Journal of Behavioral and Experimental Economics 117 (2025) 102410

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

Replication code as a cornerstone of the credibility revolution 2.0

Jan Marcus

Freie Universität Berlin, Garystr. 21, Berlin 14195, Germany

ARTICLE INFO

Keywords:
Replication code
Reproducibility
Credibility revolution 2.0
Replicability
Code availability

ABSTRACT

This opinion piece argues that code sharing is a key yet underutilized component of research transparency and the ongoing "Credibility Revolution 2.0." It outlines the benefits of providing replication code for both individual researchers and the broader scientific community, and explains why alternative methods of documenting the details of the research process fall short. The opinion piece concludes with concrete recommendations for how different groups in the scientific community can help make code sharing the norm in empirical research.

1. Setting the stage: code sharing in the credibility revolution 2.0

The public availability of the details of the research process is one of the four criteria distinguishing scientific research from non-science, as outlined by King et al. (1994). Without public availability of research details, it is impossible to thoroughly assess a study's internal validity, critique it on solid grounds, or reproduce and replicate its findings. A straightforward and practical way to adhere to this defining feature of science is to provide replication code (Claerbout & Karrenbach, 1992)—computer instructions enabling the reproduction of a scientific study's numbers, figures, and tables.

This opinion piece argues that code sharing is an essential and low-hanging fruit in the broader toolbox of research transparency practices and the ongoing "Credibility Revolution 2.0". While the first Credibility Revolution (Angrist & Pischke, 2009; 2010) centered on better and more transparent research designs to identify causal effects, the emerging "Credibility Revolution 2.0" places a stronger emphasis on transparency throughout the entire research process. It seeks to make more transparent the researchers' degrees of freedom—i.e., the many often subtle decisions made during the research process, such as variable selection and construction, sample exclusions, and how results are presented. At the same time, it aims to reduce these degrees of freedom, which is valuable because it helps limit *p*-hacking—the practice of selectively reporting results that reach statistical significance. *P*-hacking can lead to misleading conclusions and non-replicable findings, undermining the credibility of scientific research. This is particularly problematic when such findings inform evidence-based policymaking or the testing of scientific theories. Code sharing contributes to this goal by making researchers' analytic choices transparent and allowing others to easily re-run the analysis, test alternative specifications, and assess the robustness of the results. Code sharing complements other research transparency tools like pre-registrations with pre-analysis plans (see, e.g., Imai et al., 2025), data sharing, specification curve analyses, etc. (see Miguel, 2021, for an overview).

This opinion piece highlights the benefits of (replication) code provision for researchers and the scientific community, argues that it should be standard in empirical research, and explains why alternative ways of sharing research process details are ultimately inadequate. It concludes with concrete steps for making code sharing the norm.

Before proceeding, I would like to clarify some key terms. Following Dreber and Johannesson (2025), I use the term *reproducibility* to refer broadly to obtaining the same results using the original data, while *computational reproducibility*, a specific type of reproducibility, describes the ability to duplicate results using the same data *and* code; in contrast, *replicability* involves confirming results with new data. Although replication code primarily supports (computational) reproducibility, it also facilitates replication efforts. I use the term replication code—both because it is more common and to emphasize its importance for both reproducibility and replicability.

2. Code sharing benefits researchers and science

The key benefit of code sharing for the scientific system—and for the society—lies in enhanced transparency and reproducibility. By providing their code, researchers enable others to critically assess the

I would like to thank the editor Levent Neyse, two anonymous referees, Lukas Fink, Nicolas Frink, Toivo Glatz, and Jessica Rohmann for their valuable comments on earlier versions of this opinion piece.

E-mail address: jan.marcus@fu-berlin.de.

https://doi.org/10.1016/j.socec.2025.102410
Received 29 April 2025; Received in revised form 11 June 2025; Accepted 26 June 2025
Available online 27 June 2025

methodology and verify the findings. This practice not only strengthens the credibility of scientific claims but also helps to uncover errors that might otherwise go unnoticed. In an era of increasing misinformation and skepticism toward scientific expertise, transparent research practices are essential for maintaining public trust in science. They help ensure that scientific evidence can play its rightful role in informing policy, supporting democratic deliberation, and contributing to a more stable and informed society.

Making code available greatly reduces the barriers to reproduction and replication, supporting this essential yet often undervalued work and helping to address the current shortage of such efforts in economics and related fields (see, e.g., Ankel-Peters et al., 2023). Additionally, sharing code allows others to build on existing work more efficiently, fostering innovation and cumulative progress. Further, replication code serves as invaluable educational tools for students and early-career researchers. It provides concrete examples of how to implement complex methods and operationalize abstract research concepts. Lastly, providing code that allows reproducing the original results benefits not only the wider research community but also the authors themselves, by allowing for better collaboration with co-authors as well as with one's future self: Replication code helps that researchers can revisit their projects in the future without confusion, supporting ongoing work as well as collaborations.

3. Alternatives to code sharing do not work well

An inferior alternative to providing replication code is to contact the authors to request the code. Several studies have shown that this approach is often ineffective, as email addresses may be outdated, the author may be deceased, and many living authors do not respond to such requests (see, e.g., Dewald et al., 1986; Hardwicke & Ioannidis, 2018; Krähmer et al., 2023). Furthermore, it is unclear why researchers do not directly provide the code, especially given that they must prepare it anyway and, with the availability of online repositories, it is technically easy to make the code publicly available.

Another alternative is to share only the data while including all details of the empirical analyses directly in the research article. While data sharing is a core element of research transparency, it is not a substitute for code sharing but a complement. For both observational studies and Randomized Controlled Trials (RCTs), there are several reasons why code provision is clearly superior to embedding all analytical details in the text. First, code provision offers greater precision than textual descriptions, much like mathematical formulas are more exact than verbal expressions. Hence, code provision can help prevent misunderstandings and ensure clarity. Second, many observational studies involve numerous variables and complex data transformations, making it difficult to report all relevant details within an article. Including everything would not only exceed space limits but also make the article harder to read (Krähmer et al., 2023). Due to these reasons, several researchers have noted the challenges of *recreate reproducibility* (Dreber and Johannesson, 2025), i.e., reproducing results from published observational studies when only the data is provided, without access to the code (e.g., Black et al., 2024). Moreover, based on my own experience teaching graduate courses, attempts to reproduce observational studies using only the original data and information provided in published articles—without access to the code—have proven to be not only extremely time-consuming but, in most cases, ultimately unsuccessful. This is largely due to the complexity of empirical analyses and the multitude of decisions researchers make throughout the research process.

Similarly, for RCTs, which are commonly used in behavioral and experimental economics, some argue or implicitly assume that code provision is unnecessary, suggesting that for simple analyses like comparing group means, data provision alone would suffice. However, most published RCTs are more complex and, for instance, include control variables, multiple hypothesis testing, or clustered treatment assignments. Further, even in simpler cases where it would be possible to include all details of the research procedure directly in the article text (e.g., just loading raw data and computing a mean difference), providing the code requires minimal effort and can only enhance transparency.

Lastly, in both observational studies and RCTs, researchers are often unable to share the full dataset due to ownership restrictions or confidentiality concerns. In many cases, only cleaned data—not raw data—can be shared, with sensitive information such as names and identifiers removed. In these instances, providing the code helps clarify the transformation from raw to cleaned data and from the cleaned data to the final analysis data, aiding in error detection in variable generation and understanding sample restrictions.[2] When data can only be shared partially—or not at all—code sharing becomes even more critical to allow others to understand and evaluate the research process.

4. Steps for more and better code provision

Despite its benefits, the level of code sharing is rather low in the social sciences. For instance, Fink and Marcus (2025)—after reviewing 2518 peer-reviewed articles using German Socio-Economic Panel (SOEP) data through journal websites, article content, author webpages, and code repositories—report that only six percent share code publicly, with overall rates remaining below 20 % in 2021 in economics and other disciplines despite growth over time. Similarly, Rainey et al. (2025) find increased code availability in political science, but most studies still lack replication code.

To make code sharing the new norm, different groups in the scientific community must take deliberate action:

- **Academic journals** should require authors to post their code in public repositories like the Harvard Dataverse, the Open Science Framework, or the Dryad Digital Repository. This is a low-cost policy, and as Fink and Marcus (2025) show, more and more journals are adopting it across disciplines. Journals do not need to host code themselves, they only need to ensure it is available, as availability remains low without enforcement (Vilhuber, 2020). While it would certainly be desirable for journals to conduct reproducibility checks themselves, this is highly resource-intensive and only a few journals can afford to do so. When journals cannot perform such checks, it remains essential that they require code, as this enables others to assess reproducibility—effectively enabling crowd-sourced verification (as, e.g., in Brodeur et al., 2024; Fišar et al., 2024). Establishing comprehensive, centralized archives—such as the "Replication Wiki" (Höffler, 2017)—to document reproducibility information for individual studies would further support this effort.

  Journals should adopt not only explicit data sharing policies but also explicit code sharing policies. Code sharing is even more important when data cannot be shared and journals should require code sharing in such cases—restricted data access is not an excuse to withhold code.
- **Peer reviewers and editors** could actively request replication code during the review process and flag its absence when it is expected but not provided.
- **Funding agencies,** similar to journals, should require that the funded researchers not only provide reports and data, but also their code.
- **Universities and research institutions** should require code underlying dissertations and policy reports to be submitted to public repositories. They can also create rewards or incentives for researchers who share their code—like the Berlin Institute of Health

[1] Huntington-Klein et al. (2025) demonstrate that decisions made during the process of preparing the final analysis dataset can substantially influence final results, underscoring the pivotal role this often-overlooked step plays in the research process.

(2025), which uses open data as a criterion in performance-based research funding.
- **Individual researchers** should prepare replication code for every project, regardless of journal requirements. This should be seen as a professional responsibility and integrated into the workflow from the outset. The code should follow good coding practices—such as clear annotations, readable structure, appropriate versioning, manageable complexity, and explicit references to specific tables, figures, or results—to ensure that it is understandable and usable by others (Gentzkow & Shapiro, 2014). When using or adapting code from others, researchers should acknowledge this by citing the original source—both to give proper credit and to ensure that code sharing efforts are recognized and valued.
- **College and university instructors** should teach students how to prepare replication code and data. Courses on research methods and data analysis should include training on documenting and sharing code, fostering these skills early in students' academic careers.
- **PhD and master's thesis supervisors** should require students to prepare replication code and data for their research. By embedding this expectation into mentorship, supervisors can instill good practices in the next generation of researchers.

While providing code is essential for transparency, it does not guarantee computational reproducibility. As Jones et al. (2022), Fišar et al. (2024), and Herbert et al. (2024) demonstrate, even when code is provided, reproducibility can still fail—for example, due to code that cannot be executed across different computing environments (e.g., missing version controls for software packages), or results that are unstable across repeated runs (e.g., when random processes are not properly controlled). Orozco et al. (2020) and Reif (2025) provide practical guidance on structuring workflows and computing environments for reproducible research. Some associations, like the European Association of Health Economics (EuHEA), already offer courses on this topic, and others could follow by training researchers to prepare reproducible code. University instructors, thesis supervisors, and other researchers can help by teaching reproducible coding practices. Importantly, writing transparent and reproducible code is a valuable skill beyond academia, promoting efficient collaboration, sustainable projects, and faster debugging.

5. Conclusion

In Angrist and Pischke (2010), Angrist and Pischke argued that research in economics has undergone a credibility revolution due to its focus on rigorous and transparent research designs for identifying causal effects. However, the "Credibility Revolution 2.0" emphasizes that a study's credibility depends not only on sound research design but also on transparency throughout the entire research process. Such openness allows peers to scrutinize each step, supports reproducibility, and fosters collective validation through replications and extensions. Providing replication code is a cornerstone of the "Credibility Revolution 2.0", and basically any empirical researcher can and should engage in it. Code sharing is a rather low-cost measure and complements other research transparency tools. Moreover, code sharing directly supports the fundamental ethos of science by making the details of the research process publicly available. Strictly speaking, empirical research can only be regarded as scientific research if it provides replication code. Otherwise, one of the criteria that sets science apart from non-science will be violated as not all details of the research process are public (King et al., 1994).

CRediT authorship contribution statement

Jan Marcus: Writing – review & editing, Writing – original draft, Conceptualization.

Data availability

No data was used for the research described in the article.

References

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives, 24*(2), 3–30.

Ankel-Peters, J., Fiala, N., & Neubauer, F. (2023). Do economists replicate? *Journal of Economic Behavior & Organization, 212*, 219–232.

Berlin Institute of Health. (2025). *Open data lom and open data iom*. Berlin Institute of Health. https://www.bihealth.org/en/notices/einfuehrung-von-open-data-als-zusaetzlicher-indikator-fuer-die-interne-leistungsorientierte-mittelvergabe-lom-forschung. Last accessed: April 29, 2025.

Black, B. S., Desai, H., Litvak, K., Yoo, W., & Yu, J. J. (2024). The SEC's short-sale experiment: Evidence on causal channels and reassessment of indirect effects. *Management Science, 70*(8), 5131–5156.

Brodeur, A., et al. (2024). Mass reproducibility & replicability: A new hope. *The Institute for Replication (I4R) Discussion Paper Series*, (107).

Claerbout J.F., & Karrenbach M. (1992). Electronic documents give reproducible research a new meaning. SEG Technical Program Expanded Abstracts, 1992(1), 601–604.

Dewald, W. G., Thursby, J. G., & Anderson, R. G. (1986). Replication in empirical economics: The Journal of Money, Credit and Banking project. *American Economic Review, 76*(4), 587–603.

Dreber, A., & Johannesson, M. (2025). A framework for evaluating reproducibility and replicability in economics. *Economic Inquiry, 63*(2), 338–356.

Fink, L., & Marcus, J. (2025). Replication code availability over time and across fields: Evidence from the German Socio-Economic Panel. *Economic Inquiry, 63*(2), 357–386.

Fišar, M., Greiner, B., Huber, C., Katok, E., Ozkes, A. I., & Management Science Reproducibility Collaboration. (2024). Reproducibility in management science. *Management Science, 70*(3), 1343–1356.

Gentzkow M., & Shapiro J.M. (2014). Code and data for the social sciences: A practitioner's guide. https://web.stanford.edu/~gentzkow/research/CodeAndData.pdf. (Last accessed April 29, 2025).

Hardwicke, T. E., & Ioannidis, J. P. A (2018). Populating the Data Ark: An attempt to retrieve, preserve, and liberate data from the most highly-cited psychology and psychiatry articles. *PLoS One, 13*(8), Article e0201856.

Herbert, S., Kingi, H., Stanchi, F., & Vilhuber, L. (2024). Reproduce to validate: A comprehensive study on the reproducibility of economics research. *Canadian Journal of Economics, 57*(3), 961–988.

Höffler, J. H. (2017). ReplicationWiki—Improving transparency in social sciences research. *D-Lib Magazine, 23*(3/4).

Huntington-Klein N., et al. (2025). The sources of researcher variation in economics. IZA Discussion Paper No. 17744.

Imai T., Toussaert S., Baillon A., Dreber A., Ertaç S., Johannesson M., Neyse L., & Villeval M.C. (2025). Pre-registration and pre-analysis plans in experimental economics. IZA Discussion Paper No. 17821.

Jones, M., Cardoso de Andrade, L., San Martin, L. E., & Singh, A. (2022). *How to make sure your research paper is reproducible: Evidence from 55 papers*. World Bank Blogs. https://blogs.worldbank.org/en/impactevaluations/how-make-sure-your-research-paper-reproducible-evidence-55-papers-guest-blog-post. Last accessed April 29, 2025.

King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research.* Princeton University Press.

Krähmer, D., Schächtele, L., & Schneck, A. (2023). Care to share? Experimental evidence on code sharing behavior in the social sciences. *PLoS One, 18*(8), Article e0288651.

Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives, 35*(3), 193–214.

Orozco, V., Bontemps, C., Maigné, E., Piguet, V., Hofstetter, A., Lacroix, A., Levert, F., & Rousselle, J. M. (2020). How to make a pie: Reproducible research for empirical economics and econometrics. *Journal of Economic Surveys, 34*(5), 1134–1169.

Rainey, C., Roe, H., Wang, Q., & Zhou, H. (2025). Data and code availability in political science publications from 1995 to 2022. *PS: Political Science & Politics*. forthcoming.

Reif J. (2025). Stata coding guide. Retrieved from https://julianreif.com/guide/. (Last accessed April 29, 2025).

Vilhuber, L. (2020). Reproducibility and replicability in economics. *Harvard Data Science Review, 2*(4), 1–39.