

Publishing Replication Packages: Insights From the Federal Reserve Bank of Kansas City

Courtney R. Butler¹

¹**Federal Reserve Bank of Kansas City, Kansas City, Missouri, United States of America**

Published on: Jul 27, 2023

DOI: <https://doi.org/10.1162/99608f92.aba61304>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Introduction

Research institutions play a vital role in generating robust research findings. Yet their role in disseminating the evidence underlying those findings—particularly for the purposes of replication—is often less clear. Many economic journals are beginning to provide guidance and infrastructure for replication packages, but this is far from standard and does not account for organizations that self-publish. Researchers seeking to share replication materials must also navigate barriers such as controls around sensitive and proprietary data. Thus, the degree to which research institutions can and should participate in the publication of replication packages remains an open question.

The Federal Reserve Bank of Kansas City has chosen to focus our resources primarily on curation in order to balance our mission-driven orientation, security considerations, and capacity limitations. The rationale and resulting procedures described below provide one possible model for involvement in the publication process, but it is based on a unique combination of factors that influence our decision-making. Each organization's landscape and priorities will help dictate the most suitable configuration.

Mission and Priorities

The Federal Reserve Bank of Kansas City (FRB KC) operates with a mission to work in the public's interest by supporting economic and financial stability. Central to this mission is public and community engagement that emphasizes trust along with core values such as integrity and service. These characteristics point toward a strong organizational alignment with the foundational principles of replication, especially transparency.

When FRB KC first began exploring reproducibility in earnest, our researchers were sharing data on an infrequent, ad hoc basis with full responsibility for the preparation and dissemination of their own replication packages. However, growing library services and evolving journal expectations drove us to reconsider our approach. An internal preservation initiative posed questions about what types of research outputs we should preserve and, importantly, why. As we considered the historical and administrative purposes of code and data, we also examined preservation policies and practices in economic journals to better understand the state of the field.¹ There, we found a steadily increasing trend in journals requiring or strongly recommending the inclusion of replication packages as part of the publication process.

With both our mission and industry trends in mind, FRB KC ultimately decided to prioritize a high degree of transparency in our approach to sharing replication packages. Our researchers' findings inform important economic policies and decision-making, and we determined that focusing on openness was one way FRB KC could tangibly demonstrate our ongoing commitment to integrity and public trust. We therefore established a formal process to enable researchers to share underlying code and data to the extent possible.

The Publication Process

Researchers initiate the publication process, and the research library coordinates the subsequent steps. These steps include legal reviews of contracts and terms of use, information security assessments, and the creation of a comprehensive ReadMe document. The ReadMe document contains vital information such as licenses, file descriptions, instructions for usage, data source references, and fixity details to ensure long-term usability.² The curation process does not presently include a verification component. However, research findings are evaluated internally via seminars, peer advisor groups, and leadership reviews prior to final publication.

Currently, our researchers are under no organizational requirements to publish replication packages; it is up to their discretion to determine when to initiate the publication process. Those decisions are often driven by a range of motivations, including journal requirements, media inquiries, collegial requests, a desire to make research widely available, and more. These packages primarily accompany research produced for FRB KC publications or external journals, but packages occasionally correspond with items such as conference presentations, reports, or indicators.³ Even when the desire is present, however, not all data can be shared in every situation. Potential barriers include code and data complexity, contractual restrictions, privacy concerns, and many more.

Despite these challenges, FRB KC confronts obstacles to sharing head-on. While many journal policies provide exceptions for situations involving proprietary or confidential data that was not originally intended to be shared, FRB KC will often seek permission from data vendors on an ad hoc basis. We also work to incorporate explicit language during contract negotiations that allows for replication-focused sharing. In this way, we have successfully shared data from a wide range of public and proprietary sources. Appropriate sharing is also sometimes facilitated through methods such as aggregation and synthetic data. Still, sharing is not always feasible, and legal and ethical matters always outweigh desires for transparency. When the data itself cannot be shared, code files, instructional appendices, or simply data references may be published to enable as much transparency as possible.⁴

Once the reviews are complete, researchers have a wide range of distribution options available. Replication packages may be posted on the FRB KC website on a researchers' profile⁵ or alongside a research product,⁶ uploaded to an external repository,⁷ emailed to a requestor, submitted to a journal, or any other means of distribution. Research library staff are available to help facilitate any desired publication method, but researchers can also manage distribution themselves. With the wide variance in publication methods, replication packages are not currently assigned individual DOIs unless they are uploaded to an external repository such as ICPSR.

The publication process has a turnaround time of approximately 3 weeks from initiation to dissemination, but it can also be expedited or may take additional time if issues arise. Defining and operationalizing this timeline has required ongoing relationship-building and negotiations with key stakeholder functions such as legal and

information security. These discussions have centered on clearly defining researchers' needs and balancing them with the organization's risk framework and available resources. This timeline is sufficient in most cases but does occasionally cause frustration and resistance in others. To mitigate this, we continuously seek out opportunities to streamline the process and proactively manage expectations.

Resourcing Considerations

The curation process is completed entirely by internal staff since replication packages are considered public once they leave our internal systems. Whether posted publicly in a repository or emailed to a single well-known colleague, we recognize that we no longer have control over replication files once they have been distributed. As a result, our curation process ensures that all replication packages meet FRB KC's legal, information security, and ReadMe requirements before being distributed to external audiences across any medium.

At this time, we have elected to place less emphasis on the mechanisms for distribution so we can instead focus our limited resources on curation for the reasons outlined above. Operationalizing this process has required a keen awareness of available skill sets, capacity, and opportunities for investment within the research library. The effort required can seem deceptively light from the outside, and it has been important to keep a pulse on the true amount of labor this process entails.

For instance, only one librarian currently has full responsibility for administering the publication process, and that process is not their sole duty. In partnership with the data services librarian and the library manager, the technical services librarian must research and understand a wide range of unique components within each package, including the restrictions on the data sources being released, the code and data file types, the structure of the code and data files, the information security criteria being reviewed, the procedures for addressing flagged issues, organizational branding standards for ReadMe's and publications, relevant journal requirements, submission processes and standards for relevant external repositories, and more. This knowledge is in addition to the communication, relationship-building, and other soft skills required to liaise with researchers, coordinate across multiple organizational stakeholders, evaluate ongoing process improvement opportunities, and further enhance the culture around reproducibility and data sharing. In other words, it can be a relatively heavy lift, especially depending on the volume of requests and competing demands.

To evolve this process further, library staff would either need to deepen their technical skills or partner with additional stakeholders such as research associates to perform code reviews, verifications, and other advanced services. Training for these activities is certainly available, but the investment required must continuously be evaluated against both the need and the benefits. Additional library staff may also be required to expand current offerings, such as enhancing distribution mechanisms. As a result, scalability and capacity remain crucial considerations as we explore future opportunities to progress publication and preservation practices. Every replication package requires time-consuming context gathering, stakeholder coordination,

documentation, and tailored decision-making, and demand for this service is growing. While we view that as an extremely positive development, we must continue being thoughtful and intentional in deploying available resources.

Conclusion

The Federal Reserve Bank of Kansas City’s commitment to transparency and public trust has shaped our approach to publishing replication packages. By transitioning from an informal, ad hoc approach to a more structured process, we have been able to further promote secure, ethical, and effective data sharing. This is especially important because the long-lived nature of the Federal Reserve System makes it imperative that evolving FRB KC research practices continue to align with the overall mission and values of the organization. After more than 100 years of supporting economic and financial stability, trust in our research is intended to extend well into the future in addition to informing the audiences of today.

As other research institutions consider their role in the publication of replication packages, their overall approach should be influenced by factors such as organizational goals, risk assessments, known impediments, and available capacity. Comprehensive curation and publication services are not appropriate for all institutions, and many external partners and resources are available to supplement where needed. These resources are particularly noteworthy because replication best practices are maturing steadily in the field of economics and will continue to normalize with time. Research institutions that proactively consider their role and expectations in this space will be much better positioned to maintain ongoing trust and integrity in their findings.

Disclosure Statement

The views expressed herein are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

References

Butler, C. R., & Currier, B. D. (2017, May 23–26). *You can’t replicate what you can’t find: Data preservation policies in economic journals* [Paper presentation]. 43rd Annual Conference of the International Association for Social Science Information Services & Technology (IASSIST), Lawrence, KS.

<http://doi.org/10.17605/OSF.IO/HF3DS>

Doh, T., & Smith, A. L. (2022). *Code and data files for the paper “A new approach to integrating expectations into VAR models.”* Research Working Paper no. 18-13, Federal Reserve Bank of Kansas City.

<https://www.kansascityfed.org/research/research-working-papers/reconciling-var-based-forecasts-2018/>

Luo, Y., Nie, J., & Young, E. (2020). *Data and code files for the paper “Ambiguity, low risk-free rates, and consumption inequality.”* Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E118047V4>

Mustre-del-Rio, J. (2015, October 21). *Software code supplement (online only) to “Solving Heterogeneous Agent Models with GPUs”* [Paper presentation]. The Economic Research in High Performance Computing Environments Workshop. Federal Reserve Bank of Kansas City's Center for the Advancement of Data and Research in Economics. <https://www.kansascityfed.org/people/josemustre-del-rio>

Rappaport, J. (2019). *Stata code for OLS estimation with spatially correlated disturbances*. Federal Reserve Bank of Kansas City. <https://www.kansascityfed.org/research-staff/jordan-rappaport/>

©2023 Courtney R. Butler. This article is licensed under a [Creative Commons Attribution \(CC BY 4.0\) International license](https://creativecommons.org/licenses/by/4.0/), except where otherwise indicated with respect to particular material included in the article.

Footnotes

1. For more information, see Butler & Currier (2017). [↵](#)
2. The FRB KC ReadMe includes organizational branding and varies in style based on replication package contents. However, a publicly available ReadMe template with detailed guidance for social science replication packages is available here: https://social-science-data-editors.github.io/template_README/. [↵](#)
3. See, for example, Mustre-del-Rio (2015) and Rappaport (2019). [↵](#)
4. The Kansas City Financial Stress Index (KCFSI) Data Sources publication is one such example: <https://www.kansascityfed.org/data-and-trends/kansas-city-financial-stress-index/>. [↵](#)
5. For example, see <https://www.kansascityfed.org/research-staff/brent-bundick/>. [↵](#)
6. For example, see Doh & Smith (2022). [↵](#)
7. For example, see Luo et al. (2020). [↵](#)