European Commission

JRC TECHNICAL REPORTS

JRC Digital Economy Working Paper 2018-01

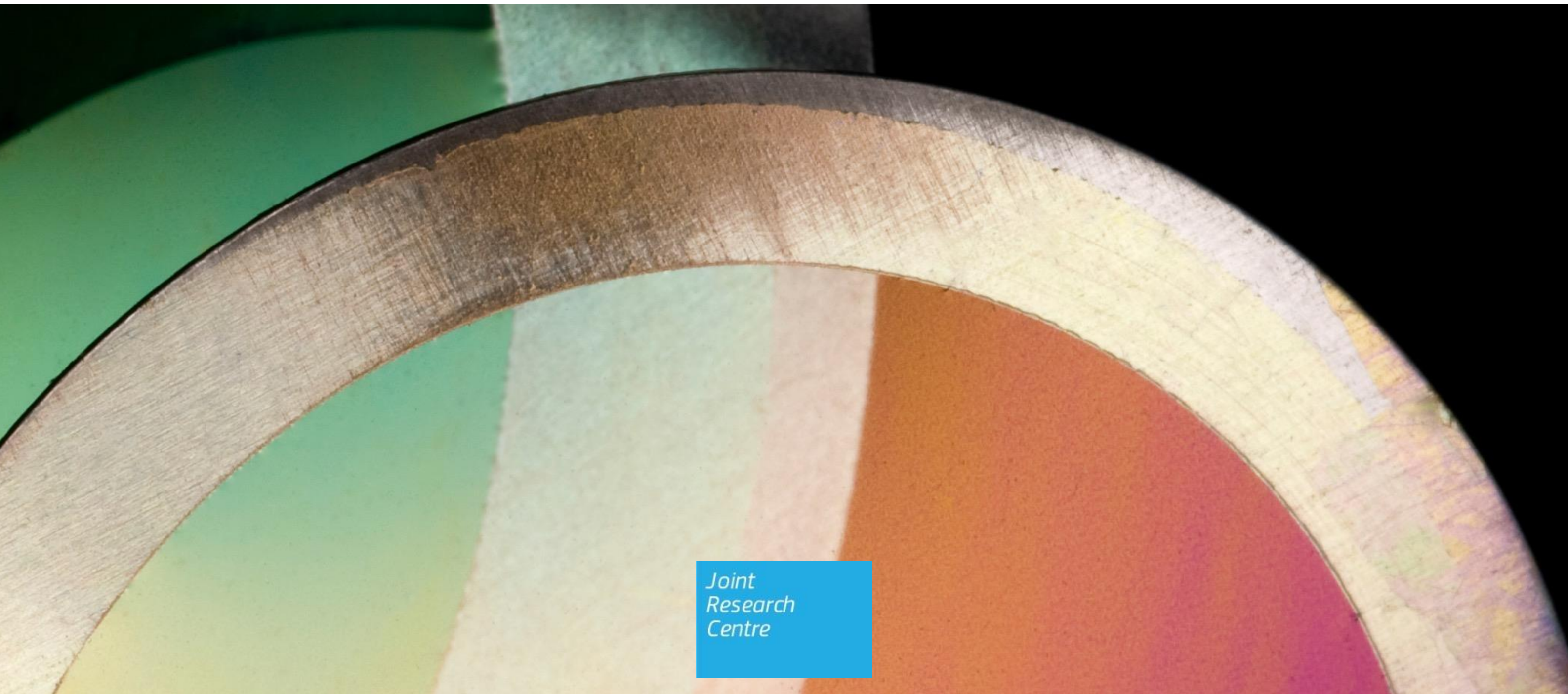
Replication Studies in Economics

How Many and Which Papers Are Chosen for Replication, and Why?

Frank Mueller-Langer[1,2], Benedikt Fecher[3], Dietmar Harhoff[2] and Gert G. Wagner[3,4,5,6]

[1] European Commission, Joint Research Center
[2] Max Planck Institute for Innovation and Competition
[3] Alexander von Humboldt Institute for Internet and Society
[4] Max Planck Institute for Human Development
[5] DIW Berlin, German Socio Economic Panel Study (SOEP)
[6] Institute of Labor Economics (IZA)

April 2018



Joint Research Centre

Contact information
European Commission, Joint Research Centre
Address: Edificio Expo. c/Inca Garcilaso, 3. 41092 Seville (Spain)
E-mail: frank.muller-langer@ec.europa.eu
Tel.: +34 9544-88731

JRC Science Hub
https://ec.europa.eu/jrc

JRC111765

Seville, Spain: European Commission, 2018

How to cite: Frank Mueller-Langer, Benedikt Fecher, Dietmar Harhoff and Gert G. Wagner, Replication Studies in Economics: How Many and Which Papers Are Chosen for Replication, and Why?; Digital Economy Working Paper 2018-01; JRC Technical Reports.

Abstract

We investigate how often replication studies are published in empirical economics and what types of journal articles are eventually replicated. We find that from 1974 to 2014 0.10% of publications in the Top 50 economics journals were replications. We take into account the results of replication (negating or reinforcing) and the extent of replication: narrow replication studies are typically devoted to mere replication of prior work while scientific replication studies provide a broader analysis. We find evidence that higher-impact articles and articles by authors from leading institutions are more likely to be subject to published replication studies whereas the probability of published replications is lower for articles that appeared in higher-ranked journals. Our analysis also suggests that mandatory data disclosure policies may have a positive effect on the incidence of replication.

Keywords: Replication, economics of science, science policy, economic methodology

JEL codes: A1, B4, C12, C13

1. Introduction

In times of increasing publication rates, cases of inadvertent errors and failure to fully report the fragility or robustness of results the mechanisms of quality control in science are put under scrutiny (McNutt, 2015). Replication studies can be considered an important post-publication quality check and have been described as a hallmark of good scientific practice (Campbell and Stanley, 1963; Jasny et al., 2011; Popper, 1959). Considering that an increasingly powerful science infrastructure makes it possible to store and retrieve data and code online at low costs, replications become more feasible. This is reinforced by the growing demand from publishers, science funders, and policy makers to make these materials available (e.g., Bohannon, 2015; Hoeffler, 2017).

The need for replicability is particular evident for economics; a field that traditionally informs political and economic decision-making as well as public discourse. While issues with the replicability of scientific research have been reported in multiple scientific fields, including economics (Anderson et al., 2008; Dewald et al., 1986; Camerer et al., 2016), systematic empirical evidence regarding the incidence of replication efforts in economics and their determinants is scarce.[1] But such evidence is a necessary precursor for designing favourable boundary conditions for professional self-control and for promoting trust in the scientific enterprise.

Here, we investigate replication in economics by examining which and how many published papers are selected for replication and what factors drive replication in these instances. The replication studies in our sample may differ in terms of (a) the results of replication, i.e. they may be negating or (partially) reinforcing the results of the replicated article, and (b) the extent of replication (narrow or scientific). Narrow replication studies are typically entirely devoted to replication of a particular result. In contrast, the replication work in scientific replication studies is often embedded in broader analysis. Note, however, that all replication studies in our sample, irrespective of replication result and extent, share the common property that their *main* purpose is to evaluate the accuracy of prior research.

We use metadata about all articles published in the Top 50 economics journals from 1974 to 2014. We find that replication is a matter of impact: high-impact articles and articles by authors from leading institutions are more likely to be replicated. Furthermore, our results suggest that the replication probability is lower for articles published in higher-ranked journals. We also find that mandatory data disclosure policies may have a positive effect on the incidence of replication. Finally, we discuss the implications of these results.

2. Background: Replication in economics

In 1982, the Journal of Money, Credit and Banking (JMCB) initiated the JMCB Data Storage and Evaluation Project (Dewald et al., 1986). Within this project, the JMCB required authors to make the data and code used in their articles available to other researchers upon request. In a second part of this project, Dewald et al. (1986) conducted replications of nine articles for which the data

[1] A notable exception is Duvendack et al. (2015).

was made fully available.[2] They were able to replicate the results of two articles in their entirety. Later, McCullough et al. (2006) tried to replicate 62 articles submitted to the same journal and could fully replicate the results of 14 articles.[3] Also, McCullough et al. (2008) tried to replicate 133 articles published by the Federal Reserve Bank of St. Louis Review and could replicate 29. These results raised concerns with respect to the technical and factual reproducibility of empirical work in economics.

The debate surrounding Reinhart and Rogoff (2010a, 2010b) and the non-supportive replications by Herndon et al. (2014) and Bell et al. (2015) has further advanced the debate about replication in the field of economics.[4] The study of Camerer et al. (2016), in which the authors attempted to replicate 18 studies that have been published in two top economic journals – the American Economic Review and the Quarterly Journal of Economics – between 2011 and 2014 drew renewed attention to the issue of replicability in economics. The authors were able to find a significant effect in the same direction as the original research proposed for only 11 of 18 replications (61%).

It should be noted that the situation that researchers are unable to replicate each other's results is well-known from other disciplines (for example biology; Begley & Ellis, 2012) where in some cases a single paper may be linked to a patent or a product (Fehder et al., 2014; Gans et al., 2017; Murray and Stern, 2007).[5] However, considering the impact of economic research outside the academic sphere (e.g., evidence-based policy making) and the increasing impact of empirical economic research–as measured by extramural citations–on other scientific disciplines (Angrist et al., 2017), there is a particular need to explore and understand replication in economics. This is necessary in order to design favourable boundary conditions for replication practice and to remove false results from the cumulative body of knowledge (see also Anderson and Kichkha, 2017; McCullough et al., 2006).

3. Conceptual framework

Towards a better understanding of the mechanics behind replication, it is helpful to review the different strands of literature on replicability in economics. Here, we focus on research on impact and competition as well as costs of replication. Examining the incidence of replication, we also

[2] From the first 54 data sets submitted to JMCB 14 were judged incomplete. For instance, the identification of the source of data was the most frequent problem in this respect.

[3] From 186 empirical articles, 69 had entries in the data archive. Replication of 7 articles was not possible because of confidentiality of data or lack of software.

[4] Notably, Carmen M. Reinhart and Kenneth S. Rogoff facilitated the detection of their coding error by making the data available upon request to Thomas Herndon, Michael Ash and Robert Pollin. In addition, Reinhart and Rogoff (2013) provide an erratum to Reinhart and Rogoff (2010b) where they correct their coding error. Reinhart et al. (2012) also address some of the methodology and exclusion issues raised by Herndon et al. (2014). Finally, Bell et al. (2015) used Reinhart and Rogoff's data, made available by Herndon et al. (2014), to re-examine the relationship between growth and debt in developed countries. Using a multilevel distributed lag model, Bell et al. (2015, p. 470) provide evidence for a reverse causal link "predominantly in the opposite direction to that mooted by Reinhart and Rogoff."

[5] In addition, the scientific community has repeatedly experienced instances of misconduct and erroneous analyses (Lacetera and Zirulia, 2011; Azoulay et al., 2015). The data fraud scandal concerning social psychologist Diedrik Stapel (Levelt et al., 2012), Hwang’s fraudulently reported breakthroughs in stem-cell research (Cyranoski, 2006), or Schoen’s entirely fabricated results on organic transistors in over 40 publications (Grant, 2002) are only the most prominent examples.

distinguish between narrow and scientific replications (Baltagi, 2003; Dewald et al., 1986; Duvendack et al., 2015; Hamermesh, 2007; Hunter, 2001).

3.1 Impact, competition, and costs

Several authors relate replication to impact and competition. For example, Hamermesh (2007, 2017) proposes that the probability that an individual will attempt a replication increases with the visibility of the published results that then become subject to replication. Furman et al. (2012) suggest that results from frequently cited articles generate more interest and scrutiny and—possibly, due to a higher probability of replication—have a higher likelihood of retraction. There is also empirical evidence that replication is a matter of impact. Sukhtankar (2017) examined replication of empirical papers in development economics published in the Top 5 and subsequent 5 general interest journals between 2000 and 2015. Of the 1056 articles, 57 (5.4%) were replicated in another published article or working paper. The strongest predictor of whether a study was replicated or not is the article's Google scholar citation count, followed by the year of publication (older articles are less likely to be replicated). Card and Della Vigna (2013) provide evidence that competition for publication in the top economics journals has significantly intensified over the last 50 years, i.e., the acceptance rate in the top economics journals has fallen from 15 to 6 percent from 1970 to 2012. The authors propose that due to increased competition researchers have improved the quality of their works. Furthermore journal editors can be more selective in choosing from a large number of submissions (Feigenbaum and Levy, 1993). It is for this reason that the editors of JPE (1975, p. 1295) suggested that "(t)he true remedy (against careless empirical research) is to resort to the powerful force of competition". In a similar fashion, Lacetera and Zirulia (2011) propose that, while referees examine the accuracy of submitted articles, the competitive environment will lead to critical review and replication. Dewald et al. (1986) and Hamermesh (2007) suggest that replication studies are more likely to get published when a central result of the original study is contradicted, e.g. they detect error or fraud, and that a replicator's main objective is to publish a correction or comment. On the basis of this strand of literature, we propose that competition in general and the "news value" of replication results in particular can serve as an explanation for the instances of replication in academic journals.

Another strand of literature focuses on the role of access to data and supplementary material for replication via mandatory data disclosure policies. Several authors suggest that data disclosure enhances the quality of articles as it reduces the cost of checking empirical results and encourages more careful research (Frisch, 1933; Lacetera and Zirulia, 2011). Dewald et al. (1986) suggest that data disclosure decreases the frequency and magnitude of errors in published articles. Analogously, the (effort) costs of replication can be considered an important factor for replication. Arguably, the cost of replication is *ceteris paribus* lower if data is made available under a data disclosure policy. In addition, the cost of replication is likely to be particularly high when confidential or self-created data used in an article is not made available under a data disclosure policy. Previous research in this area addressed the low willingness among scientists to make research data available (Andreoli-Versbach and Mueller-Langer, 2014; Haeussler, 2011; Savage and Vickers, 2009; Tenopir et al., 2011, Fecher et al., 2017), poor data documentation in published research (Ioannidis, 2005) and missing data availability policies in scientific journals (McCullough et al., 2008; Vlaeminck, 2013).[6]

[6] See also Mueller-Langer and Andreoli-Versbach (2018) for a theoretical analysis of the welfare effects of universal mandatory data disclosure.

In addition, we consider the incentives of researchers to engage in replication studies. A researcher would typically choose between allocating time to create an original research article and/or a replication study. According to Feigenbaum and Levy (1993, p. 217), the optimization problem of the researcher "is dictated by the relative returns in citations yielded by original versus replication work, as compared to the relative time cost of the two alternative endeavors." Feigenbaum and Levy (1993) also suggest that the expected citation returns to replication studies are lower than those to original research articles. This provides an explanation why replication studies are rarely conducted. Or, as Dewald et al. (1986, p. 587) put it: "replication (…) does not fit within (…) the reward structure in scientific research".[7] However, under the before-mentioned conditions, it will only be beneficial for a researcher to engage in replication studies if it is substantially less costly to produce them. Feigenbaum and Levy (1993) argue that the availability of data and code will decrease these costs.

Finally, from the perspective of journals, Feigenbaum and Levy (1993) suggest that an editor's rationale to publish replications of one's own journal's articles differs from the decision to publish replications of another journal's articles. In particular, they suggest that a rational, citation-maximizing editor of a high-ranked journal may decline the publication of a replication study on an article published in her own journal anticipating that editors of lower-ranked journals, in expectation of positive citation effects, may still find it optimal to publish this replication study. It is in this respect that the editor of a high-ranked journal "captures the citation externality without rewarding the replicator" (Feigenbaum and Levy, p. 223).

3.2. Types of replication

Following Hamermesh (2007), we distinguish between two types of replication studies: pure and scientific replication. In a "pure replication" (often also referred to as "narrow replication"), authors of a replication study use the same data and the same methods as the authors of the replicated article.[8] In contrast, authors of scientific replications use different data and, possibly, different methods than the authors of the respective replicated articles. Following the taxonomy provided by Baltagi (2003) and the *ReplicationWiki*[9], we consider three subtypes of scientific replications: (1) wide replications using different data but the same methods and models as the replicated article,[10] (2) replications using the same data but new methods and models,[11] and (3) replications using new data and new methods and models.[12] Hence, the extent of replication examined in the present study ranges between the extremes of pure replication on the one hand and scientific replication using new data and new methods on the other. Arguably, the different types of replication studies (scientific or pure) and the different purposes of replication, i.e. creation of new knowledge or mere checking of results, are relevant with respect to the question where replication studies should be published or posted. For instance, one may argue that scientific replications should undergo the same quality assessment as "regular" journal articles. Therefore, scientific replications should be

[7] See also Kuhn (1970) and Wible (1991).
[8] For example, Zhang and Ortmann (2014, p. 415) provide a "replication of Engel's (2011) study using his data and statistical methods".
[9] http://replication.uni-goettingen.de/wiki/index.php/Replication_in_economics (last accessed 21 August 2017)
[10] For example, see Taylor et al.'s (2010) wide replication of Hastings (2004).
[11] To illustrate, Rock et al. (2000) use the same data as the respective replicated article (Bhushan, 1989), but use count-data econometrics instead of OLS.
[12] See Bali et al.'s (2005, p. 906) replication of Goyal and Santa-Clara (2003) where they use an extended sample period and provide alternative measures of idiosyncratic risk.

published in peer-reviewed journals in regular issues or replication sections (Baltagi, 2003). In contrast, sites such as RePEc, SSRN, arXiv.org or ReplicationWiki appear to be suitable outlets for pure replications where the mere checking of results is at the core of the replication effort. Replication journals such as the recently established International Journal for Re-Views in Empirical Economics appear to be suitable outlets for all types of replication studies.[13] Finally, we take into account the replication result, i.e. we examine whether replication studies support or contradict prior published findings.[14]

4. Data and variables

The period under study is 1974 to 2014. We obtain our sample from the population of all articles published in the Top 50 economics journals (in total, 126,505 articles). From these 126,505 articles we identified 130 published replication studies, i.e., 0.10% were replications. Following Duvendack et al. (2015), we categorized a published journal article as a replication study if its main purpose is to test the reliability of a previously published study.[15] Consequently, this definition does not include original research articles where indirect replication, e.g. as initial "sanity check" to test the validity of a data set, either does not occur at all or *does* occur but is not the main purpose of an article. In particular, while the narrow replications under study are typically entirely devoted to replication, the replication work in the scientific replications under study are embedded in broader analysis. However, in any case, the testing of a previously published result always plays the main role in all replications under study.

Having identified 130 published replication studies, which is consistent with prior research on the frequency of published replication studies in economics,[16] we then identified the respective journal articles (henceforth, replicated articles) that were eventually replicated. In order to study differences between replicated and non-replicated articles, we selected all articles in issues of Top 50 economics journals that contained at least one replicated article. Our sample thus consists of 1,243 articles, 130 of which were replicated. Including non-replicated articles from the same journal issues allows us to control for latent effects at the level of journal issues while keeping the data collection effort manageable. Note that our approach only covers instances of formal replications, i.e., replication studies published in journals. Informal replication studies that are not published (e.g., replication efforts that are conducted in teaching or are part of an initial step in a research project) are not covered (see also Section 6.1).

---

[13] See https://www.zbw.eu/en/about-us/key-activities/research-data-management/iree/ (last accessed 14 September 2017).

[14] See also Section 4.1 where we provide detailed examples of different replication results.

[15] See also Sections 3.2 where we define replication type (pure or scientific) and Section 4.1 where we define replication result (negating or reinforcing).

[16] Notably, Duvendack et al. (2015) have also identified published replication studies in economics. They used (a) keyword searches in Google Scholar and in all 333 economics journals listed in WoS, (b) entries in the ReplicationWiki, (c) suggestions from editors and (d) their own collections of replication studies. They then performed a more systematic search within the Top 50 economics journals. Overall, Duvendack et al. (2015) found 162 published replication studies in the economics literature. The number of replication studies in our sample (130) is lower than the one in Duvendack et al. (2015) for the following reason. In our sample, the replication studies and respective replicated articles are both published in the Top 50 economics journals according to WoS impact factors. In contrast, Duvendack et al. (2015) also consider replication studies published in the 283 WoS-listed economics journals that have a lower impact factor.

4.1. Data and sample creation

We followed two distinct strategies to identify replication studies. First we considered Web of Science (WoS) metadata for all articles published in the Top 50 economics journals (126,505 articles). From WoS we retrieved information on the titles and abstracts of the 126,505 articles under study. We used this information to count, for each article, how often indicative key (stem) words such as "repli*," "reexamin*,", "comment," "revisit," "retesting," or "reappraisal" (among others) as well as references to other articles appeared in its title and abstract.[17] Both the frequency as well as the location of these key words allowed us to determine the likelihood of being a replication study for all articles under study. To illustrate, an article that contains "repli*," "retesting," or "note" in the title and contains a reference to another article in the abstract, for instance, as indicated by "(19*)" or "(20*)", has a high likelihood of being a replication study.[18] All articles were ranked in terms of the likelihood of being a replication study.[19] For the 100 highest-ranked articles of each journal, we studied the articles in detail in order to identify replication studies. In addition, we included all eligible replication studies published on the website of *ReplicationWiki*[20] in our data set. We then identified the respective journal article that was eventually replicated.[21]

For the analysis, we only considered empirical research articles and removed purely theoretical articles. We defined a purely theoretical article as an article that does not use any data. We searched for summary statistics and statistical tables in the PDFs in order to distinguish between empirical and purely theoretical articles.

We studied all replication studies to determine whether they either negate or (partially) reinforce the replicated article. A prominent example for the former is Joyce's (2009) non-supportive replication of Donohue and Levitt's (2001) results on the association between the abortion rate and age-specific crime rates. Joyce (2009, p. 112) provides empirical evidence "that there is no association between abortion and age-specific homicide rates or age-specific arrest rates for murder."[22] As another example, Dyl and Maberly (1986, p. 1149) negate Cornell's (1985) results on the weekly pattern in stock returns indicating the following: "(…) In an attempt to resolve this puzzling inconsistency, we replicated Cornell's study for the period from May 3, 1982, through July 24, 1984. We were unable to duplicate his results." In a more recent example, Abrevaya and Puzzello (2012) negate Adda and Cornaglia's (2006) empirical findings on the compensatory behaviour of smokers when cigarette taxes increase. Abrevaya and Puzzello (2012, p. 1760)

---

[17] Initially, we used a vector of 100 key words which was finally reduced to a vector of 70 stem words, e.g. to avoid double entries. For instance, we used the stem word "repli*" for both keywords "replication" and "replicable" or "reexamin*" for "reexamination" and "reexamine," respectively.

[18] However, our filtering algorithm also identifies comments and other publications which only use adjectives (and their stem words) such as "original" or verbs like "re-examine" in the abstract.

[19] We used the information obtained from the previous counting exercise to determine this ranking by journal as follows. First, we summed how often the 70 key stem words appear in the title and abstract of a given article in a given journal. Second, for a given journal, we order the articles according to the accumulated appearance of these key stem words. To illustrate, we consider the results of the ranking for the American Economic Review. Iversen and Söderström (2014) is the 2nd ranked article in terms of probability of being a replication study. The word "comment" appears in the title of the paper. In addition, stem words such as "comment," "correct*," or "(20*," appear in its abstract. In contrast, in Sheshinksi (1971) which is ranked 99th in terms of being a replication study the key word "note" appears in the title whereas all other key stem words do not appear in either title or abstract.

[20] http://replication.uni-goettingen.de/wiki/index.php/Main_Page (last accessed 30 January 2017)

[21] In the supplementary materials, we provide an overview of the replication studies and respective replicated articles.

[22] See also Joyce (2004), Donohue and Levitt (2004) and Fryer et al. (2013).

conclude that "(i)n this comment, we have re-examined this claim by AC and find little systematic evidence to support it." To give an example for a reinforcing replication study, consider Gerdtham et al. (1999, p. 117) who aim "to validate Wagstaffs and van Doorslaer's approach of constructing a continuous health measure" and whose "results (…) support the validity of the WvD method."[23] In addition, Hung and Plott (2001, p. 1518) reinforced Anderson and Holt's (1997) results on information cascades in the laboratory concluding the following: "The results of Anderson and Holt replicate (Result 1). In our experiments we observe the phenomena they report."

Among the Top 50 economics journals (according to the WoS impact factor), 23 had published at least one replicated article in the past.[24] We retrieved article metadata from WoS, i.e., publication date, number of references, pages and authors, and journal information. We gathered information on the rank position of the institutions that the authors are affiliated with from the Ranking Web of Universities 2014 and obtained author citation metrics from Scopus.

Following Andreoli-Versbach and Mueller-Langer (2014) we identified journals that have data disclosure policies in place and the first volume in which the policy was adopted in order to identify articles that are subject to a data disclosure policy.

Funding bodies may require authors to make data or program code available. Therefore, we also analyzed the funding guidelines of 36 research funding bodies worldwide regarding their data management policies. To this end, we randomly selected 27 public funding bodies from the 15 countries with the highest public expenditure for research according to the OECD (2016). We added 9 funding bodies that are not necessarily public, but that support noncommercial research (e.g., Bill and Melinda Gates Foundation), that are international (e.g., the European Commission's Horizon 2020), or that are from countries that are too small to appear in the list of the top funding countries (e.g., the Swiss National Science Foundation). In addition to screening the websites manually for the guidelines, we contacted every funding body in our sample via email in two waves and asked for relevant documents. 20 funding bodies replied to our information request. We were able to gather the guidelines from all 36 funding bodies. We regard the guidelines as textual data and code if they mention data management, if they specify where and how data should be stored, its terms of access (e.g., on request or public), documentation standards, and if they mention replication studies. From the 36 guidelines under study, 22 (61%) mention data sharing; 19 (53%) specify how or where to publish data;[25] 20 (56%) mention data documentation; 16 (44%) require data management plans; and 20 (56%) mention an embargo period for data in which the principal investigator has exclusive rights to access and publish with the data. Notably, none of the funding bodies under study has an explicit replication policy.

4.2. Variables

Table 1 provides an overview of the dependent and independent variables used in our study.[26]

---

[23] See Wagstaff and van Doorslaer (1994).
[24] Appendix 1 provides on overview of the journals under study, year of enactment of mandatory data disclosure policies and number of publications.
[25] For instance, 12 guidelines (33%) indicate that data from funded projects should be stored in a public repository while 2 (6%) mention that data needs to be made available upon request.
[26] Appendix 2 provides a correlation matrix for the dependent variable and main variables of interest.

4.2.1 Dependent variable

The dependent variable, *ReplicatedArticle*, is a binary variable indicating whether the article under study was eventually replicated. This variable measures the joint likelihood of a replication being undertaken and then being published. Merely analyzing the incidence of replication studies being undertaken (measurable by replications described in discussion papers or other pre-publication media) would not capture the most important filter mechanism in academic communication: review and ultimate publication. We obtained *ReplicatedArticle* by identifying replication studies published in the Top 50 economics journals and the respective replicated articles published in these journals.

[Table 1 HERE]

Figure 1 shows that the total number of journal articles has increased at a slightly higher rate than the total number of published replication studies. It is also noteworthy that the share of published replication studies on the total number of journal articles per year never exceeds 0.26% in the period from 1974 to 2014. From our sample we also estimate that the share of empirical articles rose from about 73% in 1975 to about 80% in 2010 (Figure 2).

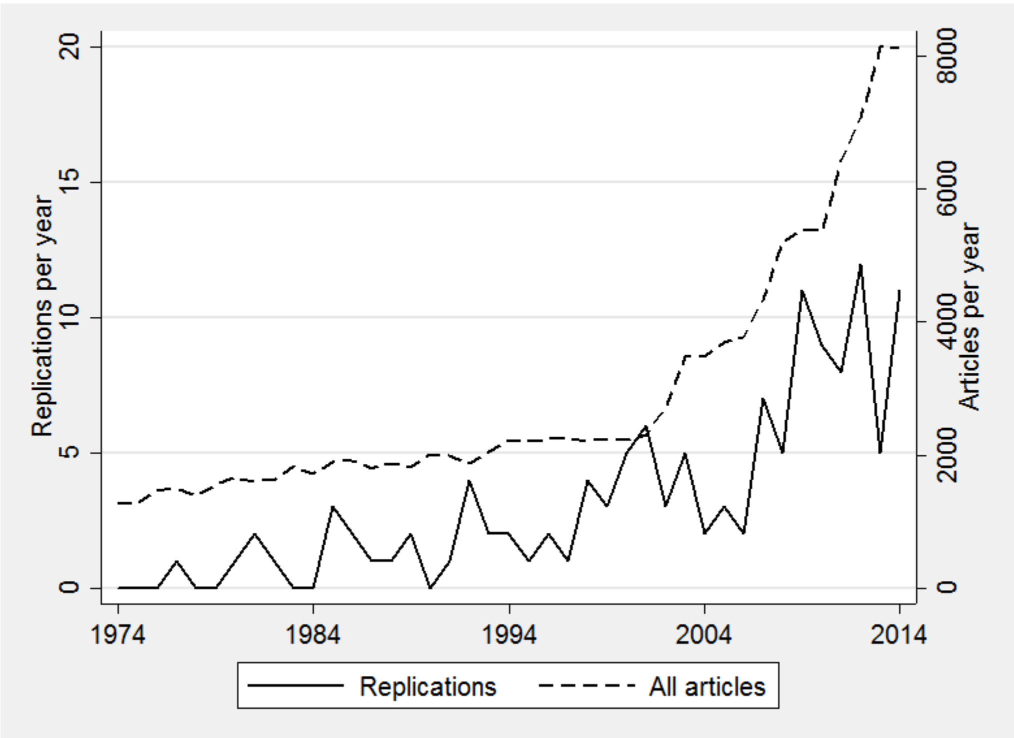


Figure 1 | Total number of published replication studies by publication year (as indicated by the vertical axis on the left-hand side) and of all journal articles by publication year (vertical axis on the right-hand side). Figure 1 suggests that the total number of journal articles has increased at a slightly higher rate than the total number of published replication studies. Sample: All articles published in the Top 50 economics journals (according to WoS) from 1974 to 2014 (including 130 published replication studies).

Replication studies may be negating or (partially) reinforcing the results of the replicated article. 61 of the 130 replication studies under study are negating (partially reinforcing: 69). In addition, 102 of the 130 replication studies under study, i.e. 78.5%, are published in the same journal as the respective replicated articles. 52.9% of the 102 replicated articles with replications in the same journal are positively replicated (negatively replicated: 47.1%). The impact factors of journals where positive and negative same-journal replications are published are, in average, very similar (positive: 3.2; negative: 3.3). In terms of positive versus negative replication, a similar picture emerges for the 28 replicated articles where the respective replication study is published in a different journal (positively replicated: 53.6%; negatively replicated: 46.4%). We now consider what fraction of these 28 different-journal replication studies is published in higher ranked (and lower ranked) journals than the respective replicated article. 32.1% of the different-journal replications are published in higher-ranked journals and 67.9% in lower-ranked journals, respectively.

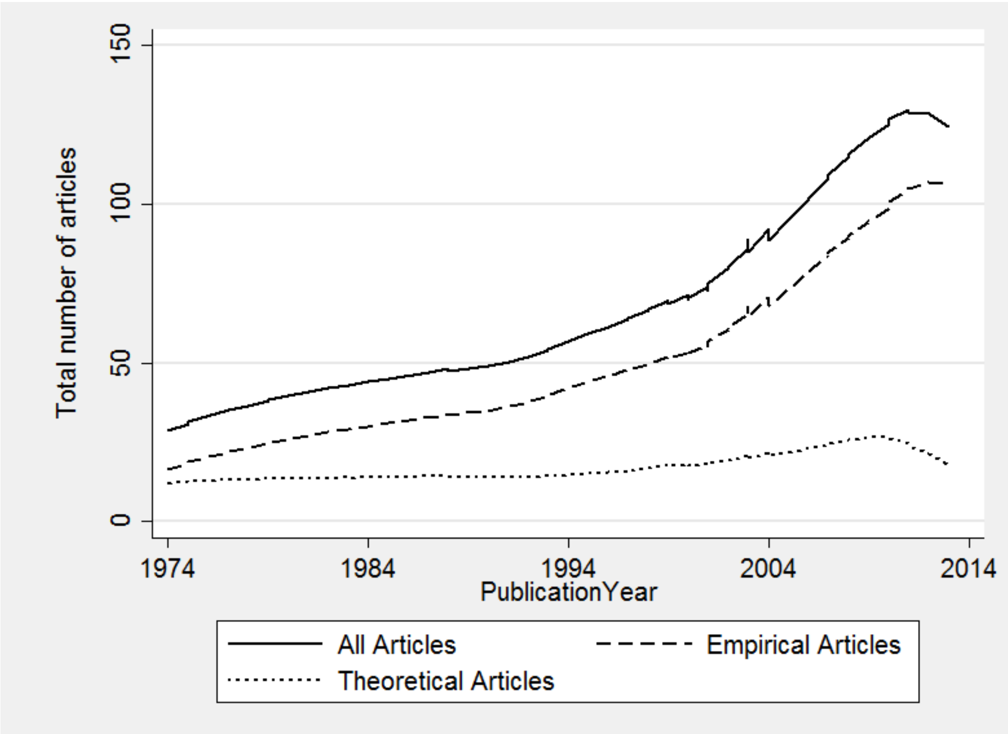


Figure 2 | Total number of empirical and purely theoretical articles by publication year. Sample: All empirical and purely theoretical articles in issues of Top 50 economics journals that contained at least one replicated article. Recall that purely theoretical articles are exluded in our main sample of 1,243 empirical articles that we use in the regression analysis. Here, we refrain from excluding 445 purely theoretical articles. We use the *lowess* command in STATA that performs locally weighted scatterplot smoothing (Cleveland, 1979). We refrain from displaying the three scatterplots to improve the readability of the figure.

Finally, we distinguish between two main types of replication studies and replicated articles, respectively. First, narrowly (or purely) replicated articles are eventually replicated in articles that use the same data and code. 20 articles under study are narrowly replicated articles. Second, 107 articles under study are scientifically replicated articles. This type of replicated article comprises widely replicated articles (new data but same method: 29 articles), articles that are replicated using the same data but new methods (48 articles) and articles that are replicated using new methods and new data (30 articles).[27]

4.2.2 Independent variables

We distinguish between the main variables of interest and control variables indicating article, author, journal, and institutional characteristics (see Table 1).

As for the main variables of interest, *CitesPreReplication* indicates the total number of citations of replicated articles and nonreplicated articles in the same issue one year prior to the publication of a replication study. To illustrate, Couch and Placzek (2010) replicate Jacobson et al. (1993). Hence, total cites one year before the publication of the replication study are given by the cumulative citations generated from 1993 to 2009 (being equal to 241 in this case). Arguably, replicated articles may attract more cites prior to the publication of a replication study if there is a longer lag between the publication of the article and the publication of the respective replication study. *LagReplication* measures this lag in years.

Table 2 provides extended descriptive statistics on *CitesPreReplication* and *LagReplication* by subgroups while distinguishing between replicated and nonreplicated articles. The descriptive statistics indicate that, prior to replication, replicated articles have attracted more cites than nonreplicated articles for virtually all subgroups.[28] Replicated and non-replicated articles published in the Top 5 journals attracted more total cites (by a factor of four) than articles published in other Top 50 journals. In addition, the lag between the publication of an article and the respective replication study is significantly longer for the Top 5 journals (6.71 years) than for other Top 50 journals (3.64 years). Replicated and non-replicated articles from authors affiliated with a Top 50 university attracted, on average, more citations than articles from authors affiliated with lower-ranked universities. While we obtain similar results for articles that received external research funding, the opposite is true for articles published in conference proceedings.

[Table 2 HERE]

The quality of a journal is given by the impact factor recorded by WoS in 2014, *ImpactFactor*. We include a dummy variable, *Top5Journal*, identifying articles published in one of the top five economics journals according to Card and DellaVigna (2013), i.e., American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, and Review of Economic Studies. For single-authored papers, *Top50University* indicates that the author is affiliated with a Top 50 university according to the Ranking Web of Universities 2014. For papers with multiple authors, it indicates that the author from the highest-ranked university of all authors is affiliated with a Top 50 university. *MandatoryDisclosure* is a binary variable indicating whether an article is subject to a mandatory data disclosure policy. It is equal to zero for articles published in journals without mandatory data disclosure. It is also equal to zero for articles published in journals with data disclosure policy if the article was published in a volume before the policy was

[27] We refrain from coding three replication studies (Salas and Raftery, 2001; Lee, 2008; Fraas and Lutter, 2012) as pure or scientific as they examine the econometric approaches adopted in the respective replicated articles theoretically without using the same or new data as the replicated articles.

[28] The two only subgroups where non-replicated articles attracted slightly more cites than replicated articles are *SelfCreatedData*=1 and *Pyear80*=1.

enacted. It equals one for articles published in the volume of enactment and all subsequent volumes. *DataOrCode* is a dummy variable indicating whether the data or program code used in an article are available on the journal website. For instance, journals with mandatory data disclosure policies typically provide a link to the data and program code together with the PDF of the article.[29]

As for the control variables, *SelfCreatedData* is a dummy variable indicating whether the data used in an article are self-created (in contrast to archived data which is re-analyzed), e.g., via laboratory or field experiments, surveys, or interviews. In order to make this classification we read all articles in detail. Publications that exclusively use publicly available data, e.g., census data, did not qualify as articles using self-created data. We read all explanatory notes in order to determine whether the data used in an article was confidential or proprietary to generate the dummy variable *ConfidentialData*. *ProceedingsArticle* is a binary variable indicating whether articles were published in conference proceedings. *References* and *Pages* are defined as the total number of references and pages. *Authors* indicates the number of authors. To control for author quality, we created the variable *BestH*, which indicates the h-index of authors of single-authored articles or the highest h-index of all co-authors in the case of multi-authored articles, respectively. *Funded* is a dummy variable indicating whether an article received third-party funding.[30] *FunderDataSupport* is a variable with range from 0 to 5 that indicates the number of data policies and data management tools that external research funders provide to the authors they support. This variable can be thought of as the extent to which external funders have policies that facilitate data availability.[31] We also create binary variables indicating the time of publication in 5-year intervals (henceforth, year fixed effects). Finally, according to WoS, articles under study fall in one the following Business & Economics research areas: General (76.4%), Mathematical Methods (10.2%), Environmental Sciences (3.6%), Health Care Sciences (0.9%) or Transportation (0.01%). Notably, WoS assigns the same research area to all articles in a given journal. By including binary journal variables in the regressions, we therefore control for WoS research areas.

5. Empirical results

We first report descriptive results on the prevalence and enforcement of mandatory data disclosure policies of journals. Then we perform probit regressions to examine the article-level correlates of the probability that a journal article is eventually replicated.

5.1 Enforcement of mandatory data disclosure policies

In our sample, nine of the 23 journals under study have a mandatory data disclosure policy (see Appendix 1). Notably, two explicit replication policies (JPE and Labour Economics) were suspended,

---

[29] Journals with such a policy typically follow the mandatory data disclosure policy of the American Economic Review which states: "Authors of accepted papers that contain empirical work, simulations, or experimental work must provide to the Review, prior to publication, the data, programs, and other details of the computations sufficient to permit replication. These will be posted on the AER Web site".

[30] We retrieved this information from the acknowledgment sections of the articles under study and from WoS article meta-data. In our sample, the five most frequent funding bodies in terms of articles funded are the National Science Foundation (12.6% of all articles), Economic and Social Research Council (1.9%), Social Sciences and Humanities Research Council of Canada (1.5%), Medical Research Council (0.5%), and Leverhulme Trust (0.5%).

[31] Appendix 3 provides on overview of the data guidelines of the funding agencies under study.

according to the editors, due to a lack of interest in replication (Hamermesh, 2007). Moreover, 286 of the empirical articles under study are subject to a mandatory data disclosure policy, i.e. they were published after the data policy was introduced. Notably, the data sets used or program codes are available on the respective journal website for 183 of these articles (63.9%). We identified one article out of 286 that was subject to a mandatory data disclosure policy, but for which the data or program code was not available because it was proprietary (which leads to an exemption from disclosure). This suggests that for 35.7% (102) of the 286 empirical articles subject to mandatory data disclosure policies, the data or program code was not available although the data was not proprietary. This result raises concerns regarding the enforcement of mandatory data disclosure policies. Figure 3 illustrates the total number of articles published under a mandatory data disclosure policy and the total number of articles published under a mandatory data disclosure policy where the policy is not strictly enforced. It suggests that both numbers increase over time and that for a large share of journals mandatory data disclosure policies are announced but not always enforced or monitored.

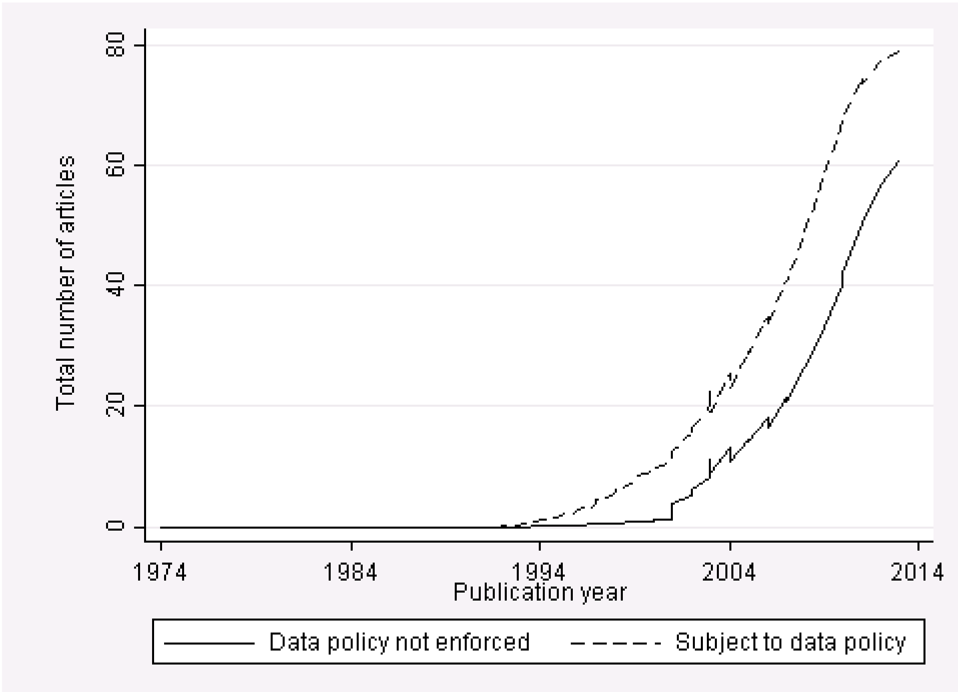


Figure 3 | Prevalence and enforcement of mandatory data disclosure by publication year. The vertical axis shows the total number of articles subject to a mandatory data disclosure policy (as given by the dotted line) and the total number of articles subject to a mandatory data disclosure policy where the policy is not strictly enforced (as given by the solid line). Figure 3 suggests that both numbers increase over time and that for a large share of journals mandatory data disclosure policies are announced but not always enforced or monitored. We use the *lowess* command in STATA that performs locally weighted scatterplot smoothing (Cleveland, 1979). We refrain from displaying the two scatterplots to improve the readability of the figure.

However, these findings should be interpreted with caution given the relatively small number of observations.

5.2 Empirical analysis of replication studies

To examine the article-level correlates of the probability that a journal article is eventually replicated we run probit (and OLS) regressions as given by:

$$prob(Replicated\,Article) = \beta_1 \cdot ProceedingsArticle + \beta_2 \cdot LogReferences + \beta_3 \cdot LogPages$$
$$+ \beta_4 \cdot Authors + \beta_5 \cdot LogBestH + \beta_6 \cdot SelfCreatedData + \beta_7 \cdot ConfidentialData$$
$$+ \beta_8 \cdot Funded + \beta_9 \cdot FunderDataSupport + \beta_{10} \cdot LogCitesPreReplication$$
$$+ \beta_{11} \cdot LogLagReplication + \beta_{12} \cdot Top5Journal + \beta_{13} \cdot LogImpactFactor$$
$$+ \beta_{14} \cdot Top50University + \beta_{15} \cdot MandatoryDisclosure + \beta_{16} \cdot DataOrCode.$$

We also include dummy variables for 23 journals and year fixed effects. Note that author citation metrics are not available for 18 out of 1,243 observations. In the regressions we use log transformations. As some researchers have an h-index of zero, we define *LogBestH* = log(*LogBestH* + 1). We follow the same procedure for *LogReferences*. We compute robust standard errors clustered at the journal level. All specifications reported in this paper are straightforward modifications of this baseline specification.

5.2.1 Full sample analysis

We run our regressions with nine different specifications as reported in Table 3. In specification [1], we consider control variables (article, author and institutional characteristics), journal and year fixed effects. Table 3 does not report marginal effects of the control variables (Appendix 4 provides the full version of Table 3). For all specifications, we provide the Wald test statistics for these variables. We include the log of total citations before publication of a replication study and the log lag between the publication of replicated articles and of the respective replication studies in specification [2] to examine the effect of article impact on the replication probability. In specification [3], we include the Top 5 economics journal dummy variable and the log of the journal impact factor to account for the effect of journal quality on the probability of replication. In specification [4], we add the affiliation with Top 50 universities. We include the binary variable for mandatory data disclosure in specification [5] to examine the impact of the existence of mandatory data disclosure policies on the replication probability. In specification [6], we include the binary variable for actual data or code availability. We separately include *MandatoryDisclosure* and *DataOrCode* in columns [5] and [6] because of their high correlation (0.69; see Appendix 2).[32] In column [7], we examine articles published in issues that contain at least one scientifically replicated article.[33] Finally, we examine the subsamples of journal articles published in issues with at least one negated replicated article (reported in column [8]) and at least one (partially) reinforced replicated article (reported in column [9]).

[Table 3 HERE]

[32] The high correlation reflects that authors of articles published in journals with mandatory data disclosure policies are more likely to make their data and code available. Results are qualitatively unchanged when we include both variables in column [5]. In particular, effects of both variables remain insignificant. Our results are also virtually the same when we run the regressions separately for the subsample of articles that fall under the WoS category "Business & Economics: General".

[33] We refrain from running the regressions separately for the subsample of articles published in issues that contain at least one narrowly replicated article because of the low number of narrowly replicated articles in our sample, i.e., 20 articles.

The results reported in Table 3 provide empirical evidence that the impact of journal articles—measured in citations—positively affects the probability that they are eventually replicated. The marginal effects at the mean of *LogCitesReplication* range from 0.040 to 0.068 and are statistically significant at the 0.1% level across all specifications. In addition, the log lag between the publication of the replicated articles and of the respective replication studies negatively affects the probability of replication. This effect is statistically significant at the 0.1% level across specifications. Interestingly, the effect of *LogLagReplication* is larger in magnitude than that of *LogCitesPreReplication*. This indicates that articles initially published further in the past are less likely to be the subject of a published replication study, irrespective of their citation performance. Our analysis further suggests that articles published in better journals are less likely to be eventually replicated. The marginal effects of both *Top5Journal* and *LogImpactFactor* are negative and statistically significant at least at the 5% level across specifications. The only exception is the negative, statistically insignificant marginal effect of *LogImpactFactor* reported in column [8]. We find empirical evidence that journal articles by authors from better institutions have a higher incidence of being replicated. The marginal effects of *Top50University* are positive across specifications and statistically significant at the 5% level in all specifications but [8].

Finally, in the full sample analysis reported in columns [5] and [6], we find no significant effect of *MandatoryDisclosure* or *DataOrCode* on the probability of replication, respectively. We further explore this nexus in the following subsample analysis.

5.2.2 Subsample analysis: Articles published in or after 2004

It is important to note that, while the period under study is 1974 to 2014, disclosure of data and code is a relatively new phenomenon. For instance, Econometrica and the AER are the first journals in our sample to adopt a mandatory data disclosure policy in 2004 and 2005, respectively (see Appendix 1). Therefore, we run the regressions separately for the subsample of articles published in or after 2004. Results are reported in Table 4 where we include *MandatoryDisclosure* in columns [1] to [3] and *DataOrCode* in columns [4] to [6], respectively. In columns [2] and [5], we refrain from including *Top5Journal* as this dummy is highly correlated with *MandatoryDisclosure* (0.69) and *DataOrCode* (0.51), respectively. In columns [3] and [6], in order to explore whether our results are driven by one journal, we include a dummy variable for AER publication (instead of *Top5Journal*) and its interaction with *MandatoryDisclosure* and *DataOrCode*, respectively. We include the AER interaction to address concerns that our results are driven by a single journal. For instance, 35 of the 130 replicated articles under study are published in the AER. In addition, 254 of the 286 articles published under a mandatory data disclosure policy are AER articles.

[Table 4 HERE]

Comparing Table 4 with columns [1] to [6] of Table 3, we can see that the results on *LogCitesReplication, LogLagReplication, Top5Journal, LogImpactFactor* and *Top50University* remain qualitatively unchanged. In contrast, for this subsample, we find evidence for a positive effect of mandatory data disclosure polices on the probability of replication. The marginal effect (+5.6%) is statistically significant at the 0.1% level in columns [1] and [2]. It remains statistically significant at the 0.1% level and increases to +7.9% when we include the interaction of *MandatoryDisclosure* with the AER dummy in column [3]. This result suggests that mandatory data disclosure policies may decrease the cost of replication thereby increasing the probability of replication.

However, we do not find evidence for a positive effect of data or code availability on the replication probability in columns [4] and [5]. However, once we include the interaction of *DataOrCode* with the AER dummy in column [6], the marginal effect of data or code availability on the replication probability (+8.3%) is statistically significant at the 1% level while the effect of the interaction term is negative (-6.6%) and statistically significant at the 1% level.

5.3 Robustness

For robustness, we also perform linear probability regressions. Table 5 reports OLS regression coefficients for the full sample. These essentially yield the same results.

[Table 5 HERE]

In addition, we ran the regressions separately by article impact (Appendix 5), H-index of the best author (Appendix 6) and university rank of the best author (Appendix 7). Thereby, we address concerns that our analysis may be limited to comparisons between same-issue articles and may not allow us to compare same-journal articles with similar impact, author quality or institution quality. However, in this respect it is important to note that for 13 journals in our sample we observe at least two issues, i.e. 1,132 articles are published in these journals (92.4% of the sample in our preferred specification).[34] For instance, we consider 478 AER articles published in 29 issues (JPE: 13 issues; Journal of Health Economics: 11 issues). Therefore, we argue that, for 92.4% of the sample, we not only compare between same-issue articles but also same-journal articles. In Appendices 5, 6, and 7, we consider all issues in columns [1], [2] and [3] while we only consider articles published in journals with at least two issues in columns [4], [5] and [6] to highlight the comparison of same-journal articles. In Appendices 5, 6, and 7 the threshold values to create the subsamples under study are given by the 25%, 50% and 75% percentiles, respectively. All specifications in these appendices are based on specification [6] of Table 3.

Appendix 5 reports marginal effects at the mean after probit regression for the subsamples of articles with more than 11 cites (columns [1] and [4]), more than 28 cites (columns [2] and [5]), and more than 78 cites (columns [3] and [6]), respectively. As in the full sample analysis, we provide evidence that articles published in better journals are less likely to be eventually replicated. We also find that articles of authors affiliated with Top 50 universities are more likely to be eventually replicated. In addition, the availability of data or code does not have a significant effect on the replication probability. These results are consistent across all specifications. We now consider the subsample of high-impact articles (>78 total cites) reported in columns [3] and [6] to examine as to why some high-impact articles are replicated and others are not.[35] First, note that we do not find evidence that the replication probability increases in article impact because there is less variation in article impact. Second, the marginal effects of *Top5Journal* and *LogImpactFactor* are negative and statistically significant at least at the 1% level. Third, a comparison of the results reported in columns [1] to [3] shows that the positive effect of *Top50University* on the replication probability is highest for the most highly cited articles (17.5% in column [3] as compared to 5.9% in column [1] and 10.3% in column [2], respectively). Finally, all results reported in columns [1] to

[34] See Appendix 8. For each of these 13 journals we have 10 or more observations.
[35] We thank an anonymous referee who suggested examining this question.

[3] are virtually identical when we run the regressions for articles published in journals with at least two issues in our sample in columns [4] to [6].

In addition, Appendix 6 reports marginal effects at the mean after probit regression for the subsamples of articles by (best) authors with an H-index higher than 8, 15 and 23, respectively. As in the full sample analysis, we find that the replication probability increases in article impact and decreases in journal quality and that the availability of data and/or code does not have a significant effect. Finally, Appendix 7 reports marginal effects at the mean after probit regression for the subsamples of articles from authors affiliated with universities ranked lower than 142, 36 and 7, respectively. Recall that better universities have a lower position in the ranking. As in the full sample analysis, we find that the replication probability increases in article impact. Notably, the replication probability decreases in journal quality for articles by authors from top-universities (*BestRank*<7, see column [3]) while there is no significant effect in the two other cases reported in columns [1] and [2].

5.4 Characteristics of replicated and replicating authors

We examine the characteristics of authors that have their papers replicated (replicated authors) and compare them with the characteristics of the authors of the respective replicating articles (replicating authors).[36] In Figure 4, we consider the best authors of the 130 replicated/replicating-article-pairs under study in terms (A) H-index, (B) total cites, (C) affiliation with a top-50 university, and (D) university rank. Figure 4 depicts the mean of these four variables for replicated and replicating authors.

In average, the H-index of replicated authors (21.54) is almost twice as high as the H-index of replicating authors (10.98). Replicated authors' total citations are, in average, four times higher than total citations of replicating authors. In addition, replicated authors are more likely to be affiliated with a Top 50 university (0.73) than replicating authors (0.33). Finally, universities of replicated authors are, in average, better ranked than the universities of replicating authors.

[36] It is beyond the scope of the present article to examine the channels through which prospective replicating authors become aware of articles that are eventually replicated. In this respect, an analysis based on citation-networks appears to be a particularly interesting avenue for further research.

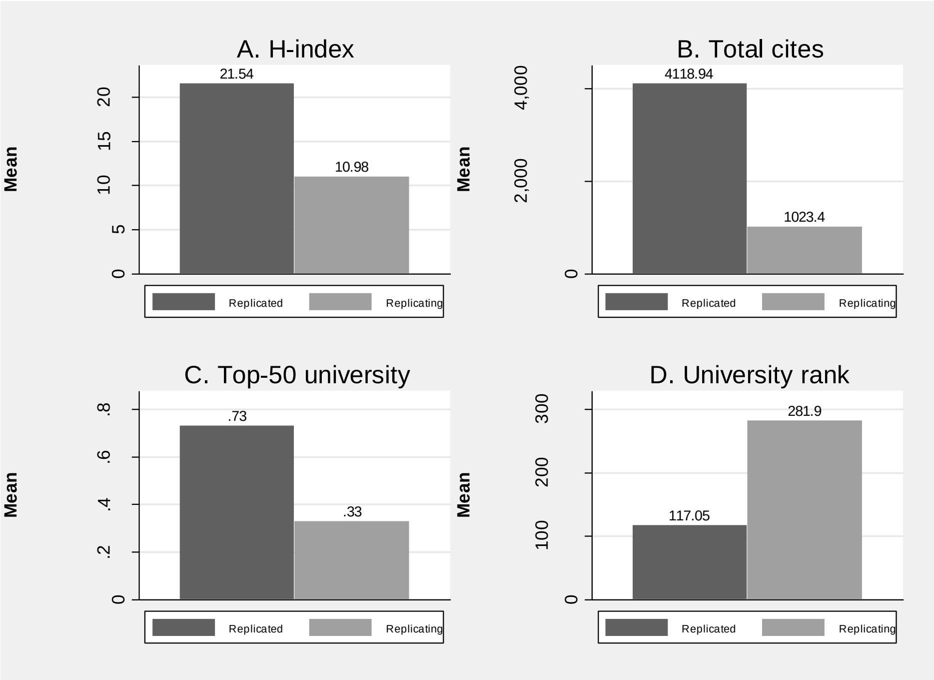


Figure 4 | Characteristics of replicated and replicating authors. Results obtained for the best authors of 130 replicated/replicating-article-pairs. Better universities have lower rank.

6. Discussion and further research

6.1 Informal vs. formal replications

Review and publication are the most important filter mechanisms in science. Our approach allows us to explore replication studies that are subject to these filter mechanisms, i.e., formal replications. However, it does not allow us to make statements about replications that have not been published in journals (e.g., replications published in working papers, initial sanity checks or replications done in course work). For example, Sukhtankar (2017) analyzed empirical papers in development economics published in the top-5 and next-5 general interest journals between 2000 and 2015. Of the 1056 papers, 57 (5.4%) were replicated in another published paper or working paper. The number drops to 29 (2.7%) if considering only published papers. Many replications might not be published at all. For example: A teacher might want her students to "learn from the best" and choose a study for a replication in class that was particularly well-performed. This replication is unlikely to be published as formal or informal paper. Furthermore, journals–in anticipating reader expectations and impact scores–publish according to a certain news value (Feigenbaum and Levy, 1993). One could argue that replications in general have little of that and become visible in working papers, but are never published in peer-reviewed journals. However, also informal replications help the field to strengthen its knowledge base. A full documentation of the instances of replications and their success in a public database, similar to the ReplicationWiki (Höffler, 2017), would be beneficial. A full documentation of informal replications enables identifying high-quality research or rewarding frequent replicators. Such a database would furthermore facilitate an answer to the question how many (more) replications are actually needed.

6.2 Optimal amount of replications

While having 0.10% of articles among the top 50 publications in economics being replications strikes us as considerably low, it is important to note that our analysis does not allow us to make a statement on the optimal amount of replications. Arguably, the optimal amount of replications strikes a balance between the opportunity cost of replication that arise as replicators exert less effort pursuing other (possibly more original) research and the benefits of replication that (a) false results are removed ex post from the cumulative body of knowledge and (b) incentives for careful empirical work are increased ex ante. As for the ex post effects of replication, note that actual replication has to take place on a case-by-case basis, i.e. a single negating replication study typically negates the results of a single article. In contrast, as for the ex ante effect of replication, note that a credible "threat" of replication could have an impact at a larger scale. However, for replications to constitute a credible threat, a sufficient level of actual replication needs to be achieved.

While an equilibrium model of replication behaviour is beyond the scope of this paper, in the following we briefly outline how such a model may determine the minimum replication needed to achieve a credible threat. Consider a researcher who chooses her effort for a new empirical research project. Thereby, she faces the trade-off between the private benefits and costs of careful empirical work. For instance, let the private benefits of carefulness be given by citations to the published article. Following Feigenbaum and Levy (1993, p. 219), in terms of citation streams, "an original article can die a violent death upon the publication of a negative replication". It is in this respect that the researcher's cost of the publication of a negative replication study can be interpreted in terms of forgone citations.[37] Arguably, under these conditions there is a minimum level of expected foregone citation streams due to the publication of a negative replication that is needed for negative replication to constitute a credible threat. To illustrate consider the following two extreme cases. First, assume that expected foregone citations equal zero because the incidence of negative replication equals zero. In this case, formal replication does not constitute a credible threat. Second, assume that the probability of publication of a negative replication equals one. This leads to the maximum level of expected foregone citations. In fact, if the negative replication is immediately published, i.e. it is published before the original article receives its first citation, a citation-maximizing researcher may not have an incentive to conduct empirical research in the first place. Formal negative replication constitutes a credible threat in this case. Arguably, the minimum level of expected foregone citation streams due to the publication of a negative replication study needed to achieve a credible threat will lie between these two extremes.

Finally, considering the growth in articles (Bornmann and Mutz, 2015), it seems highly relevant to examine further if the control mechanisms at work can cope with a higher publication output. If more replications in economics are needed, it would make sense to invest in lowering the costs of replication, for instance by promoting the availability of data and code, and increase the reputational gain from replication efforts (e.g., through awards).

[37] This provides an explanation for the observation that often the authors of original work perceive replication attempts as hostile (Mirowski and Sklivas, 1991; Maniadis et al., 2015).

6.3 Self-created vs. widely available data

Although we differentiate between self-created and widely available data in our analysis, it cannot unequivocally be concluded that non-compliance with data disclosure policies led to less replication nor if making source code available matters more than a general increase in data availability. For instance, it is possible that authors in our sample did not have to make their data available (and thereby comply with a journal's data policy) because their data was already widely available (e.g., census data).

7. Conclusions

Our results confirm previous assumptions that relate replication to impact (Dewald et al., 1986; Furman et al., 2012; Hamermesh, 2007 & 2017). Researchers tend to replicate high-impact research from renowned researchers and institutions. In this regard, private incentives are well aligned with societal interests, since high-impact publications are also the studies that are most likely to influence decision-makers in private and public organizations. To critically examine new insights that could change the discourse in a field seems expedient regarding the evolution of knowledge (Loasby 2002; Popper, 1959; Freedman et al., 2015).

Our results concerning data disclosure policies appear ambiguous at first sight. While we cannot detect any statistically strong impact of data disclosure policies on the replication probability for the full sample of articles published from 1974 to 2014, the picture changes when we restrict the analysis to the time period when data disclosure was introduced around 2004. Exploring the subsample of articles published in or after 2004, we find evidence for a statistically significant positive effect of mandatory data disclosure policies on the replication probability. Arguably, mandatory data disclosure policies may reduce the replicator's cost of replication thereby increasing the replication probability. The lack of significance in the overall sample may simply be due to the fact that a genuine regime change occurred around 2004 which is not captured in the overall regression. We also show that for 37% of the studied empirical articles subject to mandatory data-disclosure, the data or program code was not available although the data was not proprietary. This raises concerns regarding the enforcement of mandatory data disclosure policies.
Hence, our results suggest that replication efforts could be incentivized by reducing the cost of replication, for example by promoting data disclosure (see also Hoeffler, 2017). Our results further suggest that the decision to conduct a replication study is—at least partly—driven by the replicator’s reputation considerations. Other possible explanations are the importance of the topics explored in the replicated articles and their impact on public policy as well as editors' publication strategies vis-à-vis replication studies. We argue that the low number of replication studies being conducted could potentially increase if replication studies received more formal recognition e.g. through publication in (high-impact) journals, specific funding, e.g. for replications on articles with high impact on public policy, or awards[38]. A potentially viable strategy could be to promote replication studies in teaching and as an (optional) chapter of dissertations (Fecher et al., 2016). This would be easy to implement because cumulative dissertations are the standard for dissertations in economics.

[38] We find that funding agencies typically do not have replication policies (Appendix 3). We propose that research funders may introduce explicit replication policies if one of their objectives is to improve the boundary conditions for replication practice.

Acknowledgements
Some of the data used on this report were collected and first analyses done when Frank Mueller-Langer was Senior Research Fellow at Max Planck Institute for Innovation and Competition (MPI-IC). This work received support from the German Research Foundation (DFG) under the project European Data Watch Extended (DFG-Grant Number WA 547/6-2). The funders had no role in study design and in the writing of the manuscript. We thank Fabian Gässler, Bronwyn Hall, Karin Hoisl, Joachim Wagner, Stefan Wagner and Peter Winker as well as seminar and conference participants at MPI-IC, LMU Munich, TUM School of Management, and JRC Seville for comments. Moritz Mosenhauer, Wolfgang Rösch, and Christoph Winter provided research assistance.

Bibliography

Abrevaya, J. & Puzzello, L. Taxes, cigarette consumption, and smoking intensity: Comment. American Economic Review, 102(4), 1751-1763 (2012).

Adda, J. & Cornaglia, F. Taxes, cigarette consumption, and smoking intensity. American Economic Review, 96(4), 1013-1028 (2006).

Anderson, L. R. & Holt, C. A. Information cascades in the laboratory. American Economic Review, 87(5), 847-862 (1997).

Anderson, R. G., Greene, W. H., McCullough, B. D. & Vinod, H. D. The role of data/code archives in the future of economic research. Journal of Economic Methodology, 15 (1), 99-119 (2008).

Anderson, R. G. & Kichkha, A. Replication versus meta-analysis in economics: Where do we stand 30 years after Dewald, Thursby and Anderson? Paper presented at the American Economic Association Annual Meeting (2017).

Andreoli-Versbach, P. & Mueller-Langer, F. Open access to data: An ideal professed but not practised. Research Policy, 43(9), 1621-1633 (2014).

Angrist, J. D., Azoulay, P., Ellison, G., Hill, R. & Lu, S. F. Inside job or deep impact? Using extramural citations to assess economic scholarship. NBER Working Paper No. 23698 (2017).

Azoulay, P., Furman, J. L., Krieger, J. L. & Murray, F. Retractions. Review of Economics and Statistics, 97(5), 1118-1136 (2015).

Bali, T. G., Cakici, N., Yan, X. & Zhang, Z. Does idiosyncratic risk really matter? Journal of Finance, 60(2), 905-929 (2005).

Baltagi, B. Introducing a replication section. Journal of Applied Econometrics, 18(1), 111-111 (2003).

Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. Nature, 483, 531-533 (2012).

Bell, A., Johnston, R. & Jones, K. Stylised fact or situated messiness? The diverse effects of increasing debt on national economic growth. Journal of Economic Geography, 15(2), 449-472 (2015).

Bhushan, R. Firm characteristics and analyst following. Journal of Accounting and Economics, 11(2-3), 255-274 (1989).

Bohannon, J. Many psychology papers fail replication test. Science, 349, 910-911 (2015).

Bornmann, L. & Mutz, R. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. Journal of the Association for Information Science and Technology, 66(11), 2215-2222 (2015).

Camerer, C. F., Dreber, A., Forswell, E. et al. Evaluating replicability of laboratory experiments in economics. Science, 10.1126/science.aaf0918 (2016).

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research, Rand McNally, Chicago (1963).

Card, D. & DellaVigna, S. Nine facts about top journals in economics. Journal of Economic Literature, 51, 144-161 (2013).

Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association, 74(368), 829-836 (1979).

Cornell, B. The weekly pattern in stock returns: Cash versus futures: A note. Journal of Finance, 40(2), 583-588.

Couch, K. A. & Placzek, D. W. Earnings losses of displaced workers revisited. American Economic Review, 100(1), 572-589 (2010).

Cyranoski, D. Rise and fall. news@nature. doi:10.1038/news060109-8 (2006).

Dewald, W., Thursby, J. & Anderson, R. G. Replication in empirical economics: The Journal of Money, Credit and Banking Project. American Economic Review, 76(4), 587-603 (1986).
Doldirina, C., Friis-Christensen, A., Ostlaender, N. et al. JRC data policy, JRC Technical Report EUR 27163 (2015).
Donohue, J. & Levitt, S. D. The impact of legalized abortion on crime. Quarterly Journal of Economics, 116(2), 379-420 (2001).
Donohue, J. & Levitt, S. D. Further evidence that legalized abortion lowered crime: A reply to Joyce. Journal of Human Resources, 39(1), 29-49 (2004).
Duvendack, M., Palmer-Jones, R. W. & Reed, W. R. Replications in economics: A progress report. Econ Journal Watch, 12(2), 164-191 (2015).
Dyl, E. A. & Maberly, E. D. The weekly pattern in stock index futures: A further note. Journal of Finance, 41(5), 1149-1152 (1986).
Economic and Social Research Council. ESRC Research Data Policy, Retrieved from http://www.esrc.ac.uk/about-esrc/information/data-policy.aspx (2017).
Editors, Editorial Comment. Journal of Political Economy, 83, 1295-1296 (1975).
Engel, C. Dictator games: A meta study. Experimental Economics, 14(4), 583-610 (2011).
European Commission. Towards better access to scientific information: Boosting the benefits of public investments in research. European Commission: Brussels (2012).
European Commission. H2020 Programme: Guidelines on open access to scientific publications and research data in Horizon 2020, European Commission, Directorate-General for Research & Innovation: Brussels (2016).
Fecher, B., Fraessdorf, M. & Wagner, G. G. Perceptions and practices of replication by social and behavioral scientists: Making replications a mandatory element of curricula would be useful. IZA Discussion Paper Series No. 9896 (2016).
Fehder, D. C., Murray, F. & Stern, S. Intellectual property rights and the evolution of scientific journals as knowledge platforms. International Journal of Industrial Organization, 36, 83-94 (2014).
Feigenbaum, S. & Levy, D. M. The market for (ir)reproducible econometrics. Social Epistemology, 7(3), 215-232 (1993).
Fraas, A. & Lutter, R. Efficient pollution regulation: Getting the prices right: Comment, American Economic Review, 102(1), 602-607 (2012).
Freedman, L. P., Cockburn, I. M., & Simcoe, T. S. The economics of reproducibility in preclinical research. PLOS Biology, 13(6) (2015).
Frisch, R. Editor's note. Econometrica, 1(1), 1-4 (1933).
Fryer, R.G., Heaton, P.S., Levitt, S.D. & Murphy, K.M. Measuring crack cocaine and its impact. Economic Inquiry, 51(3), 1651-1681 (2013).
Furman, J. F., Jensen, K. & Murray, F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. Research Policy, 41(2), 276-290 (2012).
Gans, J. S., Murray, F. & Stern, S. Contracting over the disclosure of scientific knowledge: Intellectual property and academic publication. Research Policy, 46(4), 820-835 (2017).
Gerdtham, U. G., Johannesson, M., Lundberg, L. & Isacson, D. A note on validating Wagstaff and van Doorslaer's health measure in the analysis of inequalities in health. Journal of Health Economics, 18(1), 117-124 (1999).
Goyal., A. & Santa-Clara, P. Idiosyncratic risk matters! Journal of Finance, 58(3), 975-1008 (2003).
Grant, P. Scientific credit and credibility. Nature Materials, 1, 139-141 (2002).
Haeussler, C. Information-sharing in academia and the industry: A comparative study. Research Policy, 40(1), 105-122 (2011).
Hamermesh, D. Viewpoint: Replication in economics. Canadian Journal of Economics, 40(3), 715-733 (2007).

Hamermesh, D. What is replication? The possibly exemplary example of labor economics. Paper presented at the American Economic Association Annual Meeting (2017).
Hastings, J. Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in Southern California. American Economic Review, 94(1), 317-328 (2004).
Herndon, T., Ash, M. & Pollin, R. Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. Cambridge Journal of Economics, 38(2), 257-279 (2014).
Hoeffler, J. H. Replication and economics journal policies. Paper presented at the American Economic Association Annual Meeting in Chicago (2017).
Hung, A. A., & Plott, C. R. Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions. American Economic Review, 91(5), 1508-1520.
Hunter, J. The desperate need for replications. Journal of Consumer Research, 28(1), 149-158 (2001).
Ioannidis, J. P. A. Why most published research findings are false. PLoS Medicine 2, e124 (2005).
Jacobson, L. S., LaLonde, R. J. & Sullivan, D. G. Earnings losses of displaced workers. American Economic Review, 83(4), 685-709 (1993).
Jasny, B. R., Chin, G., Chong, L. & Vignieri, S. Data replication & reproducibility. Again, and again, and again .... Introduction. Science, 334, 1225 (2011).
Joyce, T. Did legalized abortion lower crime? Journal of Human Resources, 39(1), 1-28 (2004).
Joyce, T. A simple test of abortion and crime. Review of Economics and Statistics, 91(1), 112-123 (2009).
Kuhn T. S. The structure of scientific revolutions. Chicago: University of Chicago Press (1970).
Lacetera, N. & Zirulia, L. The economics of scientific misconduct. Journal of Law, Economics, and Organization, 27(3), 568-603 (2011).
Lee, R.H. Future costs in cost effectiveness analysis. Journal of Health Economics, 27(4), 809-818 (2008).
Levelt, W. J. M., Drenth, P. & Noort, E. Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. Commissioned by the Tilburg University, University of Amsterdam and the University of Groningen (2012).
Loasby, B. J. The evolution of knowledge: Beyond the biological model. Research Policy, 31(8-9), 1227-1239 (2002).
Maniadis, Z, Tufano, F. & List, J. A. How to make experimental economics research more reproducible: Lessons from other disciplines and a new proposal. In (C. Deck, E. Fatas and T. Rosenblat, eds.): Replication in Experimental Economics. Research in Experimental Economics, Volume 18, Bingley: Emerald Group Publishing, pp. 215-230 (2015).
McCullough, B. D., McGeary, K. A. & Harrison, T. D. Lessons from the JMCB Archive. Journal of Money, Credit and Banking, 38(4), 1093-1107 (2006).
McCullough, B. D., McGeary, K. A. & Harrison, T. D. Do economic journal archives promote replicable research? Canadian Journal of Economics, 41(4), 1406-1420 (2008).
McNutt, M. Reproducibility. Science, 343, 229-229 (2014).
Mirowski, P. & Sklivas, S. Why econometricians don't replicate (although they do reproduce). Review of Political Economy, 3(2), 146-163 (1991).
Mueller-Langer, F. & Andreoli-Versbach, P. Open access to research data: Strategic delay and the ambiguous welfare effects of mandatory data disclosure. Information Economics and Policy, 42(1) (2018).
Murray, F. & Stern, S. Do formal intellectual property rights hinder the free flow of scientific knowledge?: An empirical test of the anti-commons hypothesis. Journal of Economic Behavior & Organization, 63(4), 648-687 (2007).

Organisation for Economic Co-Operation and Development (OECD): Gross domestic expenditure on R-D by sector of performance and field of science. At <http://stats.oecd.org/Index.aspx?DataSetCode=GERD_SCIENCE> (2016).

Organisation for Economic Co-operation and Development. OECD principles and guidelines for access to research data from public funding. OECD, Paris (2007).

Popper, K. R. The logic of scientific discovery. Hutchinson & Co, London (1959).

Reinhart, C. M. & Rogoff, K. S. Growth in a time of debt. National Bureau of Economic Research. At <http://www.nber.org/papers/w15639.pdf> (2010a).

Reinhart, C. M. & Rogoff, K. S. Growth in a time of debt. American Economic Review Papers and Proceedings, 100(2), 573-578 (2010b).

Reinhart, C. M., Reinhart, V. R. & Rogoff, K. S. Public debt overhangs: Advanced-economy episodes since 1800. Journal of Economic Perspectives, 26(3), 69-86 (2012).

Reinhart, C. M. & Rogoff, K. S. Errata: Growth in a time of debt. Harvard University. At <http://scholar.harvard.edu/rogoff/publications/growth-time-debt> (2013).

Rock, S., Sedo, S. & Willenborg, M. Analyst following and count-data econometrics. Journal of Accounting and Economics, 30(3), 351-373 (2000).

Salas, D. & Raftery, J.P. Econometric issues in testing the age neutrality of health care expenditure. Health Economics, 10(7), 669-671 (2001).

Savage, C. J. & Vickers, A .J. Empirical study of data sharing by authors publishing in PLoS journals. PLoS ONE, 4, e7078 (2009).

Sukhtankar, S. Replications in development economics. Paper presented at the American Economic Association Annual Meeting in Chicago (2017).

Taylor, C. T., Kreisle, N. M. & Zimmerman, P. R. Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in Southern California: Comment. American Economic Review, 100(3), 1269-1276 (2010).

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M. & Frame, M. Data sharing by scientists: Practices and perceptions. PLoS ONE, 6, e21101 (2011).

Vlaeminck, S. Data management in scholarly journals and possible roles for libraries – Some insights from EDaWaX. LIBER Quarterly, 23(1), 48-79 (2013).

Wagstaff, A. & Van Doorslaer, E. Measuring inequalities in health in the presence of multiple-category morbidity indicators. Health Economics, 3(4), 281-291 (1994).

Wible, J. R. Maximization, replication, and the economic rationality of positive economic science, Review of Political Economy, 3(2), 164-186 (1991).

Zhang, L. & Ortmann, A. The effect of the take-option in dictator-game experiments: A comment on Engel's (2011) meta-study. Experimental Economics, 17(3), 414-420 (2014).

Table 1 | Descriptive statistics

	mean	sd	min	max	N
Dependent variable					
Replicated article	0.105		0	1	1,243
Negated replicated article	0.049		0	1	1,243
(Partially) reinforced replicated article	0.056		0	1	1,243
Main variables of interest					
Total cites before publication of replication study	20.89	64.18	0	1,508	1,243
Lag between publication of replicated article and of respective replication	4.851	3.601	0	23	1,243
Journal impact factor	3.516	1.153	2.137	6.033	1,243
Top 5 economics journal	0.512		0	1	1,243
Top 50 university	0.606		0	1	1,243
Mandatory data disclosure policy	0.230		0	1	1,243
Data or program code available	0.169		0	1	1,243
Control variables					
Self-created data	0.124		0	1	1,243
Confidential or proprietary data	0.012		0	1	1,243
Article published in conference proceedings	0.118		0	1	1,243
Number of references	29.28	17.83	0	130	1,243
Number of pages	19.51	10.94	1	65	1,243
Number of authors	2.057	1.097	1	16	1,243
h-index of best author	17.42	12.90	0	106	1,225
Third party funding	0.185		0	1	1,243
Funder's support for data availability	0.598	1.366	0	5	1,243
Year variables					
Publication year 1970-1974	0.006		0	1	1,243
1975-1979	0.039		0	1	1,243
1980-1984	0.043		0	1	1,243
1985-1989	0.106		0	1	1,243
1990-1994	0.091		0	1	1,243
1995-1999	0.109		0	1	1,243
2000-2004	0.207		0	1	1,243
2005-2009	0.223		0	1	1,243
2010-2014	0.175		0	1	1,243
Research areas					
Business & Economics: General	0.764		0	1	1,243
Mathematical Methods	0.102		0	1	1,243
Environmental Sciences	0.036		0	1	1,243
Health Care Sciences	0.091		0	1	1,243
Transportation	0.007		0	1	1,243

Table 2 | Extended descriptive statistics

	Replicated articles		Nonreplicated articles	
	CitesPreReplication mean	*LagReplication* mean	*CitesPreReplication* Mean	*LagReplication* mean
Main variables of interest				
Top5Journal=1	100.57	6.71	24.51	6.10
Top5Journal=0	25.23	3.64	7.40	3.45
Top50University=1	73.07	5.04	19.67	5.07
Top50University=0	26.69	5.17	11.32	4.47
MandatoryDisclosure=1	35.22	3.89	12.06	4.04
MandatoryDisclosure=0	64.66	5.26	17.58	5.07
DataOrCode=1	29.00	4.26	20.15	3.87
DataOrCode=0	65.99	5.22	15.45	5.02
Control variables				
SelfCreatedData=1	13.00	3.93	13.65	4.42
SelfCreatedData=0	66.79	5.23	16.63	4.88
ProceedingsArticle=1	27.40	2.60	6.92	4.33
ProceedingsArticle=0	61.91	5.18	17.62	4.90
Funded=1	118.23	5.67	23.66	5.06
Funded=0	43.29	4.90	14.63	4.77
Year variables				
Pyear1970=1	78.00	8.00	12.43	8.00
Pyear1975=1	3.25	3.50	3.02	3.98
Pyear1980=1	10.29	6.57	11.25	5.04
Pyear1985=1	71.62	5.85	10.73	4.48
Pyear1990=1	49.40	7.13	22.25	8.32
Pyear1995=1	106.57	5.71	21.11	5.75
Pyear2000=1	99.48	5.28	30.04	5.57
Pyear2005=1	24.54	3.96	14.39	3.90
Pyear2010=1	19.25	2.08	5.08	3.17

Table 3 | Marginal effects at the mean after probit regression

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Issues/Sample:	All	All	All	All	All	All	With sci. repl. art.	With neg. repl. art.	With reinf. repl. art.
Dependent variable:	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Scient. repl. art.	Negated repl. art.	Reinf. repl art.
Log total cites before publication of replication study		0.057***	0.057***	0.054***	0.053***	0.053***	0.053***	0.068***	0.040***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.011)	(0.008)
Log lag between publication of replicated article and replication		-0.110***	-0.110***	-0.104***	-0.103***	-0.102***	-0.106***	-0.152***	-0.070***
		(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.024)	(0.016)
Top 5 economics journal			-0.040***	-0.052***	-0.054***	-0.058***	-0.045***	-0.052*	-0.058***
			(0.008)	(0.010)	(0.010)	(0.010)	(0.011)	(0.021)	(0.010)
Log impact factor			-0.079***	-0.097***	-0.097***	-0.096***	-0.095***	-0.053	-0.102***
			(0.016)	(0.017)	(0.017)	(0.018)	(0.018)	(0.043)	(0.013)
Top 50 university				0.042*	0.042*	0.042*	0.048*	0.021	0.047*
				(0.021)	(0.021)	(0.021)	(0.020)	(0.027)	(0.020)
Mandatory data disclosure policy					0.004				
					(0.018)				
Data or program code available						0.013	-0.000	0.014	0.013
						(0.018)	(0.028)	(0.033)	(0.010)
Observations	1,225	1,225	1,225	1,225	1,225	1,225	966	559	714
Pseudo R-squared	0.0763	0.118	0.118	0.125	0.125	0.125	0.143	0.134	0.152
Log Pseudo Likelihood	-382.8	-365.7	-365.7	-362.8	-362.8	-362.7	-288.1	-166.8	-192.3
Journal FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Wald Test Stat., Control Vars. †									
Chi-squared	133.6	23.68	23.68	23.31	23.22	21	29.36	5.497	266
Degrees of freedom	9	9	9	9	9	9	8	8	9
p-value	0.000	0.005	0.005	0.006	0.006	0.013	0.000	0.703	0.000
Wald Test Statistics, Journal FE									
Chi-squared	506,595	4.521e+06	6.570e+06	1.760e+07	1.106e+06	1.310e+09	32,925	7.140e+08	2,232
Degrees of freedom	16	18	18	19	19	20	17	18	11
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE									
Chi-squared	202.2	724.7	724.7	998.1	634.1	740.2	95.28	101.7	60.52
Degrees of freedom	8	8	8	8	8	8	8	8	7
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Robust standard errors clustered at the journal level in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.
† Control variables (not reported): *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Table 4 | Marginal effects at the mean after probit regression: Articles published in or after 2004

	[1]	[2]	[3]	[4]	[5]	[6]
Sample:	≥2004	≥2004	≥2004	≥2004	≥2004	≥2004
Dependent variable:	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article
Log total cites before publication of replication study	0.050*** (0.006)	0.050*** (0.006)	0.050*** (0.005)	0.049*** (0.006)	0.049*** (0.006)	0.050*** (0.006)
Log lag between publication of replicated article and of respective replication	-0.117*** (0.013)	-0.117*** (0.013)	-0.117*** (0.013)	-0.114*** (0.012)	-0.114*** (0.012)	-0.117*** (0.013)
Top-5 economics journal	-0.077*** (0.017)			-0.056*** (0.016)		
AER publication			-0.073*** (0.019)			-0.040*** (0.012)
Log impact factor	-0.077*** (0.021)	-0.077*** (0.021)	-0.076*** (0.021)	-0.075*** (0.022)	-0.075*** (0.022)	-0.076*** (0.022)
Top-50 university	0.034** (0.012)	0.034** (0.012)	0.034** (0.012)	0.034** (0.011)	0.034** (0.011)	0.033** (0.011)
Mandatory data-disclosure policy	0.056*** (0.008)	0.056*** (0.008)	0.079*** (0.013)			
Mandatory data disclosure interacted with AER publication			-0.027* (0.013)			
Data or program code available				0.035 (0.020)	0.035 (0.020)	0.083** (0.025)
Data or program code available interacted with AER publication						-0.066** (0.025)
Observations	560	560	560	560	560	560
Pseudo R-squared	0.188	0.188	0.188	0.187	0.187	0.191
Log Pseudo Likelihood	-131.1	-131.1	-131	-131.2	-131.2	-130.6
Journal FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Wald Test Stat., Control Vars. †						
Chi-squared	600	600	589.5	51.84	51.84	92.60
Degrees of freedom	9	9	9	9	9	9
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Journal FE						
Chi-squared	22645	45018	24362	31503	36217	105719
Degrees of freedom	12	13	12	12	13	12
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE						
Chi-squared	49.39	49.39	48.58	4.338	4.338	5.149
Degrees of freedom	2	2	2	2	2	2
p-value	0.000	0.000	0.000	0.114	0.114	0.0762

Notes: Robust standard errors clustered at the journal level in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.
† Control variables (not reported): *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Table 5 | Ordinary least squares (OLS) regression coefficients

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Issues:	All	All	All	All	All	All	With sci. repl. art.	With neg. repl. art.	With reinf. repl. art.
Dependent variable:	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Scient. repl. art.	Negated repl. art.	Reinf. repl art.
Article published in conference proceedings	-0.006	0.012	0.012	0.007	0.004	0.007	0.008	-0.022	0.013
	(0.026)	(0.029)	(0.029)	(0.029)	(0.031)	(0.029)	(0.030)	(0.104)	(0.037)
Log number of references	-0.002	-0.017	-0.017	-0.014	-0.013	-0.014	-0.016	0.049	-0.061*
	(0.027)	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)	(0.027)	(0.035)	(0.022)
Log number of pages	0.057	0.031	0.031	0.025	0.024	0.025	0.047	-0.039	0.069*
	(0.031)	(0.029)	(0.029)	(0.030)	(0.030)	(0.031)	(0.028)	(0.044)	(0.023)
Number of authors	0.018	0.015	0.015	0.015	0.015	0.015	0.018	0.014	0.019
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.021)	(0.012)
Log H-index of the best author	0.028	0.008	0.008	0.002	0.002	0.002	0.006	0.031	-0.019
	(0.016)	(0.017)	(0.017)	(0.015)	(0.016)	(0.016)	(0.020)	(0.022)	(0.016)
Self-created data	-0.012	0.003	0.003	0.007	0.007	0.007	0.026	0.003	0.004
	(0.014)	(0.015)	(0.015)	(0.016)	(0.016)	(0.017)	(0.015)	(0.036)	(0.024)
Data proprietary according to notes on data&code	-0.012	-0.016	-0.016	-0.014	-0.017	-0.014	-0.087***	-0.091***	-0.003
	(0.091)	(0.076)	(0.076)	(0.080)	(0.082)	(0.080)	(0.018)	(0.018)	(0.099)
Third party funding	-0.012	-0.018	-0.018	-0.015	-0.016	-0.015	-0.051	-0.007	-0.036
	(0.062)	(0.055)	(0.055)	(0.052)	(0.052)	(0.053)	(0.068)	(0.084)	(0.056)
Funder's support for data availability	0.009	0.008	0.008	0.007	0.007	0.007	0.018	-0.002	0.018
	(0.016)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.018)	(0.021)	(0.016)
Log total cites before publication of replication study		0.063***	0.063***	0.061***	0.061***	0.061***	0.062***	0.075***	0.051***
		(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.012)
Log lag between publication of replicated article and replication		-0.109***	-0.109***	-0.106***	-0.105***	-0.106***	-0.109***	-0.159***	-0.084**
		(0.017)	(0.017)	(0.016)	(0.017)	(0.016)	(0.016)	(0.022)	(0.021)
Top 5 economics journal			-0.040***	-0.056***	-0.062***	-0.056***	-0.043**	-0.037	-0.070***
			(0.010)	(0.012)	(0.015)	(0.012)	(0.012)	(0.024)	(0.016)
Log impact factor			-0.081***	-0.103***	-0.104***	-0.103***	-0.094***	-0.039	-0.116***
			(0.017)	(0.020)	(0.019)	(0.020)	(0.021)	(0.037)	(0.021)
Top 50 university				0.051	0.050	0.051	0.058*	0.029	0.056
				(0.026)	(0.026)	(0.026)	(0.026)	(0.029)	(0.031)
Mandatory data-disclosure					0.014				
					(0.021)				
Data or program code available						-0.001	-0.017	0.004	-0.002
						(0.019)	(0.029)	(0.039)	(0.012)
Constant	-0.218**	-0.011	0.134	0.162*	0.158*	0.162*	0.087	0.053	0.241
	(0.060)	(0.066)	(0.076)	(0.070)	(0.072)	(0.072)	(0.114)	(0.102)	(0.133)
Observations	1,225	1,225	1,225	1,225	1,225	1,225	973	563	714
R-squared	0.049	0.074	0.074	0.079	0.080	0.079	0.092	0.092	0.088
Journal FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Wald Test Stat., Control Vars. †									
Chi-squared	7.913	1.765	1.765	1.279	1.146	1.274	22.16	17.69	13.15
Degrees of freedom	9	9	9	9	9	9	9	9	9
p-value	0.000	0.133	0.133	0.302	0.374	0.305	0.000	0.000	0.000
Wald Test Statistics, Journal FE									
Chi-squared	1,143	22,636	12,730	26,923	33,183	1.740e+08	13,648	5.321e+06	122.8
Degrees of freedom	16	18	18	19	19	20	17	18	11
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE									
Chi-squared	17.31	22.06	22.06	30.46	12.81	17.40	5.209	104.3	9.041
Degrees of freedom	8	8	8	8	8	8	8	8	7
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000

Notes: Robust standard errors clustered at the journal level in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

† Control variables: *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Appendix 1 | Overview of the journals, data policies and articles under study

Journal	Impact Factor Rank	Mandatory Data Disclosure Policy	Year / Volume of Enactment	# Articles	# Repl. Articles
JOURNAL OF ECONOMIC LITERATURE	1				
JOURNAL OF FINANCE	2			118	10
QUARTERLY JOURNAL OF ECONOMICS	3			72	10
JOURNAL OF ECONOMIC PERSPECTIVES	4				
TRANSPORTATION RESEARCH PART B-METHODOLOGICAL	5				
JOURNAL OF FINANCIAL ECONOMICS	6			59	9
JOURNAL OF POLITICAL ECONOMY	7	YES	2005/113	72	13
REVIEW OF FINANCIAL STUDIES	8			4	1
ECONOMETRICA	9	YES	2004/72	10	3
JOURNAL OF THE EUROPEAN ECONOMIC ASSOC.	10	YES	2011/9	28	1
REVIEW OF ENVIRONMENTAL ECONOMICS AND POLICY	11				
PHARMACOECONOMICS	12				
AMERICAN ECONOMIC REVIEW	13	YES	2005/95	478	35
ECONOMIC GEOGRAPHY	14				
REVIEW OF ECONOMIC STUDIES	15	YES	2006/73		
JOURNAL OF ECONOMIC GROWTH	16				
BROOKINGS PAPERS ON ECONOMIC ACTIVITY	17	YES	2008/39	6	1
AMERICAN ECONOMIC JOURNAL-APPLIED ECONOMICS	18	YES	2009/1		
JOURNAL OF FINANCIAL STABILITY	19				
VALUE IN HEALTH	20				
AMERICAN ECONOMIC JOURNAL-MACROECONOMICS	21	YES	2009/1		
ECONOMIC POLICY	22			5	1
JOURNAL OF ACCOUNTING & ECONOMICS	23			58	9
JOURNAL OF ECONOMIC GEOGRAPHY	24				
TECHNOL. AND ECONOMIC DEVELOPMENT OF ECONOMY	25				
REVIEW OF ECONOMICS AND STATISTICS	26	YES	2010/92	73	4
ECONOMIC JOURNAL	27	YES	2012/122	10	1
ENERGY ECONOMICS	28				
AMERICAN ECONOMIC JOURNAL-ECONOMIC POLICY	29	YES	2009/1		
EXPERIMENTAL ECONOMICS	30			9	1
TRANSPORT. RES. PART A-POLICY AND PRACTICE	31			4	1
JOURNAL OF ENVIR. ECONOMICS AND MANAGEMENT	32			12	2
ECOLOGICAL ECONOMICS	33			33	4
ANNUAL REVIEW OF ECONOMICS	34				
ECONOMICS & HUMAN BIOLOGY	35				
JOURNAL OF INTERNATIONAL ECONOMICS	36			10	1
JOURNAL OF DEVELOPMENT ECONOMICS	37	YES	2014/111	12	1
ECONOMIC SYSTEMS RESEARCH	38				
FOOD POLICY	39				
JOURNAL OF BUSINESS & ECONOMIC STATISTICS	40	YES	2011/29	34	3
JOURNAL OF AGRARIAN CHANGE	41				
JOURNAL OF POLICY ANALYSIS AND MANAGEMENT	42				
JOURNAL OF HEALTH ECONOMICS	43			72	12
JOURNAL OF TRANSPORT GEOGRAPHY	44				
TRANSPORTATION RESEARCH PART E-LOGISTICS AND TRANSPORTATION REVIEW	45			5	1
HEALTH ECONOMICS	46			41	6
WORLD BANK RESEARCH OBSERVER	47				
JOURNAL OF MONETARY ECONOMICS	48				
JOURNAL OF REGIONAL SCIENCE	49				
JOURNAL OF LABOR ECONOMICS	50	YES	2010/28		
			TOTAL:	1,225	130

Notes: Table based on the 1,225 observations used in the regressions. 23 journals under study that published at least one replicated article in bold.

Appendix 2 | Correlation matrix for the dependent variable and main variables of interest

	Replicated Article	*CitesPre Replication*	*Lag Replication*	*Top5 Journal*	*Impact Factor*	*Top50 University*	*Mandatory Disclosure*	*DataOr Code*
ReplicatedArticle	1.00							
CitesPreReplication	0.21	1.00						
LagReplication	0.02	0.36	1.00					
Top5Journal	-0.03	0.17	0.37	1.00				
ImpactFactor	-0.01	0.11	0.08	0.11	1.00			
Top50University	0.09	0.11	0.08	0.20	0.15	1.00		
MandatoryDisclosure	-0.07	-0.06	-0.12	0.45	-0.12	0.14	1.00	
DataOrCode	-0.02	0.00	-0.12	0.35	-0.10	0.10	0.69	1.00

Appendix 3 | Overview of data guidelines of funding agencies

Funding Agency	[1] Policy/ guideline mentions data sharing	[2] Policy/ guideline specifies how or where to publish data	[3] Policy/ guideline mentions data documen- tation/ metadata	[4] Policy/ guideline requires data manage- ment plan	[5] Policy/ guideline mentions embargo period
National Science Foundation (NSF)	1	0	0	1	1
US National Institutes of Health (NIH)	1	1	1	1	1
National Natural Science Foundation of China (NSFC)	0	0	0	0	0
Deutsche Forschungsgemeinschaft (DFG)	1	1	1	0	0
Bundesministerium für Bildung und Forschung (BMBF)	0	0	0	0	0
Japanese society for the promotion of science (JSPS)	0	0	0	0	0
Japan Science and Technology Agency (JST)	0	0	0	0	0
Russian Foundation for Basic Research (RFBR)	0	0	0	0	0
French National Research Agency (ANR)	1	0	0	0	0
National Research Foundation of Korea (NRF)	0	0	0	0	0
Ministry of Knowledge Economy (MKE)	0	0	0	0	0
Science & Technology Facilities Council (STFC)	1	1	1	1	1
Medical Research Council (MRC)	1	1	1	1	1
Engineering and Physical Science Research Council (EPSRC)	1	1	1	0	1
Biotechnology and Biological Science Research Council	1	1	1	1	1
Natural Environment Research Council (NERC)	1	1	1	1	1
Economic and Social Research Council (ESRC)	1	1	1	1	1
Arts and Humanities Research Council (AHRC)	1	1	1	0	1
National Research Council (CNR)	0	0	0	0	0
Ministry of Education, Universities and Research (MIUR)	0	0	0	0	0
Spanish Ministry of Economy and Competitiveness	0	0	0	0	0
Social Sciences and Humanities Research Council of Canada	1	1	1	1	1
Natural Science and Engineering Research Council of Canada	1	1	1	1	1
Netherlands Organisation for Scientific Research (NWO)	1	1	1	1	1
Swedish Research Council (VR)	1	1	1	0	1
Swedish Foundation for Strategic Research (SSF)	0	0	0	0	0
Swedish Agency for Innovation Systems (VINNOVA)	0	0	0	0	0
Australian Research Council (ARC)	1	1	1	0	1
Norwegian Research Council	1	1	1	1	1
Portuguese Found. Sci. & Techn. (FCT)	1	1	1	1	1
Bill and Melinda Gates Foundation	1	1	1	1	1
Wellcome Trust	1	1	1	1	1
NASA	1	0	1	1	1
Leverhulme Trust	0	0	0	0	0
European Commission (Horizon 2020)	1	1	1	1	1
Swiss National Science Foundation	0	0	0	0	0

Notes: FunderDataSupport is a variable with range from 0 to 5 that indicates the number of data policies and data management tools that external research funders provide to the authors they support. This variable can be thought of as the extent to which external funders have policies that facilitate data availability. To obtain this variable, we analyzed the funding guidelines of 36 research funding bodies worldwide regarding their data management policies. We regard the guidelines as textual data and code if they mention data management, if they specify where and how data should be stored, its terms of access (e.g., on request or public), documentation standards, and if they mention replication studies.

Appendix 4 | Full version of Table 3 (Marginal effects at the mean after probit)

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Issues:	All	All	All	All	All	All	With sci. repl. art.	With neg. repl. art.	With reinf. repl. art.
Dependent variable:	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Repl. article	Scient. repl. art.	Neg. repl. art.	Reinf. repl art.
Article published in conference proceedings	-0.054	-0.043	-0.043	-0.045	-0.046	-0.042	-0.035	-0.003	-0.049
	(0.028)	(0.032)	(0.032)	(0.033)	(0.034)	(0.032)	(0.031)	(0.068)	(0.027)
Log number of references	-0.005	-0.015	-0.015	-0.012	-0.011	-0.011	-0.011	0.046	-0.047*
	(0.029)	(0.025)	(0.025)	(0.026)	(0.026)	(0.026)	(0.027)	(0.034)	(0.020)
Log number of pages	0.067*	0.038	0.038	0.032	0.031	0.030	0.053*	-0.036	0.059**
	(0.034)	(0.028)	(0.028)	(0.029)	(0.029)	(0.030)	(0.026)	(0.047)	(0.018)
Number of authors	0.012	0.009	0.009	0.009	0.009	0.009	0.010	0.008	0.011
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.013)	(0.007)
Log H-index of the best author	0.029*	0.007	0.007	0.004	0.004	0.004	0.006	0.030	-0.014
	(0.014)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.017)	(0.021)	(0.014)
Self-created data	-0.016	0.003	0.003	0.006	0.006	0.005	0.018	0.006	-0.002
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.012)	(0.032)	(0.017)
Data proprietary according to notes on data&code	-0.003	-0.010	-0.010	-0.011	-0.012	-0.015	[dropped]	[dropped]	-0.003
	(0.106)	(0.088)	(0.088)	(0.091)	(0.093)	(0.089)			(0.081)
Third party funding	-0.012	-0.024	-0.024	-0.023	-0.023	-0.023	-0.046	-0.002	-0.067
	(0.050)	(0.043)	(0.043)	(0.041)	(0.041)	(0.041)	(0.053)	(0.055)	(0.063)
Funder's support for data availability	0.009	0.010	0.010	0.009	0.009	0.009	0.016	-0.004	0.025
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.014)	(0.014)	(0.019)
Log total cites before publication of replication study		0.057***	0.057***	0.054***	0.053***	0.053***	0.053***	0.068***	0.040***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.011)	(0.008)
Log lag between publication of replicated article and replication		-0.110***	-0.110***	-0.104***	-0.103***	-0.102***	-0.106***	-0.152***	-0.070***
		(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.024)	(0.016)
Top 5 economics journal			-0.040***	-0.052***	-0.054***	-0.058***	-0.045***	-0.052*	-0.058***
			(0.008)	(0.010)	(0.010)	(0.010)	(0.011)	(0.021)	(0.010)
Log impact factor			-0.079***	-0.097***	-0.097***	-0.096***	-0.095***	-0.053	-0.102***
			(0.016)	(0.017)	(0.017)	(0.018)	(0.018)	(0.043)	(0.013)
Top 50 university				0.042*	0.042*	0.042*	0.048*	0.021	0.047*
				(0.021)	(0.021)	(0.021)	(0.020)	(0.027)	(0.020)
Mandatory data-disclosure					0.004				
					(0.018)				
Data or program code available						0.013	-0.000	0.014	0.013
						(0.018)	(0.028)	(0.033)	(0.010)
Observations	1,225	1,225	1,225	1,225	1,225	1,225	966	559	714
Pseudo R-squared	0.0763	0.118	0.118	0.125	0.125	0.125	0.143	0.134	0.152
Log Pseudo Likelihood	-382.8	-365.7	-365.7	-362.8	-362.8	-362.7	-288.1	-166.8	-192.3
Journal FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Wald Test Stat., Control Vars. †									
Chi-squared	133.6	23.68	23.68	23.31	23.22	21	29.36	5.497	266
Degrees of freedom	9	9	9	9	9	9	8	8	9
p-value	0.000	0.005	0.005	0.006	0.006	0.013	0.000	0.703	0.000
Wald Test Statistics, Journal FE									
Chi-squared	506,595	4.521e+06	6.570e+06	1.760e+07	1.106e+06	1.310e+09	32,925	7.140e+08	2,232
Degrees of freedom	16	18	18	19	19	20	17	18	11
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE									
Chi-squared	202.2	724.7	724.7	998.1	634.1	740.2	95.28	101.7	60.52
Degrees of freedom	8	8	8	8	8	8	8	8	7
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Robust standard errors clustered at the journal level in parentheses. *** *p*<0.001, ** *p*<0.01, * *p*<0.05.
† Control variables: *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Appendix 5 | Marginal effects at the mean after probit regression (by article impact)

	[1]	[2]	[3]	[4]	[5]	[6]
Sample:	Total cites >11	Total cites >28	Total cites >78	Total cites >11 ≥2 issues	Total cites >28 ≥2 issues	Total cites >78 ≥2 issues
Dependent variable:	Replicated article	Replicated article	Replicated article	Replicated article	Replicated article	Replicated article
Log total cites before publication of replication study	0.063***	0.065***	0.048	0.061***	0.064***	0.048
	(0.012)	(0.019)	(0.029)	(0.012)	(0.019)	(0.029)
Log lag between publication of replicated article and of respective replication	-0.114***	-0.120***	-0.125	-0.108***	-0.118***	-0.125
	(0.029)	(0.030)	(0.079)	(0.029)	(0.030)	(0.079)
Top 5 economics journal	-0.096***	-0.094***	-0.091**	-0.097***	-0.097***	-0.091**
	(0.016)	(0.020)	(0.029)	(0.016)	(0.020)	(0.029)
Log impact factor	-0.120***	-0.209***	-0.283***	-0.118***	-0.211***	-0.283***
	(0.028)	(0.047)	(0.082)	(0.027)	(0.046)	(0.082)
Top 50 university	0.059*	0.103***	0.175**	0.062*	0.100**	0.175**
	(0.026)	(0.031)	(0.063)	(0.027)	(0.032)	(0.063)
Data or program code available	0.010	-0.055	-0.034	0.006	-0.057	-0.034
	(0.020)	(0.041)	(0.064)	(0.020)	(0.041)	(0.064)
Observations	890	599	288	848	580	288
Pseudo R-squared	0.102	0.111	0.137	0.097	0.108	0.137
Log Pseudo Likelihood	-317.9	-241.7	-135.8	-298.9	-232.1	-135.8
Journal FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Wald Test Statistics, Control Vars. †						
Chi-squared	10.95	23.88	279.9	145	67.07	279.9
Degrees of freedom	8	8	7	8	8	7
p-value	0.204	0.002	0.000	0.000	0.000	0.000
Wald Test Statistics, Journal FE						
Chi-squared	3.220e+07	1,152	31,060	327	2,471	31,060
Degrees of freedom	19	15	8	10	10	8
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE						
Chi-squared	151	158.6	159.6	151.3	187.1	159.6
Degrees of freedom	8	8	7	8	8	7
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Total cites: 25% percentile: 11; 50% percentile: 28: 75% percentile: 78. Columns [1] – [3] report results using articles published in all 23 journals under study; columns [4] – [6]: 13 journals with at least two issues. Robust standard errors clustered at the journal level in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

† Control variables (not reported): *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Appendix 6 | Marginal effects at the mean after probit regression (by H-index of best author)

	[1]	[2]	[3]	[4]	[5]	[6]
Sample:	BestH >8	BestH >15	BestH >23	BestH >8 ≥2 issues	BestH >15 ≥2 issues	BestH >23 ≥2 issues
Dependent variable:	Replicated article	Replicated article	Replicated article	Replicated article	Replicated article	Replicated article
Log total cites before publication of replication study	0.055***	0.073***	0.067**	0.053***	0.069***	0.061**
	(0.007)	(0.013)	(0.023)	(0.008)	(0.015)	(0.022)
Log lag between publication of replicated article and replication	-0.102***	-0.120***	-0.084	-0.096***	-0.111**	-0.074
	(0.013)	(0.034)	(0.052)	(0.014)	(0.036)	(0.050)
Top 5 economics journal	-0.073***	-0.116***	-0.078	-0.072***	-0.117***	-0.073
	(0.012)	(0.022)	(0.049)	(0.011)	(0.021)	(0.047)
Log impact factor	-0.141***	-0.139***	-0.159**	-0.138***	-0.142***	-0.174***
	(0.020)	(0.039)	(0.053)	(0.019)	(0.038)	(0.052)
Top 50 university	0.044	0.043	0.027	0.046	0.042	0.030
	(0.025)	(0.034)	(0.045)	(0.025)	(0.034)	(0.042)
Data or program code available	-0.016	-0.011	-0.044	-0.020	-0.017	-0.046
	(0.033)	(0.027)	(0.056)	(0.031)	(0.026)	(0.052)
Observations	867	526	243	803	489	238
Pseudo R-squared	0.153	0.180	0.223	0.160	0.181	0.229
Log Pseudo Likelihood	-267.8	-173.8	-84.39	-245.2	-161.5	-80.58
Journal FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Wald Test Statistics, Control Vars. †						
Chi-squared	76.94	25.43	61.24	140.7	43.63	111.4
Degrees of freedom	8	8	8	8	8	8
p-value	0.000	0.001	0.000	0.000	0.000	0.000
Wald Test Statistics, Journal FE						
Chi-squared	174,074	50,147	1,273	475.2	3,923	66.60
Degrees of freedom	17	14	8	9	9	6
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE						
Chi-squared	210.4	128.3	121.1	232.1	189.5	166
Degrees of freedom	8	7	6	8	7	6
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Notes: H-index of the best author: 25% percentile: 8; 50% percentile: 15: 75% percentile: 23. Columns [1] – [3] report results using articles published in all 23 journals under study; columns [4] – [6]: 13 journals with at least two issues. Robust standard errors clustered at the journal level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

† Control variables (not reported): *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Appendix 7 | Marginal effects at the mean after probit regression (by university rank of best author)

	[1]	[2]	[3]	[4]	[5]	[6]
Sample:	BestRank <142	BestRank <36	BestRank <7	BestRank <142 ≥2 issues	BestRank <36 ≥2 issues	BestRank <7 ≥2 issues
Dependent variable:	Replicated article	Replicated article	Replicated article	Replicated article	Replicated article	Replicated article
Log total cites before publication	0.072***	0.074***	0.088***	0.070***	0.073***	0.088***
of replication study	(0.008)	(0.005)	(0.010)	(0.009)	(0.006)	(0.010)
Log lag between publication of	-0.138***	-0.134***	-0.157***	-0.135***	-0.131***	-0.157***
replicated article and replication	(0.019)	(0.012)	(0.029)	(0.021)	(0.013)	(0.029)
Top 5 economics journal	0.017	-0.001	-0.149***	0.014	-0.010	-0.149***
	(0.016)	(0.022)	(0.031)	(0.017)	(0.021)	(0.031)
Log impact factor	0.011	-0.048	-0.155***	0.008	-0.055	-0.155***
	(0.027)	(0.034)	(0.017)	(0.027)	(0.032)	(0.017)
Top 50 university	0.030	[omitted]	[omitted]	0.034	[omitted]	[omitted]
	(0.040)			(0.043)		
Data or program code available	0.003	0.004	-0.015	0.002	0.005	-0.015
	(0.024)	(0.022)	(0.017)	(0.024)	(0.022)	(0.017)
Observations	902	592	243	859	565	243
Pseudo R-squared	0.145	0.230	0.365	0.142	0.223	0.365
Log Pseudo Likelihood	-282.4	-165.7	-65.78	-270.4	-158.3	-65.78
Journal FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Wald Test Statistics, Control Vars. †						
Chi-squared	77.08	54.32	51.94	413.2	73.39	51.94
Degrees of freedom	9	9	7	9	9	7
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Journal FE						
Chi-squared	52,459	68,771	60.77	3,160	648	60.77
Degrees of freedom	15	13	5	10	9	5
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Wald Test Statistics, Year FE						
Chi-squared	304.2	55.02	241.1	326.9	52.43	241.1
Degrees of freedom	8	8	6	8	8	6
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Notes: University rank of best author (better universities have lower rank): 75% percentile: 142; 50% percentile: 36: 25% percentile: 7. Columns [1] – [3] report results using articles published in all 23 journals under study; columns [4] – [6]: 13 journals with at least two issues. Robust standard errors clustered at the journal level in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

† Control variables (not reported): *ProceedingsArticle, LogReferences, LogPages, Authors, LogBestH, SelfCreatedData, ConfidentialData, Funded, FunderDataSupport*.

Appendix 8 | Total number of articles in journals with at least two issues

Number of issues:	**2**	**3**	**4**	**6**	**8**	**9**	**11**	**13**	**29**	Total
AMERICAN ECONOMIC REVIEW									478	478
ECOLOGICAL ECONOMICS			33							33
ECONOMETRICA		10								10
HEALTH ECONOMICS				41						41
JOURNAL OF ACCOUNTING & ECONOMICS					58					58
JOURNAL OF BUSINESS & ECONOMIC STATISTICS		34								34
JOURNAL OF ENVIRONMENTAL ECON. AND MANAG.	12									12
JOURNAL OF FINANCE						118				118
JOURNAL OF FINANCIAL ECONOMICS						59				59
JOURNAL OF HEALTH ECONOMICS							72			72
JOURNAL OF POLITICAL ECONOMY								72		72
QUARTERLY JOURNAL OF ECONOMICS						72				72
REVIEW OF ECONOMICS AND STATISTICS			73							73
Total	12	44	106	41	58	249	72	72	478	1,132

Notes: In our preferred specification, we use 1,225 observations when we consider all journals. Hence, 92.4% of this sample, i.e. 1,132 articles, are published in journals with at least two issues.

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.