

REPLICATION AND ETHICS IN ECONOMICS: THIRTY YEARS AFTER DEWALD, THURSBY, AND ANDERSON[‡]

What Is Meant by “Replication” and Why Does It Encounter Resistance in Economics?[†]

By MAREN DU VENDACK, RICHARD PALMER-JONES, AND W. ROBERT REED^{*}

In recent years, the lack of reproducibility of scientific research has received much attention (National Academy of Sciences 2015). This has led to increased interest in replications. Four reasons are commonly given for why replication is necessary.

First is “HARKing,” or hypothesizing after the results are known (Kerr 1998). This practice turns hypothesis testing on its head, with theories being developed only after empirical results have been obtained, and then the same empirical results are used to “test” the theories. Second is data mining and estimation manipulation, commonly known as “*p*-hacking,” by which researchers torture the data until they are able to produce the elusive $p < 0.05$. According to Ziliak and McCloskey (2008), the preoccupation with *p*-values represents a corruption of the scientific process by which statistical significance—rather than economic importance—becomes the focus. Third is data error and outright fraud. The popular website Retraction Watch publishes a leaderboard that tracks researchers with the most retractions in academic journals.¹ In good news,

only one economist makes its Top 30 list. The last reason is publication bias (Stanley and Doucouliagos 2012), by which false positives are disproportionately reported in the literature (Ioannides 2005).

Replication is a counter-weight to these challenges as it can expose questionable or fragile analyses, error, and fraud. By re-doing the original data analysis, by adjusting model specifications, exploring the influence of unusual observations, using different estimation methods, and alternative datasets, replication can identify spurious or fragile results. Further, replication has a deterrent effect on questionable practices because the knowledge that their work may be replicated provides an incentive for researchers to take extra precautions in ensuring their results will stand up to independent scrutiny.

I. What Is Replication?

There are different conceptions of what a replication is. Pesaran (2003) distinguishes two types of replications: replications in the “narrow sense” and the “wide sense.” The former consists of checking for errors or computational discrepancies in the original study. The latter investigates whether the results are sustained when using other data. Hamermesh (2007) proposes grouping replications into three categories: pure replications (re-analysis of the same dataset using the same model and estimation methods); statistical replications (use of alternative comparable data, variable constructions, statistical software, or estimation methods); and scientific replications (use of alternative theoretical or conceptual approaches). Clemens (2017) identifies four categories, being careful to distinguish

[‡]*Discussants:* Bruce McCullough, Drexel University; Jack Tatom, Johns Hopkins University; Stan Liebowitz, University of Texas-Dallas; Maren Duvendack, University of East Anglia.

^{*}Duvendack: University of East Anglia, Norwich, NR4 7TJ, United Kingdom (e-mail: m.duvendack@uea.ac.uk); Palmer-Jones: University of East Anglia, Norwich, NR4 7TJ, United Kingdom (e-mail: r.palmer-jones@uea.ac.uk); Reed: University of Canterbury, Private Bag 4800, Christchurch, New Zealand (e-mail: bob.reed@canterbury.ac.nz).

[†]Go to <https://doi.org/10.1257/aer.p20171031> to visit the article page for additional materials and author disclosure statement(s).

¹See <http://retractionwatch.com/the-retraction-watch-leaderboard/>.

between replication and robustness tests. And Hubbard (2016) identifies six different types of replications.

These different conceptualizations are consistent with the National Academy of Science’s (2016) conclusion that there is no consensus in the scientific literature on what is meant by reproducibility, replicability, and robustness. Correspondingly, there is no generally accepted, scientific standard for determining whether previous research is reproducible/replicable. Among other things, this makes it difficult to determine replication rates within a discipline.²

For the purposes of this article, we operationalize replication as any study whose main purpose is to determine the validity of one or more empirical results from a previously published study. Using this broad definition, we identify 188 replications that have been published in the top 50 economics journals since the late 1960s. While the number of replication studies have increased in frequency, they are still relatively uncommon and have not increased in recent years (see Figure 1).

II. What are the Obstacles to Replication in Economics?

Dewald, Thursby, and Anderson’s (1986) seminal study identified a number of reasons why economists do not undertake replications. Replication research is usually not well regarded, as it is commonly considered to be derivative and lacking in methodological and conceptual novelty.³ Further, researchers who replicate other scholars’ research may be suspected of having distrustful and/or malevolent motivations. Replications can generate feelings of abuse, bullying, and persecution in both replicators and replicatees (Duvendack and Palmer-Jones 2013; Bohannon 2014). This creates an environment that inhibits sharing and cooperation. Feigenbaum and Levy (1993) note that journals may not want to publish replications

²For example Patil, Peng, and Leek (2016) and Johnson et al. (2016) produce conflicting replication rates for psychology using the findings of Open Science Collaboration (2015).

³The editors of the *New England Journal of Medicine*, widely considered to be the top medical journal in the world, famously characterized researchers who use other researchers’ data as “research parasites” (Longo and Drazen 2016).

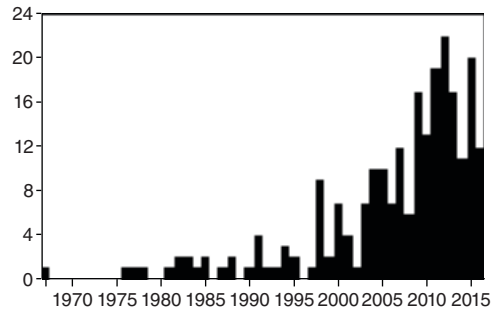


FIGURE 1. NUMBER OF REPLICATIONS PUBLISHED PER YEAR IN ECONOMICS

Source: Updated from Duvendack, Palmer-Jones, and Reed (2015).

because of a belief that they will not be cited as frequently as original research.

A major obstacle to undertaking replication research is the paucity of publication outlets. Table 1 groups the 188 replication studies of Figure 1 by publishing journal. Only 16 journals have ever published more than three replication studies. Five journals account for more than half of all published replications. As some of the journals in Table 1 are specialty journals, the effective number of possible publication outlets for a given replication study can be very limited. This creates a major disincentive to undertake replication research, especially given other obstacles identified above.

One area in which there has been marked improvement is the spread of data availability policies (DAPs), with many journals now adopting such policies. Vlaeminck and Herrmann (2015, p. 153) note:

While McCullough (2009) ... was able to find only 10 journals equipped with such policies, Vlaeminck (2013) was able to find ... 29 journals with [DAPs]. Two years later, we identified 49 economics journals outfitted with such policies.

However, Vlaeminck and Herrmann (2015) also report that enforcement is lax. Of the 49 journals with DAPs, approximately 20 percent did not have a single article for which data was actually available.

Our analysis of 333 economics journals listed in Web of Science confirms that availability

TABLE 1—DISTRIBUTION OF REPLICATIONS ACROSS ECONOMICS JOURNALS

Journal	Frequency percent (number)	Cumulative percent
1. <i>Journal of Applied Econometrics</i>	19.7 (46)	19.7
2. <i>American Economic Review</i>	12.0 (28)	31.6
3. <i>Journal of Human Resources</i>	8.5 (20)	40.2
4. <i>Econ Journal Watch</i>	6.0 (14)	46.2
5. <i>Journal of Development Studies</i>	4.3 (10)	50.4
6. <i>Public Finance Review</i>	4.3 (10)	54.7
7. <i>Empirical Economics</i>	3.8 (9)	58.5
8. <i>Experimental Economics</i>	3.8 (9)	62.4
9. <i>Applied Economics</i>	3.4 (8)	65.8
10. <i>Journal of Political Economy</i>	3.4 (8)	69.2
11. <i>Journal of Economic and Social Measurement</i>	3.0 (7)	72.2
12. <i>Public Choice</i>	2.6 (6)	74.8
13. <i>Economic Inquiry</i>	1.7 (4)	76.5
14. <i>Economics Bulletin</i>	1.3 (3)	77.8
15. <i>Labour Economics</i>	1.3 (3)	79.1
16. <i>Quarterly Journal of Economics</i>	1.3 (3)	80.3

Source: Updated from Duvendack, Palmer-Jones, and Reed (2015).

of data and code lags significantly behind the adoption of DAPs. Only 28 of 333 economics journals regularly⁴ make data and code available (see Table 2). While there are increasing calls for journals to improve data sharing and transparency, there is also significant resistance among researchers, as evidenced by opposition to the adoption of a DAP at top finance journals (Harvey 2014); and the online petition against the data access and research transparency (DA-RT) initiative in political science. The latter has been signed by over 1,000 political science scholars, including 10 former presidents of the American Political Science Association (Dialogue on DA-RT 2015). The issue is important because facilitating access to an original study's data and code greatly lowers the cost of replicating that study.

III. Replication in other Social Sciences

Among the social sciences, psychology, and to a lesser extent, political science have played a leading role in promoting the practice of replication. One of the largest replication exercises ever undertaken was a collaborative project involving 270 researchers replicating 100 experimental

and correlational studies that were published in three top psychology journals (Open Science Collaboration 2015, p. 943). The headline conclusion was that only “39 percent of effects were subjectively rated to have replicated the original result.” This ignited a fierce controversy regarding the interpretation of the project's findings (Gilbert et al. 2016; Anderson et al. 2016).

Another initiative spearheaded in psychology, but which has since gained traction in political science, is that of registered replications. In 2013, concerned that journals were biased against the publication of statistically insignificant results, *Social Psychology* put out a call for a special issue on replications. Invited replication formats included registered replications, described thusly, “Authors submit the introduction, methods, and analysis plan for a replication study or studies. These proposals will be reviewed for their importance and soundness. Once provisionally accepted, the results will be published without regard to the outcome, provided the authors complete the study as proposed” (Nosek and Lakens 2013, p. 59; see also Nosek and Lakens 2014). Relatedly, the journal *Perspective in Psychological Science* announced that a new article type called Registered Replication Reports (RRR) would be a regular feature in the journal. The first RRR was published in 2014 (Simons and Holcombe 2014; Simons, Holcombe, and Spellman 2014).

⁴“Regularly” means that at least 50 percent of empirical articles in recent issues of the journal supply their data and code.

TABLE 2—ECONOMICS JOURNALS THAT REGULARLY PUBLISH DATA AND CODE

1. <i>Agricultural Economics</i>
2. <i>American Economic Journal: Applied Economics</i>
3. <i>American Economic Journal: Economic Policy</i>
4. <i>American Economic Journal: Macroeconomics</i>
5. <i>American Economic Journal: Microeconomics</i>
6. <i>American Economic Review</i>
7. <i>Brookings Papers on Economic Activity</i>
8. <i>Econometrica</i>
9. <i>Econometrics Journal</i>
10. <i>Economic Journal</i>
11. <i>Economics: The Open-Access, Open-Assessment E-Journal</i>
12. <i>Energy Economics</i>
13. <i>European Economic Review</i>
14. <i>Explorations in Economic History</i>
15. <i>Jahrbücher für Nationalökonomie und Statistik/ Journal of Economics and Statistics</i>
16. <i>Journal of Applied Econometrics</i>
17. <i>Journal of Development Economics</i>
18. <i>Journal of Economic Growth</i>
19. <i>Journal of Economic Perspectives</i>
20. <i>Journal of Labor Economics</i>
21. <i>Journal of Law and Economics</i>
22. <i>Journal of Political Economy</i>
23. <i>Journal of the European Economic Association</i>
24. <i>Review of Economic Dynamics</i>
25. <i>Review of Economic Studies</i>
26. <i>Review of Economics and Statistics</i>
27. <i>Review of International Organizations</i>
28. <i>Studies in Nonlinear Dynamics and Econometrics</i>

Source: Updated from Duvendack, Palmer-Jones, and Reed (2015).

Since then, the concept of accepting articles based solely on importance of research question and quality of research design, before research results are known, has rapidly expanded to include a large number of journals in psychology, political science, and other disciplines, including many of the top journals in the fields. The Center for Open Science keeps a running count of the number of journals that have adopted registered reports.⁵ At the time of this writing, there were 41 journals. No economics journals appear on the list.

IV. Conclusion

Economics has made some advances in promoting the practice of replication. In addition

⁵An updated Google Docs spreadsheet can be found here: https://docs.google.com/spreadsheets/d/1D4_k-8C_UENTRtbPzXfhjEyu3BfLxdOsn9j-otrO870/edit#gid=0.

to those mentioned above, noteworthy efforts include Camerer et al. (2016), whose teams replicated 18 studies in experimental economics. While similar to Open Science Collaboration (2015) in many respects, they report a higher replication rate.⁶ Chang and Li (2015) and Duvendack, Palmer-Jones, and Reed (2015) both measure replication rates in economics, with both reporting relatively low rates of replication success. Also noteworthy is the Impact Evaluation Replication Programme of 3ie which funds replications of important papers in the area of development economics (International Initiative for Impact Evaluation 2016). The websites Replication in Economics wiki⁷ and The Replication Network⁸ provide updated information on other replication efforts in economics. Overall, however, the practice of replication in economics lags behind a number of other fields. Whether this is because the problems that plague those disciplines are less severe in economics, or because economics is more resistant to replications, is arguable.

REFERENCES

- Anderson, Christopher J., Štěpán Bahník, Michael Barnett-Cowan, Frank A. Bosco, Jesse Chandler, Christopher R. Chartier, Felix Cheung, et al. 2016. “Response to Comment on ‘Estimating the Reproducibility of Psychological Science.’” *Science* 351 (6277): 1037–39.
- Bohannon, John. 2014. “Replication Effort Provokes Praise—and ‘Bullying’ Charges.” *Science* 344 (6186): 788–89.
- Camerer, Colin F., Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, et al. 2016. “Evaluating Replicability of Laboratory Experiments in Economics.” *Science* 351 (6280): 1433–36.
- Chang, Andrew C., and Phillip Li. 2015. “Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say ‘Usually Not.’” Board of Governors of the Federal Reserve System Finance and Economics Discussion Series Paper 2015–83.

⁶A similar team, composed of many economists, is currently engaged on a related project (The Social Sciences Replication Project 2016).

⁷See http://replication.uni-goettingen.de/wiki/index.php/Main_Page.

⁸See <https://replicationnetwork.com/>.

- Clemens, Michael A.** 2017. "The Meaning of Failed Replications: A Review and Proposal." *Journal of Economic Surveys* 31 (1): 326–42.
- Dewald, William G., Jerry G. Thursby, and Richard G. Anderson.** 1986. "Replication in Empirical Economics: The Journal of Money, Credit and Banking Project." *American Economic Review* 76 (4): 587–603.
- Dialogue on DA-RT.** 2015. *Petition to Delay DA-RT Implementation*. https://docs.google.com/forms/d/e/1FAIpQLSc_cmfNLxUeUY-F6ElgtvNHtnfQaoEY0-T_ocpcu9KUXduj0cA/viewform (accessed December 2, 2016).
- Duvendack, Maren, and Richard W. Palmer-Jones.** 2013. "Replication of Quantitative Work in Development Studies: Experiences and Suggestions." *Progress in Development Studies* 13 (4): 307–22.
- Duvendack, Maren, Richard W. Palmer-Jones, and W. Robert Reed.** 2015. "Replications in Economics: A Progress Report." *Econ Journal Watch* 12 (2): 164–91.
- Feigenbaum, Susan, and David M. Levy.** 1993. "The Market for (Ir)reproducible Econometrics." *Social Epistemology* 7 (3): 215–32.
- Gilbert, Daniel T., Gary King, Stephen Pettigrew, and Timothy D. Wilson.** 2016. "Comment on 'Estimating the Reproducibility of Psychological Science.'" *Science* 351 (6277): 1037.
- Hamermesh, Daniel S.** 2007. "Viewpoint: Replication in Economics." *Canadian Journal of Economics* 40 (3): 715–33.
- Harvey, Campbell R.** 2014. "Reflections on Editing the Journal of Finance, 2006–2012." In *Secrets of Economics Editors*, edited by Michael Szenberg and Lall Ramrattan, 67–82. Cambridge: MIT Press.
- Hubbard, Raymond.** 2016. *Corrupt Research: The Case for Reconceptualizing Empirical Management and Social Science*. Thousand Oaks: SAGE Publications, Inc.
- International Initiative for Impact Evaluation.** 2016. Impact Evaluation Replication Programme. <http://www.3ieimpact.org/en/evaluation/impact-evaluation-replication-programme/> (accessed December 2, 2016).
- Ioannidis, John P.A.** 2005. "Why Most Published Research Findings Are False." *PLoS Med* 2: 124.
- Johnson, Valen E., Richard D. Payne, Tianying Wang, Alex Asher, and Soutrik Mandal.** 2016. "On the Reproducibility of Psychological Science." *Journal of the American Statistical Association*. <http://dx.doi.org/10.1080/1621459.2016.1240079>.
- Kerr, Norbert L.** 1998. "HARKing: Hypothesizing After the Results Are Known." *Personality and Social Psychology Review* 2 (3): 196–217.
- Longo, Dan L., and Jeffrey M. Drazen.** 2016. "Data Sharing." *New England Journal of Medicine* 374: 276–77.
- McCullough, B. D.** 2009. "Open Access Economics Journals and the Market for Reproducible Economic Research." *Economic Analysis and Policy* 39 (1): 117–26.
- National Academy of Sciences.** 2016. *Statistical Challenges in Assessing and Fostering the Reproducibility of Scientific: Summary of a Workshop*. Washington, DC: National Academies Press.
- Nosek, Brian A., and Daniël Lakens.** 2013. "Replications of Important Results in Social Psychology: Special Issue of Social Psychology." *Social Psychology* 44 (1): 59–60.
- Nosek, Brian A., and Daniël Lakens.** 2014. "Registered Reports: A Method to Increase the Credibility of Published Results." *Social Psychology* 45 (3): 137–41.
- Open Science Collaboration.** 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349 (6251).
- Patil, Prasad, Roger D. Peng, and Jeffrey T. Leek.** 2016. "What Should Researchers Expect When They Replicate Studies? A Statistical View of Replicability in Psychological Science." *Perspectives on Psychological Science* 11 (4): 539–44.
- Pesaran, Hashem.** 2003. "Introducing a Replication Section." *Journal of Applied Econometrics* 18 (1): 111.
- Simons, Daniel J., and Alex O. Holcombe.** 2014. "Registered Replication Reports." *Observer* 27 (3).
- Simons, Daniel J., Alex O. Holcombe, and Barbara A. Spellman.** 2014. "An Introduction to Registered Replication Reports at *Perspectives on Psychological Science*." *Perspectives on Psychological Science* 9 (5): 552–55.
- Social Sciences Replication Project.** 2016. "Social Sciences Replication Project." <http://www.socialsciencesreplicationproject.com/> (accessed December 2, 2016).
- Stanley, T. D., and Hristos Doucouliagos.** 2012. *Meta-Regression Analysis in Economics and*

- Business*. London: Routledge.
- 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?” *EconStor Open Access Articles 2015*: 145–55.
- Ziliak, Stephen T., and Deirdre N. McCloskey.** 2008. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. Ann Arbor: University of Michigan Press.