" produces qualitatively similar results to those using "same original counterfactuals". In both cases, focusing on the preferred *Ratio* measures, we calculate a range of estimates, including both citation penalties and citation benefits. The estimates from the TABLE 3 analysis ("same original counterfactuals") range from a citation penalty for replications of 48% to a citation benefit of 119%. The estimates from the TABLE 4 analysis ("same author counterfactuals") range from a citation penalty of 22% to a citation benefit of 227%. In both tables, a substantial minority, and even a slight majority (Column 5, TABLE 4) receive more citations than their matched counterfactuals.

Tests of Equality of Citations. Lacking from the previous discussion is an analysis of the extent to which sampling error could be responsible for the observed differences in citations between replications and their counterfactuals. TABLE 5 collects our findings from previous tables and adds standard errors for the *Ratios* of replication citations to counterfactual citations. The column headings indicate the respective counterfactual approach and specific *Ratio* measure. A *Ratio* of 1 means replications and their counterfactuals receive an equal number of citations, so our significance tests will focus on whether a *Ratio* is significantly different from 1 (as opposed to the usual null hypothesis of 0). Given that citation counts tend to have outliers with extreme values, we analyse median (the even columns) as well as mean *Ratios* (the odd columns).

The standard errors come from regression analyses. For the mean, we obtain the robust standard error from an OLS regression of the respective *Ratio* on a constant. For the median, we use quantile regression (for the median quantile). The estimated coefficient of the constant term reflects the mean (median) *Ratio* and the corresponding standard error can be used to test whether that mean (median) is statistically different from 1.

TABLE 5
Test for Differences in Citations Between Replications and Their Counterfactuals: Ratios

	Same Original Counterfactuals (TABLE 3)				Same Author Counterfactuals (TABLE 4)			
	Column 4 – Mean (1)	Column 4 – Median (2)	Column 5 – Mean (3)	Column 5 – Median (4)	Column 4 – Mean (5)	Column 4 – Median (6)	Column 5 – Mean (7)	Column 5 – Median (8)
Ratio	1.12	0.52	2.19	0.95	2.32	0.78	3.27	1.00
Standard Error	0.12	0.04	0.21	0.07	0.27	0.06	0.44	0.07
$H_0\!: Ratio = 1$ **(t-stat)**	0.99	-13.06	5.79	-0.72	4.85	-3.90	5.18	-0.00
$H_0\!: Ratio = 1$ **(p-value)**	0.32	0	0	0.47	0	0	0	1.00
Observations	395	395	395	395	495	495	495	495

NOTE: This table tests for differences in citations between replications and their counterfactuals ("same original counterfactual" and "same author counterfactual"). The Ratios have been previously reported in TABLES 3 and 4. For example, the ratio "1.12" in Column (1) comes from the first row ("Mean") of Column (4) in TABLE 3. The ratio "0.50: comes from the second row "Median" of Column (4) in TABLE 3. The ratios "2.19" and "0.95" come from the "Mean" and "Median" rows of Column (5) in TABLE 3. In each case, we test whether the Ratio = 1, as Ratio = 1 indicates no difference in the citations received by replications and their matched counterfactuals. Ratios greater than 1 indicated that replications receive more citations than the counterfactuals.

Of the eight tests of hypotheses, five find significant differences (Columns 2, 3, 5, 6, and 7). In two cases (Columns 2 and 6), we find that replications have significantly fewer citations than their respective matched counterfactuals. In three cases (Columns 3, 5, and 7), we find that replications have significantly more citations than their matched counterfactuals. Overall, TABLE 5 does not provide much additional illumination concerning relative citation rates. Nevertheless, it was necessary for us to conduct the tests in order to address concerns about the role of sampling error in the previous results. For example, if all the significant differences had been one way, either all indicating a citation penalty or a citation benefit, that would have been useful in clarifying the previous results.

In conclusion, our results are mixed whether an author who publishes a replication should expect to receive more, or fewer citations compared to publishing an original study. The wide range of citation impacts observed—ranging from substantial penalties to substantial benefits—suggests that there is no single answer for all researchers. While our analyses usually, but not always, find that a majority of replications receive fewer citations than their matched counterfactuals, we always find that a substantial share of replications receive more citations. The results provide strong evidence that the prevailing belief that replications receive fewer citations than other academic research does not consistently hold true.

5. Results: Journals' Perspective

So far, we have focused on estimating the citation penalties/benefits for authors. We now study the incentives for journals to publish replications. To do that, we compare the citations of replications to the citations of other papers published in the same issue in which the replication was published ("same issue counterfactual"). In particular, we address the question, is there a citation penalty for journals publishing replications rather than non-replication papers?

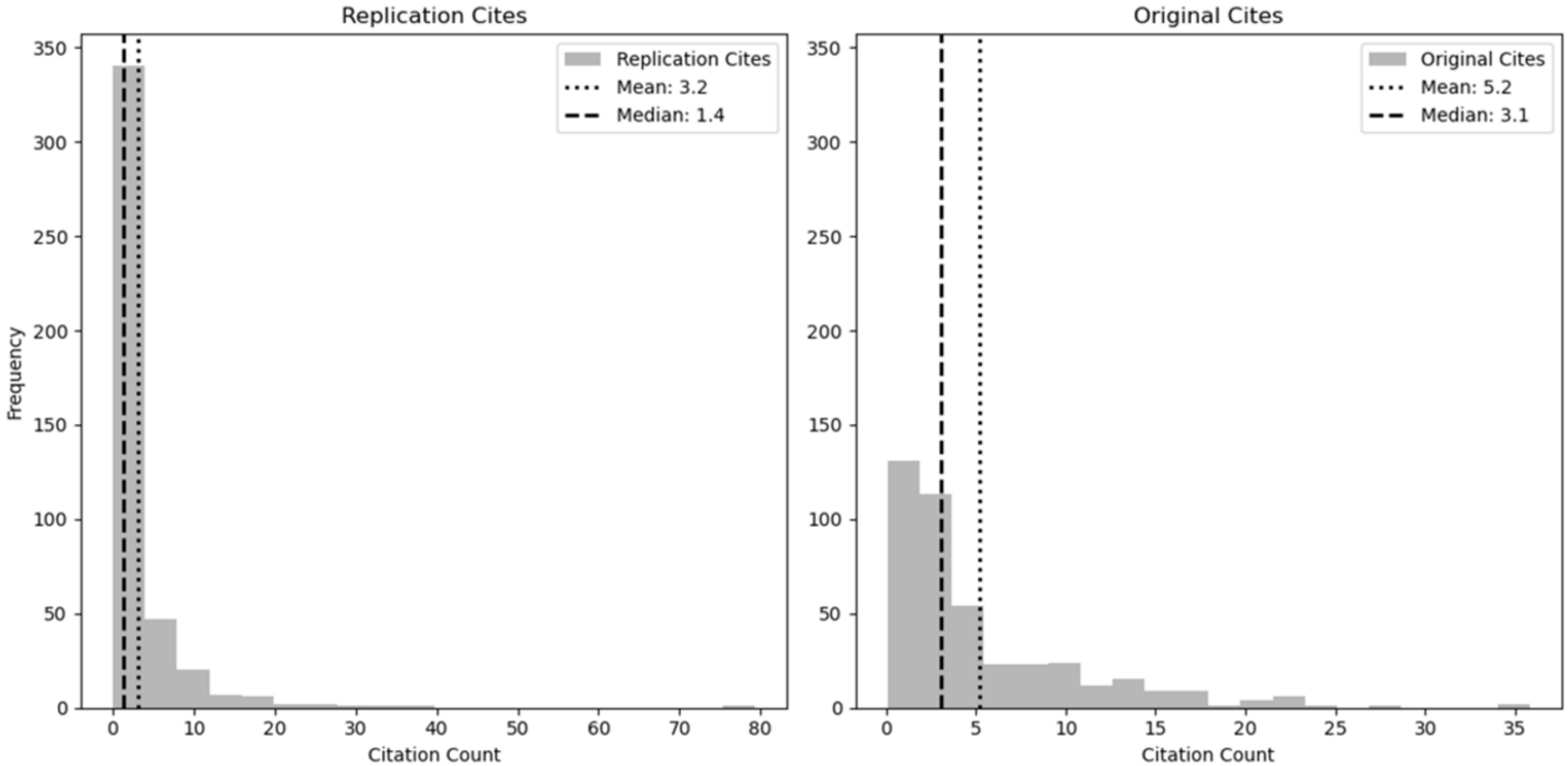

FIGURE 3

A Comparison of the Distribution of Citations: Replications versus Same Issue Counterfactuals

TABLE 6
A Comparison of Citations: Same Issue Counterfactual

	Citations			Ratio of Citations	
	Replication (1)	Same Issue (Mean) (2)	Same Issue (Median) (3)	Same Issue (Mean) (4)	Same Issue (Median) (5)
Mean	3.19	5.23	3.33	1.01	1.66
Median	1.38	3.1	1.89	0.49	0.75
Ratio < 1 (%)	----	----	----	73.8%	58.4%
Min	0.00	0.09	0.01	0.00	0.00
Max	79.42	35.94	24.15	29.26	47.9
Observations	428	428	428	428	428

NOTE: This table compares citations of replications with the citations of the papers that were published in the same issue as the replication. Since each replication is matched to several counterfactuals, we explore both the distribution when using the mean citations of the matched counterfactuals for each replication (Column 2), and the median citations of the matched counterfactual (Column 3). Column (4) calculates the ratio of citations of the replication to mean citation of the matched counterfactuals, and Column (5) does the same thing using the median of the matched counterfactuals. "Ratio < 1 (%)" reports the percent of replications having fewer citations than the counterfactual papers they were matched with.

A Comparison of Citations for Replications and "Same Issue Counterfactuals". As before, we first provide a visual comparison of the two sets of distributions in FIGURE 3. For the original studies drawn from the same issue as the replication study, we take the mean of their citations over the same period as the replication study. The distribution of citations for replications has both lower mean (3.2 versus 5.2) and lower median annual citations (1.4 versus 3.1). However, as we have seen before, the distribution of replications is characterized by larger outliers.

TABLE 6 repeats the now familiar exercise of conducting a more in-depth analysis of the two distributions. Column (1) reports the distribution of citations for replications, which in this case is identical to Column (1) of TABLE 2. Columns (2) and (3) quantify the citations of the counterfactuals using, respectively, the mean and median to represent the "typical" number of citations from a study published in the same journal issue as the replication.

A comparison of the first two rows of Columns (1) and (2) reproduces the results from FIGURE 3, showing a clear citation penalty for replications. We get a different picture if we use Column (3) to measure the citations of the counterfactuals. The means are very close (3.19 versus 3.33), and the medians, while indicating a citation penalty (1.38 versus 1.89) show a much smaller citation penalty than Column (2).

However, we know that these types of comparisons can be misleading because they do not match replications with their respective counterfactuals. Therefore we turn to Columns (4) and (5) to look at the ratio of replication citations to matched counterfactual citations, measured using the mean (Column 4) or the median (Column 5) of the citations from the same journal issue as the replication.

Using the mean of the distribution of *Ratios* (first row), we find citation benefits for replications: a slight 1% in Column (4) and a larger 66% in Column (5). However, when we compare the medians of the respective distributions of *Ratios*, we find citation penalties of 51%

(Column 4) and 25% (Column 5). 73.8% of the *Ratios* in Column (4), and 58.4% of the *Ratios* in Column (5) are less than 1, indicating that a majority of replications are cited less than their matched counterfactuals.

TABLE 7
Test for Differences in Citations Between Replications and Same Issue Counterfactuals: Ratios

	Same Issue Counterfactuals (TABLE 6)			
	Column 4 – Mean (1)	**Column 4 – Median (2)**	**Column 5 – Mean (3)**	**Column 5 – Median (4)**
Ratio	1.01	0.50	1.66	0.75
Standard Error	0.10	0.04	0.17	0.05
$\boldsymbol{H_0: Ratio = 1}$ **(t-stat)**	0.15	-13.82	3.84	-4.77
$\boldsymbol{H_0: Ratio = 1}$ **(p-value)**	0.88	0.00	0.00	0.00
Observations	428	428	428	428

NOTE: This table tests for differences in citations between replications and their counterfactuals ("same issue counterfactual"). The Ratios have been previously reported in TABLE 6. For example, the ratio "1.01" in Column (1) comes from the first row ("Mean") of Column (4) in TABLE 6. The ratio "0.50: comes from the second row "Median" of Column (4) in TABLE 6. The ratios "1.66" and "0.75" come from the "Mean" and "Median" rows of Column (5) in TABLE 6. In each case, we test whether the Ratio = 1, as Ratio = 1 indicates no difference in the citations received by replications and their matched counterfactuals. Ratios greater than 1 indicated that replications receive more citations than the matched counterfactuals.

Tests of hypotheses. As before, we investigate to what extent these differences can be explained by sampling error. TABLE 7 tests the respective *Ratios* to see if they are significantly different from 1, indicating a significant difference in the ratio of citations for replications and their "same issue counterfactuals". Of the four *Ratios*, three are statistically significant (Columns 2, 3, and 4). However, the results are split again between where statistical significance indicates a citation benefit (Column 3), and where a citation penalty is indicated

(Columns 2 and 4). Thus the results from TABLE 7 provide little additional insight beyond TABLE 6.

In conclusion, when comparing citations for replications and non-replications from the perspective of journals, we once again obtain mixed results. Based on Columns (4) and (5) of TABLE 6, we find citation penalties as large as 50%, and citation benefits as large as 66%. Most replications receive fewer citations than papers published in the same journal issue, but a large minority receive more.[9] The next section explores whether these results differ across replication outcomes.

The Effect of Replication Outcomes. Up to this point we have ignored how the findings of a replication might affect their citations, and thus their attractiveness to journals. While researchers who replicate a paper do not know in advance whether their replication will confirm the original results, journals are able to see the outcome of a replication. If a journal thinks that replications that fail to confirm the original study are more likely to be cited, this could incentivize journals towards publishing negative replications. One way to investigate this type of selection bias is to separate out replications by outcome.

A team of research assistants led by one of the co-authors of this study classified each of the 428 replications in this study into three categories: confirmed, disconfirmed, and mixed/unclear. The team generally took the replicating author's own assessment of the outcome of their replication. For example, if a replicating author stated in the abstract or conclusion that their paper failed to confirm the original study, we did not second-guess their assessment. Each paper was independently reviewed multiple times by the research assistants,

[9] It should be noted that replications are generally shorter than original studies. Accordingly, we repeated the analyses of TABLES 6 and 7 using per-page citations, since journals may be interested in per-page citation counts along with per-article citations. The results are reported in the Online Appendix. These present a more favorable picture for replications. The "Ratio < 1 (%)" values were substantially lower (60.0% and 45.8% versus 73.8% and 58.4%; cf. TABLES 6B/Online Appendix and TABLE 6, respectively). The tests of significant differences also were more favorable to replications (cf. TABLE 7A/Online Appendix and TABLE 7).

often two or more times. The co-author leading the team also independently assessed each study.

There was general agreement among the team when it came to categorizing replications that confirmed the original study. However, distinguishing between Mixed/Unclear and Disconfirmed was more challenging. For example, if a study investigated the role of an intervention on the energy consumption of poor households in Brazil, and a study identifying itself as a replication extended the study to poor households in Mexico but obtained different results, did that disconfirm the original study? Or was it unclear how it should affect one's interpretation of the original study? This depended on the target population of the original study. Was the target population of the original study poor households in Brazil? Or poor households in a developing country in South America? Unfortunately, authors of the original studies rarely explicitly stated their intended target populations. In these cases, we did our best to infer this from the context of the article's research question. As a result, while the category of "Confirm" was clearly demarcated, the lines between "Mixed/Unclear" and "Disconfirm" were not so clear.

To investigate whether the outcomes of replications affected the citation incentives for journals to publish replications, we conducted regression analysis. The dependent variable was the ratio of replication citations to counterfactual citations. *Ratio* was measured either using the mean ("Same Issue (Mean)") or median ("Same Issue (Median)") as reported in Columns (4) and (5) of TABLE 6. For each of the two samples, we regressed *Ratio* on dummy variables for Mixed/Unclear and Disconfirmed, with the reference category being Confirmed. We conducted both OLS with robust standard errors and quantile regression to investigate the impact of replication outcome on the mean and median of the two distributions, respectively.

TABLE 8

A Comparison of Citations of Replications by Outcome of Replication

A. Same Issue (Mean): OLS	
Ratio = .725 + 0.411·Mixed/Unclear + 0.344·Disconfirmed (0.106) (0.188) (0.191)	
1) $H_0: B_{M/U} = 0$	$t = 2.18$; p-value = 0.03
2) $H_0: B_D = 0$	$t = 1.80$; p-value = 0.07
3) $H_0: B_{M/U} = B_D$	$t = 0.30$; p-value = 0.76
4) $H_0: B_{M/U} = B_D = 0$	$F = 3.08$; p-value = 0.05
B. Same Issue (Mean): Quantile Regression	
Ratio = 0.366 + 0.253·Mixed/Unclear + 0.121·Disconfirmed (0.075) (0.101) (0.089)	
1) $H_0: B_{M/U} = 0$	$t = 2.50$; p-value = 0.01
2) $H_0: B_D = 0$	$t = 1.37$; p-value = 0.17
3) $H_0: B_{M/U} = B_D$	$t = 1.61$; p-value = 0.11
4) $H_0: B_{M/U} = B_D = 0$	$F = 3.16$; p-value = 0.04
C. Same Issue (Median): OLS	
Ratio = 1.30 + 0.427·Mixed/Unclear + 0.464·Disconfirmed (0.278) (0.376) (0.392)	
1) $H_0: B_{M/U} = 0$	$t = 1.14$; p-value = 0.26
2) $H_0: B_D = 0$	$t = 1.18$; p-value = 0.24
3) $H_0: B_{M/U} = B_D$	$t = -0.1$; p-value = 0.92
4) $H_0: B_{M/U} = B_D = 0$	$F = 0.88$; p-value = 0.41
D. Same Issue (Median): Quantile Regression	
Ratio = 0.600 + 0.200·Mixed/Unclear + 0.165·Disconfirmed (0.114) (0.153) (0.134)	
1) $H_0: B_{M/U} = 0$	$t = 1.31$; p-value = 0.19
2) $H_0: B_D = 0$	$t = 1.23$; p-value = 0.22
3) $H_0: B_{M/U} = B_D$	$t = 0.28$; p-value = 0.78
4) $H_0: B_{M/U} = B_D = 0$	$F = 0.99$; p-value = 0.37

NOTE: The table reports the results of four regressions, with four corresponding tests of hypotheses. The dependent variable in each case is the ratio of replication citations to counterfactual citations. There are two samples -- "Same Issue (Mean)" and "Same Issue (Median)", corresponding to the data reported in Columns (4) and (5) of TABLE 6. For each of the two samples, we regressed Ratio on dummy variables for Mixed/Unclear and Disconfirmed, with the reference category being Confirmed. We conducted both OLS and quantile regression to investigate the impact of replication outcome on the mean and median of the two distributions, respectively.

TABLE 8 reports the results of our analysis. The top panel ("Panel A") uses the 428 *Ratio* values summarized in Column (4) of TABLE 6 and regresses them on dummy variables for Mixed/Unclear and Disconfirmed. It estimates a mean value for *Ratio* of 0.725 for replication studies that confirm the original study. From the perspective of the journals, this indicates that replications that confirm the original study have an overall citation penalty of approximately 27%. In contrast, replication studies that are not fully supportive of the original study, either because the results are mixed/unclear or because they directly fail to confirm the original study, have average *Ratios* of 1.136 (=0.725+0.411) and 1.069 (=0.725+0.344). This indicates citation benefits of 14% and 7% associated with publishing replications that are either mixed/unclear or that disconfirm the original study.

The four rows below the reported regression test the following hypotheses:

1) Do replications that are Mixed/Unclear have a different effect on citations compared to replications that confirm the original study ($H_0: B_{M/U} = 0$)?

2) Do replications that disconfirm the original study have a different effect on citations compared to replications that confirm the original study ($H_0: B_D = 0$)?

3) Do replications that are Mixed/Unclear have a different effect on citations compared to replications that disconfirm the original study ($H_0: B_{M/U} = B_D$)?

4) Does the outcome of replications have no impact on how they are cited relative to their counterfactuals ($H_0: B_{M/U} = B_D = 0$)?

In most cases, we cannot reject the null hypotheses of no differences. The exception is for the outcome Mixed/Unclear. We find a significant difference between Mixed/Outcome and

Confirmed ($H_0: B_{M/U} = 0$) with a p-value of 0.03. This is weakly confirmed when we test for differences across all three outcomes ($H_0: B_{M/U} = B_D = 0$) and marginally reject equality of citation effects for all three outcomes (p-value of 0.05).

Panel B uses the same sample of *Ratios* as Panel A ("Same Issue (Mean)"), but estimates the regression using quantile regression. Quantile regression is a robust alternative to OLS when the data are non-normal/skewed and/or characterized by outliers. Columns (4) and (5) of TABLE 6 indicate that both characteristics are evident in our data. The quantile regression estimates for Panel B indicate a consistent citation penalty across all three types of replication outcomes.

Replications that confirm the original study have an estimated median *Ratio* value of 0.366. The corresponding values for Mixed/Unclear and Disconfirmed are 0.619 and 0.487. The associated citation penalties are 64%, 38%, and 51%. These results are in line with the overall results reported in Column (4) of TABLE 6. As in Panel A, we find some evidence that the citation impact of Mixed/Unclear is different from the other two outcomes. Mixed/Unclear is statistically different from Confirmed with a p-value of 0.01 ($H_0: B_{M/U} = 0$). In addition, a test of equality in effects across all three replication outcomes marginally rejects the null hypothesis of no difference with a p-value of 0.04 ($H_0: B_{M/U} = B_D = 0$).

The next two panels present similar results, though with some differences. They use the 428 *Ratio* values summarized in Column (5) of TABLE 6. The OLS results in Panel C find that all three types of replication outcomes are associated with substantial citation benefits for replications compared to their matched counterfactuals. The corresponding *Ratio* values are 1.3, 1.727, and 1.764. However, we cannot reject the null hypothesis that any of the replication outcomes are statistically different from the others. The quantile regression from Panel D likewise reveals no significant differences across replication outcomes. However, in contrast to Panel C, it shows that all three replication outcomes are associated with citation penalties, with estimated *Ratios* of 0.600, 0.800, and 0.765.

Summarizing the results from TABLE 8, we consistently find that replications that do not fully support the original studies generally have larger citation benefits/smaller citation penalties than studies that confirm the original studies. However, the differences are not consistently statistically significant. These findings add to the mixed results from TABLES 6 and 7. They support the conclusion that there is no one-size-fits-all answer to the question of whether replications receive fewer citations. From the journals' perspective, most replications receive fewer citations than papers published in the same issue, but a large minority receive more. Further, the share receiving more citations may be higher for replications that do not unambiguously support the original study.

6. Conclusion

Do replications receive fewer citations than so-called "original" research? There is a common belief that they do. This paper makes two main points. First, the common belief rests on an inappropriate comparison. It stems from the huge difference in the number of citations between replications and the originals they replicated. We also find a large disparity in citations between replications and replicated originals. However, it would be wrong to generalize this disparity to all original studies. Instead, we argue in this paper that a proper comparison requires developing "counterfactuals" that represent the citations the author or journal could have received had they conducted/published a non-replication study. Second, when one does that, the answer is mixed.

Using a sample of 428 replications in economics published between 1958 and 2021, and applying three measures of counterfactuals – two from the perspective of authors and one from the perspective of journals – we obtain a wide range of estimates for whether replications incur a citation penalty or enjoy a citation benefit. Our preferred estimates use the ratio of citations of a replication to the citations of its matched counterfactuals. By this measure, we obtain citation penalties as large as 51% and citation benefits as large as 227%. Measured from when the replication was published, most replications receive fewer citations than their matched counterfactuals, but a sizable portion, and sometimes even a majority, receive more. There is some evidence that replications that wholly or in part fail to support the original study have more favorable relative citations than those that confirm the original study.

While our analysis does not produce an unambiguous answer to whether replications receive less citations than their counterfactuals, it does challenge the widely held, one-sided view that replications receive fewer citations. To the extent that expected citations affect the incentives of authors and journals to produce/publish replications, we hope that our findings will help authors and journals to better align their expectations with actual citation patterns and promote the use of replication results whenever original studies are cited.

REFERENCES

Begley, C.G. & Ellis, L.M. (March 2012). Drug development: Raise standards for preclinical cancer research. *Nature* (Comment article). 483 (7391): 531–533.

Biagioli, M. (2016). Watch out for cheats in citation game. *Nature*, *535*(7611), 201-201.

Birkmaier, D., & Wohlrabe, K. (2014). The Matthew effect in economics reconsidered. *Journal of Informetrics*, 8(4), 880-889.

Bornmann, L., & Wohlrabe, K. (2019). Normalisation of citation impact in economics. *Scientometrics*, 120(2), 841-884.

Burman, L. E., Reed, W. R., & Alm, J. (2010). A call for replication studies. *Public Finance Review*, 38(6), 787-793. https://doi.org/10.1177/1091142110385210

Byington, E. K. & W. Felps (2017). Solutions to the credibility crisis in management science. *Academy of Management Learning & Education,* 16(1): 142-162.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., ... & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433-1436.

Camerer CF, Dreber A, Holzmeister F, Ho TH, Huber J, Johannesson M, et al. (September 2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*. 2 (9): 637–644. doi:10.1038/s41562-018-0399-z. PMID 31346273. S2CID 52098703.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, *52*(1), 57-82.

Chang, C. and McAleer, M. (2016). Quality weighted citations versus total citations in the sciences and social sciences, with an application to finance and accounting. *Managerial Finance*, 42(4), 324-337. https://doi.org/10.1108/mf-12-2014-0315

Choi, J. J., & Zhao, K. (2021). Carhart (1997) mutual fund performance persistence disappears out of sample. *Critical Finance Review*, *10*(2), 263-70. http://dx.doi.org/10.1561/104.00000103

Coffman, L. C. & M. Niederle (2015). "Pre-analysis plans have limited upside, especially where replications are feasible." *Journal of Economic Perspectives,* 29(3): 81-98.

Dennis, A. R. & Valacich, J. S. (2014). A replication manifesto. *AIS Transactions on Replication Research,* 1, 1-4. https://doi.org/10.17705/1atrr.00001

Duvendack, M., Palmer-Jones, R. W., & Reed, W. R. (2015). Replications in economics: A progress report. *Econ Journal Watch*, *12*(2), 164-191.

Duvendack, M., Palmer-Jones, R., & Reed, W. R. (2017). What is meant by “replication” and why does it encounter resistance in economics?. *American Economic Review*, *107*(5), 46-51.

Errington, T. M., Mathur, M., Soderberg, C. K., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021). Investigating the replicability of preclinical cancer biology. *eLife, 10*, e71601.

Feigenbaum, S. & D. M. Levy (1993). The market for (ir)reproducible econometrics. *Social Epistemology,* 7(3): 215-232.

Heirene, R. (2020). A call for replications of addiction research: which studies should we replicate and what constitutes a 'successful' replication?. *Addiction Research & Theory*, 29(2), 89-97. https://doi.org/10.1080/16066359.2020.1751130

Höffler, J. H. (2017). "ReplicationWiki: Improving transparency in social sciences research." D-Lib Magazine 23(3): 1.
Ioannidis, J.P. (August 2005). "Why most published research findings are false". *PLOS Medicine*. 2 (8): e124. doi:10.1371/journal.pmed.0020124.

Ioannidis, J.P. (July 2005). "Contradicted and initially stronger effects in highly cited clinical research". *JAMA*. 294 (2): 218–228. doi:10.1001/jama.294.2.218.

Köhler, T. and Cortina, J. M. (2019). Play it again, Sam! An analysis of constructive replication in the organizational sciences. *Journal of Management*, 47(2), 488-518. https://doi.org/10.1177/0149206319843985

Koole, S. L. & D. Lakens (2012). Rewarding replications: A sure and simple way to improve psychological science. *Perspectives on Psychological Science,* 7(6): 608-614.

Krauss, J. (2007). Journal self-citation rates in ecological sciences. *Scientometrics* **73**, 79–89. https://doi.org/10.1007/s11192-007-1727-7

Lemons, C. J., King, S. A., Davidson, K. A., Berryessa, T. L., Gajjar, S. A., & Sacks, L. H. (2016). An inadvertent concurrent replication. *Remedial and Special Education*, 37(4), 213-222. https://doi.org/10.1177/0741932516631116

Lillquist, E., & Green, S. (2010). The discipline dependence of citation statistics. *Scientometrics*, *84*(3), 749-762.

Lockwood, M. (2020). Editorial: citation malpractice. *Proceedings of the Royal Society A*, *476*(2242), 20200746.

Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, *43*(6), 304-316.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur?. *Perspectives on Psychological Science*, *7*(6), 537-542.

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, *68*(2), 321-391.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? what does “failure to replicate” really mean?. American Psychologist, 70(6), 487-498. https://doi.org/10.1037/a0039400

Mehlum, H., Moene, K., & Torvik, R. (2006). Institutions and the resource curse. *The Economic Journal*, *116*(508), 1-20.

Mueller-Langer, F., Fecher, B., Harhoff, D., & Wagner, G. G. (2019). Replication studies in economics—How many and which papers are chosen for replication, and why?. *Research Policy*, *48*(1), 62-83.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Pridemore, W. A., Makel, M. C., & Plucker, J. A. (2018). Replication in criminology and the social sciences. *Annual Review of Criminology*, *1*(1), 19-38.

Prinz F, Schlange T, Asadullah K (August 2011). Believe it or not: how much can we rely on published data on potential drug targets?. *Nature Reviews/Drug Discovery*. 10 (9): 712. doi:10.1038/nrd3439-c1.

Sachs, J. D., & Warner, A. M. (1997). Sources of slow growth in African economies. *Journal of African Economies*, *6*(3), 335-376.

Seeber, M., Cattaneo, M., Meoli, M., & Malighetti, P. (2019). Self-citations as strategic response to the use of metrics for career decisions. *Research Policy*, 48(2), 478-491. https://doi.org/10.1016/j.respol.2017.12.004

Van Noorden R (July 2023). Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed?. *Nature*. 619 (7970): 454–458. doi:10.1038/d41586-023-02299-w.

Wilhite, A. W., & Fong, E. A. (2012). Coercive citation in academic publishing. *Science*, 335(6068), 542-543.

Zimmermann, C. (2015). On the need for a replication journal. FRB St. Louis Working Paper(2015-16).

ONLINE APPENDIX

TABLE 3A: Identical to TABLE 3 in the text except it winsorizes the citations of replications

TABLE 4A: Identical to TABLE 4 in the text except it winsorizes the citations of replications

TABLE 6A: Identical to TABLE 6 in the text except it winsorizes the citations of replications

TABLE A: A Comparison of Citations of Authors Who Published a Non-Replication in the Same Year as the Replication and Those Who Did Not

TABLE 6B: Identical to TABLE 6 in the text except it that it is based on citations per page

TABLE 7A: Identical to TABLE 7 in the text except it that it is based on citations per page

TABLE 3A
A Comparison of Citations: Same Original Counterfactuals (Winsorized Replication Citations)

	Citations			Ratio of Citations	
	Replication (1)	Same Original (Mean) (2)	Same Original (Median) (3)	Same Original (Mean) (4)	Same Original (Median) (5)
Mean	2.50	4.30	2.17	0.88	1.69
Median	1.50	3.38	1.71	0.51	0.92
Ratio < 1 (%)	----	----	----	73%	51%
Min	0	0.08	0.08	0	0
Max	7.93	38.13	24.47	21.50	21.50
Observations	395	395	395	395	395

NOTE: This table is identical to TABLE 3 in the text except it winsorizes the annual citations for replications. We winsorize values at Quartile 3 + (1.5 × the interquartile range).

TABLE 4A
A Comparison of Citations: Same Author Counterfactuals (Winsorized Replication Citations)

	Citations			Ratio of Citations	
	Replication (1)	**Same Author (Mean) (2)**	**Same Author (Median) (3)**	**Same Author (Mean) (4)**	**Same Author (Median) (5)**
Mean	2.93	3.76	3.30	2.05	2.82
Median	1.64	2.14	1.83	0.77	0.94
Ratio < 1 (%)	----	----	----	57%	51%
Min	0	0.03	0.03	0	0
Max	9.42	55.57	55.57	101	101
Observations	497	497	497	497	497

NOTE: This table is identical to TABLE 4 in the text except it winsorizes the annual citations for replications. We winsorize values at Quartile 3 + (1.5 × the interquartile range).

TABLE 6A
A Comparison of Citations: Same Issue Counterfactuals (Winsorized Replication Citations)

	Citations			Ratio of Citations	
	Replication (1)	**Same Author (Mean) (2)**	**Same Author (Median) (3)**	**Same Author (Mean) (4)**	**Same Author (Median) (5)**
Mean	2.33	5.23	3.33	0.83	1.37
Median	1.38	3.1	1.89	0.47	0.74
Ratio < 1 (%)	----	----	----	77%	60%
Min	0	0.09	0.01	0	0
Max	7.52	35.94	24.15	8.36	30.06
Observations	428	428	428	428	428

NOTE: This table is identical to TABLE 6 in the text except it winsorizes the annual citations for replications. We winsorize values at Quartile 3 + (1.5 × the interquartile range).

TABLE A
A Comparison of Citations of Authors Who Published a Non-Replication in the Same Year as the Replication and Those Who Did Not

	Replication Citations		Original Citations		Ratio	
	Published Non-Replication (1)	Did Not Publish Non-Replication (2)	Published Non-Replication (3)	Did Not Publish Non-Replication (4)	Published Non-Replication (5)	Did Not Publish Non-Replication (6)
Mean	3.86	2.25	33.34	24.26	0.22	0.79
Median	1.48	1.17	13.91	11.25	0.13	0.11
Ratio < 1 (%)	----	----	----	----	0.97	0.97
Min	0	0	0.3	0.02	0	0
Max	79.42	34.69	371	283	3.01	65
Observations	156	105	156	105	156	105

NOTE: This table is based on replication papers written by authors who were included in the "Same Authors Counterfactuals" or not written by authors included in the "Same Authors Counterfactuals". It compares citations of their replication paper, the originals they replicated, and the ratio of these.

TABLE 6B
A Comparison of Citations: Same Issue Counterfactuals (Per Page Citations)

	Citations			Ratio of Citations	
	Replication (1)	**Same Author (Mean) (2)**	**Same Author (Median) (3)**	**Same Author (Mean) (4)**	**Same Author (Median) (5)**
Mean	0.26	0.24	0.16	1.39	2.31
Median	0.12	0.17	0.11	0.77	1.11
Ratio < 1 (%)	----	----	----	60.0%	45.8%
Min	0	0.01	0	0	0
Max	3.97	2.24	1.04	35.7	54.2
Observations	417	417	417	417	417

NOTE: This table is similar to TABLE 6 in the text except it calculates per page citations, given that replications are typically shorter than original papers. For 11 papers we did not have page numbers for all counterfactuals.

TABLE 7
Test for Differences in Citations Between Replications and Same Issue Counterfactuals: Ratios (Per Page Citations)

	Same Issue Counterfactuals (TABLE 6B)			
	Column 4 – Mean (1)	**Column 4 – Median (2)**	**Column 5 – Mean (3)**	**Column 5 – Median (4)**
Ratio	1.39	0.77	2.31	1.11
Standard Error	0.12	0.05	0.21	0.07
$\boldsymbol{H_0: Ratio = 1}$ **(t-stat)**	3.34	-4.32	6.24	1.41
$\boldsymbol{H_0: Ratio = 1}$ **(p-value)**	0.00	0.00	0.00	0.16
Observations	417	417	417	417

NOTE: This table is similar to TABLE 7 in the text except that it is based on per page citations, given that replications are typically shorter than original papers. For 11 papers we did not have page numbers for all counterfactuals.