Harvard Data Science Review • Issue 5.3, Summer 2023

Reinforcing Reproducibility and Replicability: An Introduction

Lars Vilhuber[1] Ian Schmutte[2] Aleksandr Michuda[3,4,5] Marie Connolly[6]

[1]Department of Economics, ILR School, Cornell University, Ithaca, New York, United States of America,

[2]Department of Economics, Terry College of Business, University of Georgia, Athens, Georgia, United States of America,

[3]Center for Data Science for Enterprise and Society, Cornell University, Ithaca, New York, United States of America,

[4]Department of Information Science, Cornell Ann S. Bowers College of Computing and Information Science, Cornell University, Ithaca, New York, United States of America,

[5]Dyson School of Applied Economics and Management, College of Agriculture and Life Sciences and Cornell S.C. Johnson College of Business, Cornell University, Ithaca, New York, United States of America,

[6]Department of Economics, School of Management, Université du Québec à Montréal, Montreal, Quebec, Canada

Published on: Jul 27, 2023

DOI: https://doi.org/10.1162/99608f92.9ba2bd43

The purpose of scientific publishing is the dissemination of robust research findings, exposing them to the scrutiny of peers. The key to this endeavor is documenting the provenance of those findings. Scientific practices during the course of research and subsequent publication, peer review, and dissemination practices and tools, all interact to (hopefully) enable a meaningful discourse about the veracity of scientific claims. However, while all practices and tools contribute to the final output, some are less often discussed than others, and perceptions, usage, and acceptance differ in myriad ways across disciplines. In this special theme, and in a subsequent column called "Reinforcing Reproducibility and Replicability,"we will explore these topics, with expert providers and expert users providing their input. While we will start within the economics discipline in this special theme, the column will not be as narrowly focused, providing context and voice from other disciplines over time.

Whether or not one actually believes there is a "replication crisis" (Fanelli, 2018), some doubts have been expressed in recent years about the reliability of research. Partially in response, there has been an increased emphasis on various methods that support improved provenance documentation. In the social sciences, this includes preregistration (Nosek et al., 2018, 2019), pre-analysis plans (Banerjee et al., 2020; Olken, 2015), registered reports (Chambers, 2014; Hardwicke & Ioannidis, 2018; *Journal of Development Economics*, 2019), greater availability of working papers and preprints across disciplines other than economics, statistics, and physics (Vilhuber, 2020), and increasingly more stringent journal policies surrounding data and code availability, including active review and verification of replication packages (Christian et al., 2018, 2020; Editors, 2021; Vilhuber, 2019).

A bit of terminology first. The terms 'reproducibility,' 'replicability,' and even 'transparency' are not defined universally the same way. We adopt in the special theme and later for the new column the National Academies of Sciences, Engineering, and Medicine (NASEM) definition of [*computational*] *reproducibility* as "obtaining consistent results using the same input data, computational steps, methods, and code, and conditions of analysis" and *replicability* as "obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data" (NASEM, 2019, Chapter 3). A key component of the current landscape, and what will be a recurring topic in this column, are 'replication packages,' which here are defined as those materials (data, computer code, and instructions) linked to a specific publication that facilitate the replication of the manuscript's results by others, but should also be computationally reproducible. Together with the actual manuscript, typically preserved or published elsewhere, they constitute the 'research compendium' (Buckheit & Donoho, 1995). The focus here on the infrastructure surrounding reproducibility—more so than replicability—is intentional; after all, if an article's methods are not even reproducible, why bother attempting to replicate or extend the research in that article?

The verification of replication packages, which includes not just checks of the computational reproducibility of the provided materials but also documented data provenance and completeness of such materials, is not a

magical solution that will solve the 'replicability crisis.' Replication packages may be reproducible, but wrong (see, e.g., the recent discussion surrounding Simonsohn et al., 2021). Verification also faces educational and procedural barriers. Should journals, which act at the tail end of the scientific production process, be the verifiers of reproducibility, as some have been doing (Christian et al., 2018; Vilhuber, 2021), or should verification be a natural part of the post-publication assessment by the scientific community, with nonreproducible articles being cited less (as claimed by Hamermesh, 2007) or being retracted (Journal of Finance, 2021)? Should scientists' work be reproducible at every stage of the research process, even prior to submission to journals, and what does that imply for funding, technical infrastructure, and the training of undergraduate and graduate students?

The consensus on answers to these questions is still emerging and needs to be discussed by all researchers in the discipline, because such a consensus will guide how disciplinary and interdisciplinary research is conducted. Most discussions on these topics, however, occur in workshops and conferences that are not the core disciplinary conferences attended by the typical social scientist. For instance, the Research Data Alliance (RDA) plenaries, CODATA (Committee on Data of the International Science Council) conferences, or conferences that cater to information specialists, data scientists, librarians, and so on, are rarely attended by disciplinary specialists.

We attempt to remedy this lack of exposure. Since August 2022, we have been organizing an extended conference via a series of webinars called the Conference on Reproducibility and Replicability in Economics and Social Sciences (CRRESS). The goal of CRRESS is to make the topics described above accessible to all researchers by pulling them out of specialized conferences, and making them available to a broad audience, through a consistent and logical sequence of sessions. The topics covered were selected to inform researchers about themes, tools, infrastructure, and approaches that are not typically known, taught, or learned in current or past disciplinary curricula in the social sciences. The recordings of each hour-long panel discussion are available.[1] Presenters could also submit a written record of their discussion, many of which will now appear as part of this special theme and in subsequent columns.

The first panel within CRRESS discussed whether economics journals should be the institutions responsible for verifying reproducibility. The panel, moderated by an active data editor (Vilhuber), consisted of editors-in-chief of various journals in economics. All were in favor of the ultimate goals of reproducibility of scientific articles, but had differing views on the role of journals in that context. Both Toni Whited (2023; *Journal of Financial Economics*) and Tim Salmon (2023; *Economic Inquiry*) contribute their thoughts on the topic to this issue. The discussion and the various viewpoints are useful to authors as well as to journal editors seeking guidance on this key question.

A later CRRESS session on the status and acceptance of reproducibility also emphasized the role that journals, and in particular society journals, play in economics, sociology, and political science. Hilary Hoynes (2023) reflects on the current status in economics in this theme, whereas the situation in sociology and political

science will appear in a future column. A key theme there, however, is that reproducibility is not just a top-down topic dictated by journals and society leadership but also one that has a very strong bottom-up component. Other topics include the tricky interaction of reproducibility and transparency with the use of confidential data.

CRRESS explored parts of the research lifecycle that explain the bottom-up component. Ethics approval is usually obtained from ethics committees or institutional review boards at the start of a project, and may hinder reproducibility in some cases. But late-stage consent withdrawal may also impact the ability to conduct reproducible research. The CRRESS session on how reproducibility and research ethics interacted will echo in a future column.

Creating reproducible and transparent research requires training academic personnel in appropriate tools. One piece of transparent research is properly accounting for data provenance, and data citations are key to this (Data Citation Synthesis Group, 2014). However, data provenance and data citation practices are all too often neglected in the training of social scientists. Diego Mendez-Carbajo and Alejandro Dellachiesa (2023), in this theme, explore the training of undergraduates in data provenance and data citations, reporting on the experience from several assignments in an undergraduate economics class. Richard Ball (2023) presents a series of feasible exercises introducing reproducible methods to economics (or social science) undergraduates. In that same CRRESS session, one of us (Vilhuber) reported on the employment and training of undergraduates as part of the reproducibility verification process at the American Economic Association (AEA), which is published elsewhere (Vilhuber et al., 2022). Graduate education, of course, is just as important, and was the topic of the last CRRESS session of 2022–2023. It will be the topic of a future column.

Hoynes (2023) also highlights the importance of confidential data. Often seen as an impediment to broad reproducibility, we nevertheless observe many ways in which confidential data can be part of a reproducible research process (for an overview, see the discussion in Vilhuber, 2023). Paulo Guimarães (2023) describes how the research laboratory of the Banco do Portugal, the country's central bank, supports reproducible and accessible analysis of highly confidential data through a set of tools, infrastructure, and processes. Christophe Pérignon and coauthors have also demonstrated how reproducibility services can provide value in the context of confidential data when they have persistent and possibly privileged access (Pérignon et al., 2019). Pérignon and others presented more generally on how verification services work, both when data are open and when they are confidential. These verification services will be the topic of future columns, when we will explore this fertile area in a variety of contexts relevant to the broader social sciences and in other disciplines.

Empirical social scientists do not work alone. They work within institutions, rely on many infrastructure components along the way, and are often funded by sponsors. What role do they play in enabling, supporting, or even requiring reproducible research? Graham MacDonald (2023) describes the role of open data and open science in a nonacademic research institution (the Urban Institute), finding both challenges and opportunities, including the challenge of hiring qualified researchers (and thus the importance of undergraduate education in

universities). Courtney Butler (2023) explores how a federal reserve bank can balance its primary objectives with limited resources to make sharing of replication packages easier, and the institution's research practices more transparent. There is an increasing interest in providing such internal services within research institutions, and we provide several more case studies in future columns. The role of funders in this process, and the connected role of research policy, was discussed in CRRESS sessions as well, and will appear in future columns.

The complete CRRESS collection (short articles, videos, and presentation slides) are meant to serve as a persistent resource for social scientists seeking guidance on how to understand and implement reproducible and replicable research, across multiple fields and research phases, and independent of the journal where their own work may end up being published. Readers of this special theme and of future column contributions will gain insights into the full gamut of topics related to the initiation of research, the conduct of research, the preparation of research for publication, and possibly the post-publication scrutiny related to reproducibility and replicability.

Videos From CRRESS Webinar Series

The articles in this special theme are continuations from the authors' panel discussions during the CRRESS webinar series. Below are links to the authors' articles and videos containing the corresponding panel discussions.

Whited (2023): "Costs and Benefits of Reproducibility in Finance and Economics": https://youtu.be/-dc4xxCIeqQ?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=1060

Salmon (2023): "The Case for Data Archives at Journals": https://youtu.be/-dc4xxCIeqQ?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=517

Hoynes (2023): "Reproducibility in Economics: Status and Update": https://youtu.be/WRwxOM15Zgk?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=1101

Mendez-Carbajo & Dellachiesa (2023): "Data Citations and Reproducibility in the Undergraduate Curriculum": https://youtu.be/DkSkp5svRY4?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=2

Ball (2023): "'Yes We Can!': A Practical Approach to Teaching Reproducibility to Undergraduates": https://youtu.be/DkSkp5svRY4?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=730

Guimarães (2023): "Reproducibility With Confidential Data: The Experience of BPLIM": https://youtu.be/ChR_0_zmQwk?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=988

MacDonald (2023): "Open Data and Code at the Urban Institute": https://youtu.be/Rvpy49rjGeQ?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=192

Butler (2023): "Publishing Replication Packages: Insights From the Federal Reserve Bank of Kansas City": https://youtu.be/Rvpy49rjGeQ?list=PLdcNmwWYeA7XY35YV9zV8zPTbE7twjz4S&t=2100

Acknowledgments

The Conference on Reproducibility and Replicability in Economics and the Social Sciences (CRRESS) webinar series is funded by National Science Foundation Grant #2217493.

Disclosure Statement

Lars Vilhuber, Ian Schmutte, Aleksandr Michuda, and Marie Connolly have no financial or non-financial disclosures to share for this article.

References

Ball, R. (2023). "Yes We Can!": A practical approach to teaching reproducibility to undergraduates. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.9e002f7b

Banerjee, A., Duflo, E., Finkelstein, A., Katz, L., Olken, B., & Sautmann, A. (2020). *In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics* (No. w26993; p. w26993). National Bureau of Economic Research. https://doi.org/10.3386/w26993

Buckheit, Jonathan B., and David L. Donoho. 1995. WaveLab and reproducible research. In A. Antoniadis & G. Oppenheim (Eds.), *Wavelets and Statistics* (pp. 55–81). Springer. https://doi.org/10.1007/978-1-4612-2544-7_5.

Butler, C. (2023). Publishing replication packages: Insights from the Federal Reserve Bank of Kansas City. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.aba61304

Chambers, C. (2014, November 13). Registered Reports: A step change in scientific publishing. *Reviewers' Update*. https://www.elsevier.com/reviewers-update/story/innovation-in-publishing/registered-reports-a-step-change-in-scientific-publishing

Christian, T.-M., Gooch, A., Vision, T., & Hull, E. (2020). Journal data policies: Exploring how the understanding of editors and authors corresponds to the policies themselves. *PLOS ONE*, *15*(3), Article e0230281. https://doi.org/10.1371/journal.pone.0230281

Christian, T.-M., Lafferty-Hess, S., Jacoby, W., & Carsey, T. (2018). Operationalizing the replication standard: A case study of the data curation and verification workflow for scholarly journals. *International Journal of Digital Curation*, *13*(1). https://doi.org/10.2218/ijdc.v13i1.555

Data Citation Synthesis Group. (2014). *Joint declaration of data citation principles* (M. Martone, Ed.). FORCE11. https://doi.org/10.25490/a97f-egyk

Fanelli, D. (2018). Opinion: Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences*, *115*(11), 2628–2631. https://doi.org/10.1073/pnas.1708272114

Guimarães, P. (2023). Reproducibility with confidential data: The experience of BPLIM. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.54a00239

Hamermesh, D. S. (2007). Viewpoint: Replication in economics. *Canadian Journal of Economics*, *40*(3), 715–733. https://doi.org/10.1111/j.1365-2966.2007.00428.x

Hardwicke, T. E., & Ioannidis, J. P. A. (2018). Mapping the universe of registered reports. *Nature Human Behaviour*, *2*(11), Article 11. https://doi.org/10.1038/s41562-018-0444-y

Hoynes, H. (2023). Reproducibility in economics: Status and update. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.80a1b88b

Journal of Development Economics. (2019, July 17). *Registered Reports at JDE: Lessons learned so far*. Retrieved June 14, 2020, from https://www.journals.elsevier.com/journal-of-development-economics/announcements/registered-reports-at-jde

Journal of Finance. (2021). Retracted: Risk management in financial institutions. *The Journal of Finance*, *n/a*(n/a). https://doi.org/10.1111/jofi.13064

MacDonald, G. (2023). Open data and code at the Urban Institute. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.a631dfc5

Mendez-Carbajo, D., & Dellachiesa, A. (2023). Data citations and reproducibility in the undergraduate curriculum. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.c2835391

National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and replicability in science*. National Academies Press. https://doi.org/10.17226/25303

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., Veer, A. E. van ’t, & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815–818. https://doi.org/10.1016/j.tics.2019.07.009

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, *29*(3), 61–80. https://doi.org/10.1257/jep.29.3.61

Pérignon, C., Gadouche, K., Hurlin, C., Silberman, R., & Debonnel, E. (2019). Certify reproducibility with confidential data. *Science*, *365*(6449), 127–128. https://doi.org/10.1126/science.aaw2825

Salmon, T. C. (2023). The case for data archives at journals. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.db2a2554

Simonsohn, U., Nelson, L., Simmons, J., & Anonymous. (2021, August 17). *[98] Evidence of fraud in an influential field experiment about dishonesty*. Data Colada. https://datacolada.org/98

Supporting computational reproducibility through code review. (2021). *Nature Human Behaviour*, *5*(8), 965–966. https://doi.org/10.1038/s41562-021-01190-w

Vilhuber, L. (2019). Report by the AEA Data Editor. *AEA Papers and Proceedings*, *109*, 718–729. https://doi.org/10.1257/pandp.109.718

Vilhuber, L. (2020). Reproducibility and Replicability in Economics. *Harvard Data Science Review*, *2*(4). https://doi.org/10.1162/99608f92.4f6b9e67

Vilhuber, L. (2021). Report by the AEA Data Editor. *AEA Papers and Proceedings*, *111*, 808–817. https://doi.org/10.1257/pandp.111.808

Vilhuber, L. (2023). Reproducibility and transparency versus privacy and confidentiality: Reflections from a data editor. *Journal of Econometrics*, *235*(2), 2285–2294. https://doi.org/10.1016/j.jeconom.2023.05.001

Vilhuber, L., Son, H. H., Welch, M., Wasser, D. N., & Darisse, M. (2022). Teaching for large-scale reproducibility verification. *Journal of Statistics and Data Science Education*, *30*(3), 274–281. https://doi.org/10.1080/26939169.2022.2074582

Whited, T. (2022). *Comments on reproducibility in finance and economics*. *Harvard Data Science Review, 5*(3). https://doi.org/10.1162/99608f92.63de8e58

Footnotes

1. The conference website is https://labordynamicsinstitute.github.io/crress/, and all recordings are available for free at https://youtu.be/-dc4xxCIeqQ. ↩