American Economic Review: Papers & Proceedings 2017, 107(5): 52–55
https://doi.org/10.1257/aer.p20171032

Replication and Economics Journal Policies[†]

By JAN H. HÖFFLER*

Sharing data and code that underlie empirical studies improves transparency, facilitates comparisons of results, and can help society to more easily build on existing knowledge. For individual scientists, documenting their research well enough so that others can understand it means additional work. There is a widespread fear that data collection, which is often tedious, becomes less attractive if there is an easier strategy to publish, "free-riding" on others' data (Herrnson 1995). On the other hand, sharing data could be a means to attract the interest of others and thus help to receive more citations (King 1995).

Piwowar, Day, and Fridsma (2007) found evidence that cancer microarray clinical trials for which data was available were cited more often, Piwowar and Vision (2013) confirmed this finding with a larger number of studies that created gene expression microarray data, and Bueno de Mesquita et al. (2003) came to similar conclusions for International Studies Research. Thus, for other sciences, there seems to be an indication that sharing research data tends to increase the number of citations for single articles. This can be because readers cite when they re-use the data for their own investigation. Another reason might be that sharing data makes research more reliable and other scientists acknowledge this by citing such investigations more often.

I. What Does the Data Tell Us?

In this article I investigate the impact of the introduction of replication policies for economics journals on citations. The *Journal of Applied Econometrics* has kept a mandatory data archive of all studies published since 1994.[1] After the *American Economic Review* introduced a mandatory online archive for replication material in 2005, a number of journals followed this example. Lamentably, only a minority of journals so far enforce their policies in a way that ensures replicability of most of the empirical work.

II. The Role of Software

Some journals like the *Journal of Applied Econometrics* or the *Journal of the American Statistical Association* introduced special positions in their editorial board responsible for the replicability of published results. It might, however, be too much to expect journals to have experts for every single topic and software package who can test all the submissions. The website ReplicationWiki gives an idea of the variety of software used.[2] For the more than 1,400 studies for which information on software was collected, Stata was used most often by far, in more than 900 studies. It was followed by MATLAB that was used in more than 280 studies, SAS in about 60, GAUSS in 60, Excel in 50, R in 30, FORTRAN in 30, Mathematica in 19, EViews in 18, z-Tree in 16, dynare in 15, RATS in 12, C in 8, C++ in 6, python in 5, and SPSS in 5. In total, 26 software packages or programming languages were found to have been used for more than one study.[3] For a substantial

*University of Göttingen, Humboldtallee 3, 37073 Göttingen (e-mail: jhoeffl@uni-goettingen.de). Many thanks to Areerat Kichkha for the organization of our session at the AEA general meetings, Jack Tatom as discussant, conference participants for a fruitful exchange of thoughts, Xiaohua Yu for his support, and Björn Zakula and Christoph Weißer for their research assistance.

[†] Go to https://doi.org/10.1257/aer.p20171032 to visit the article page for additional materials and author disclosure statement.

[1] Sharing code is encouraged but not required.

[2] http://replication.uni-goettingen.de/wiki/index.php/Category:Article_software_used (accessed January 13, 2017). The wiki website that is described in more detail in Höffler (2017) serves as a database of replications in economics as well as of published empirical studies and provides information on data and code availability, topics, methods, and data and software used. This can especially be used to identify examples for the teaching of methods.

[3] The majority of the listed studies were published in the *American Economic Review*, the four *American Economic Journals*, and the *Journal of Political Economy* in the years

number of studies, more than one software package or programming language was used, which can be a further challenge for replicators.

Journals cooperating to build teams of reproducible research editors could be a way to overcome the obstacles. This could also help to find standards for replicability.[4]

III. How Many Journals Have What Kind of Policy?

For Figure 1, the Thomson Reuters Social Science Citation Index for economics journals for the year 2015 were related to the journals' data availability policies.

In line with previous findings (Vlaeminck 2013 and Vlaeminck and Herrmann 2015),[5] Figure 1 shows that on average journals with a mandatory and enforced data availability policy are cited more than those that have no policy, a voluntary data archive, or the policy that authors should share their data on request. Of the 343 journals for which impact factors are available in 2015, 158 had no data availability policy at all. Twenty-six had a mandatory policy, and replication material could actually be found in online journal archives. One-hundred-ten had voluntary archives. Fifteen asked their authors to make replication material available upon request. Thirty-four both kept voluntary archives and asked authors to make their material available upon request.

IV. Publishing Replications

Sixty-eight different journals published at least one of the 330 replications that were listed in the ReplicationWiki at the time of the completion of this text in January 2017.[6] Thus, there are noticeably more journals that publish replications than there are journals with mandatory data and code archives. Some journals only publish replications of studies that had been published in the same journal, the *Journal of Applied Econometrics* publishes replications of studies from a selection of journals, and some journals publish replications of any study that seems relevant to the reviewers.

Ironically, many of the replications are published without any code or data and are thus themselves not easily replicable. Even the 2011 Report on the *American Economic Review* Data Availability Compliance Project was not replicable as, out of the sample of 220 studies, it did not explicitly list those nine that were actually chosen for replication attempts. It also did not contain the information on which result was found in which case. Unlike this 2011 report, other researchers investigating the replicability of research published in the same journal could not confirm the conclusion that "the *AER* data submission policy has largely been successful" (Chang and Li 2017, Höffler 2016).

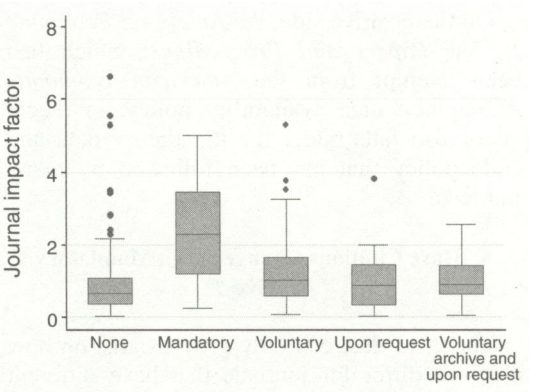

FIGURE 1. JOURNAL IMPACT FACTORS AND DATA AVAILABILITY POLICIES

Notes: Journal impact factors according to the Thomson Reuters Social Science Citation Index for economics journals for the year 2015 and data availability policies in the same year. Journals that officially had a mandatory policy but did not archive replication material for a clear majority of empirical studies in that year are treated as if they had a policy to voluntarily submit to their archive.

2008–2013. The numbers on software just give a rough indication as the sample of studies is biased toward replications and replicated studies, and the information of the software is often missing, especially when it was not easy to obtain. Certain software packages are easier to identify for those who collected the information, and this will have influenced the results.

[4] The project Teaching Integrity in Empirical Research (TIER) already makes a valuable suggestion in this direction. http://www.projecttier.org/tier-protocol/ (accessed January 13, 2017).

[5] The cited authors had also already reported that the number of journals that have introduced data availability policies has been increasing.

[6] http://replication.uni-goettingen.de/wiki/index.php/Category:Replication (accessed January 13, 2017).

On the positive side, the *American Economic Review Papers and Proceedings*, which had been exempt from the *American Economic Association* data availability policy, in recent years also falls under the mandatory data and code policy that has been followed by many journals.

V. Have Citations Changed with Mandatory Archives?

The 2015 data can only give insights on how citations differ for journals that have different data availability policies. Using the 4,267 observations of citation data for 378 journals in the years 1997–2015 helps determine how these citations developed differently with the introduction of mandatory online archives for data and code. Impact factors peaked at 9.24 for the *Journal of Economic Literature* in 2011, and the mean impact factor was measured at 0.93 with a standard deviation of 0.89. A simple regression model with time and journal fixed effects gives a coefficient of 0.55 for implemented mandatory data availability policies, which is significant at the 1 percent level. Time dummy coefficients rise over the years, which will be due to the fact that impact factors were rising as more and more journals were included in the sample. Obviously many factors influencing citations are missing from this regression. As an example, in previous research on citations for single studies it was shown that more established researchers were cited more often, just as studies that had a larger number of coauthors (Bueno de Mesquita et al. 2003, Piwowar, Day, and Fridsma 2007). Such factors should not be systematically associated with the introduction of data availability policies. There is significant variation in the point of time when the archives were introduced for different journals. Given that the journals that introduced the policies were predominantly more cited before they had the data archives, their citations may however follow a time trend different from those of other journals, and this cannot be controlled for with the assembled data. The result could be seen as a first indication that data availability policies may help journals to increase their citation counts.

VI. Open Questions

Among the factors that remain to be investigated is the number of empirical studies published in a journal. Journals that exclusively or nearly exclusively publish theoretical research should not be affected by changes in data availability policies. Their number in the sample is small, so it should not change the results substantially. The degree to which voluntary archives are used varies much. Some are hardly used at all while others are used nearly as if there was a mandatory policy. Taking this into account may have an influence.

It would be relevant to see the extent to which "upon request" policies are efficient. Do researchers who share their data easily when contacted receive a similar amount of extra citations than those that publish their data in journal archives? Many authors also archive data on their personal websites, and this might have a further effect.

Proprietary data is a controversial topic. Those who argue that data policies reduce the incentives for data collection may think that authors will try to use such data more often now. This could be a strategy to reap more benefits of their work without the risk of others using it before they can. Some journals have clauses in their policies that indicate proprietary data should only be used if they are needed for specific insights that could not be gotten from data that could be shared. This could have particular effects on citations.

While, on the one hand, our investigation shows that not all policies are actually enforced, on the other hand, new strategies are developed to ensure replicability. As an incentive for data collection, several journals offer the possibility of publications focused on the description of new datasets, sometimes as an add-on to articles using the data. For such studies it would be interesting to see if both articles receive more citations or if the data article takes citations away from the study in which the data was first used and that would traditionally have been cited.

Journals that want to keep an archive for replication material now have the option to choose among several free and professional data repositories such as the Dataverse,[7] RunMyCode,[8] ResearchCompendia,[9] or the Open Science

[7] http://dataverse.org/ (accessed January 13, 2017).

[8] http://www.runmycode.org/ (accessed January 13, 2017).

[9] http://researchcompendia.org/ (accessed January 13, 2017).

Framework.[10] This reduces the effort needed for data infrastructure.

VII. Outlook

It remains to be seen if the scientific community will find better ways to measure the quality of research than with citations. Some studies attract citations because they contain flaws and are used as common examples why more research transparency is needed.

One proposal of how to further improve replicability of scientific research is the Peer Reviewers' Openness Initiative that calls for reviewers to refuse to comprehensively review empirical work that does not guarantee fully replicable empirical results. While this approach requires a large step from the individual researcher, the Midwest Political Science Association took a collective action. In 2015 they announced that for empirical studies to be published in their *American Journal of Political Science*, replication materials need to be submitted and prior to publication an external research institution will verify if the results can be reproduced.

As long as post publication review does not effectively work as a quality control for empirical research, this might be the gold standard for ensuring replicability of a journal's research. The above results can be seen as a motivation to move in that direction.

REFERENCES

Bueno de Mesquita, Bruce, Nils Petter Gleditsch, Patrick James, Gary King, Claire Metelits, James Lee Ray, Bruce Russett, Havard Strand, and Brandon Valeriano. 2003. "Symposium on Replication in International Studies Research." *International Studies Perspective* 4 (1): 72–107.

Chang, Andrew, and Philipp Li. 2017. "A Preanalysis Plan to Replicate Sixty Economics Research Papers that Worked Half of the Time." *American Economic Review*. https://doi.org/10.1257/aer.p20171034.

Glandon, Philip J. 2011. "Appendix to the Report of the Editor: Report on the *American Economic Review* Data Availability Compliance Project." *American Economic Review* 101 (3): 695–99.

Herrnson, P. S. 1995. "Replication, Verification, Secondary Analysis, and Data Collection in Political Science." *PS: Political Science and Politics* 28 (3): 452–455.

Höffler, Jan H. 2016. "Teaching Replication in Quantitative Empirical Economics." Paper presented at the annual meeting of the Allied Social Science Associations, San Francisco.

Höffler, Jan H. 2017. "ReplicationWiki—Improving Transparency in Social Sciences Research." *D-Lib Magazine* March/April DOI: 10.1045/march2017-hoeffler.

King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* 28 (3): 444–52.

Peer Reviewers' Openness Initiative. 2017. https://opennessinitiative.org/ (accessed January 13, 2017).

Piwowar, Heather A., and Todd J. Vision. 2013. "Data Reuse and the Open Data Citation Advantage." *PeerJ* 1: e175.

Piwowar, Heather A., Roger S. Day, and Douglas B. Fridsma. 2007. "Sharing Detailed Research Data is Associated with Increased Citation Rate." *PLoS ONE* 2 (3): e308.

Vlaeminck, Sven. 2013. "Data Management in Scholarly Journals and Possible Roles for Libraries—Some Insights from EDaWaX." *LIBER Quarterly* 23 (1): 48–79.

Vlaeminck, Sven, and Lisa-Kristin Herrmann. 2015. "Data Policies and Data Archives: A New Paradigm for Academic Publishing in Economic Sciences?" In *New Avenues for Electionic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust*, edited by Birgit Schmidt, Milena Dobreva, 145–55. Washington, DC: IOS Press.

[10] https://osf.io/ (accessed January 13, 2017).