

CRRESS

**Conference on Reproducibility and Replicability in Economics and the Social
Sciences**

Table of contents

1	Home	3
2	Events	4
3	Contact Us	5
I	Session 1 - Institutional support: Should journals verify reproducibility?	6
4	Replication Packages for Journals: For and Against	8
4.1	Main Thoughts	9
4.1.1	Why Should Journal Require Replication Packages?	9
4.1.2	Why Shouldn't Journals Require Replication Packages?	11
4.2	Conclusion	13
4.3	References	14
II	Session 2 - Replication and IRB	15
5	Webinar 1 Write-up (Or put title here)	17
5.1	Background	17
5.2	Main Thoughts	17
5.3	Conclusion	18
5.4	References	18
6	Webinar 1 Write-up (Or put title here)	19
6.1	Background	19
6.2	Main Thoughts	19
6.3	Conclusion	20
6.4	References	20

1 Home

The Conference on Reproducibility and Replicability in Economics and the Social Sciences is a series of virtual and in-person panels on the topics of reproducibility, replicability, and transparency in the social sciences. The purpose of scientific publishing is the dissemination of robust research findings, exposing them to the scrutiny of peers and other interested parties. Scientific articles should accurately and completely provide information on the origin and provenance of data and on the analytical and computational methods used. Yet in recent years, doubts about the adequacy of the information provided in scientific articles and their addenda have been voiced. The conferences will address the following topics: the initiation of research, the conduct of research, the preparation of research for publication, and the scrutiny after publication. Undergraduates, graduate students, and career researchers will be able to learn about best practices for transparent, reproducible, and scientifically sound research in the social sciences.

2 Events

Upcoming and past webinars can be found [here](#)

3 Contact Us

For more information, or if you are a presenter and have questions, please [contact us](#).

CRRESS is managed by co-PIs **Lars Vilhuber and Aleksandr Michuda** (Cornell University).

The organizing committee is composed of **Vilhuber, Michuda, Ian Schmutte (UGA), and Marie Connolly (UQAM)**.

Support is provided by Sara Brooks (Cornell University) as well as the staff at the Cornell University ILR School.

Part I

Session 1 - Institutional support: Should journals verify reproducibility?

Different journals have different approaches towards enforcement of their data availability policies, ranging from a thorough and complete verification including running code and checking the output, to a cursory review of the files provided to make sure they appear satisfactory, to simply receiving the data and code package and archiving it on a website or a repository. What drives the choice of approach? What are the reasons behind such choices?

In this webinar, held on September 27th, 2022, and moderated by Lars Vilhuber, we had three panelists who are experts on this topic:

1. Guido Imbens, Professor of Economics at the School of Humanities and Sciences; Senior Fellow at the Stanford Institute for Economic Policy Research; Coulter Family Faculty Fellow at Stanford University, and editor of *Econometrica*,
2. Tim Salmon, Professor of Economics at Southern Methodist University and the editor of *Economic Inquiry*, and
3. Toni Whited, Dale L. Dykema Professor of Business Administration at the Ross School of Business at the University of Michigan and editor-in-chief at the *Journal of Financial Economics*.

4 Replication Packages for Journals: For and Against

Tim Salmon

It is vital to the integrity of our field to push for greater transparency in the research that we produce. In empirical work there is a substantial opportunity to hide, even unintentionally, very subtle but important details of a project in a long series of decisions regarding how to clean and merge data sets, how to calculate variables, how to calculate summary measures and how tests are actually performed. If all it is possible to see of a paper is a set of final tables, it is often quite difficult to understand exactly how authors achieved those results and to verify that they were produced accurately. By having authors make all of those decisions transparent and available for review it allows for the possibility of others to validate the work that has been performed and it allows for future researchers to be better able to build off of that work to advance our understanding of the issues. Journals can and should facilitate that process by requiring authors to provide full details about all empirical work performed. This is not an uncontroversial viewpoint and so in what follows I will talk about some of the key arguments for and against that position.

I have been on the editorial boards of many different journals for over 10 years. That experience, and my experience trying to publish in journals for much longer, has made me frequently question the editorial process, how to improve it and how journals can maintain high standards for work which they publish. In July of 2021, I took over as Editor of *Economic Inquiry* and was then in position to begin putting in place some policies which I thought would be beneficial in this regard. One of the first policies that I began working on was a policy requiring authors to share data and code related to papers published in the journal. I, of course, borrowed liberally from other journals which had already adopted such policies as there were many good models out there to borrow from. When the policy was finalized, we had chosen to fund a repository on OPENICPSR for both journals operated by the Western Economic Association International (*Contemporary Economic Policy* being the other journal) and establish a policy that requires all papers published by EI which include data analysis to publish a replication archive on that or a suitable alternative site. I had many discussions along the way to arrive at that policy and here I will explain some of the considerations which helped me to make the final choice.

4.1 Main Thoughts

4.1.1 Why Should Journal Require Replication Packages?

We can first examine the case in favor of journals operating data archive sites like ours or in general of requiring authors to post replication packages which will allow others to reproduce their work. The main point behind this push for reproducible science is that such measures are necessary not just to maintain the credibility of individual research papers but to maintain the credibility of all academic research. There have been many examples of fraudulent work being published in academic journals over the years including many cases of researchers faking data. Two of the more famous incidences of this type of fraud were by Michael LaCour and Diedrik Stapel. In the case of LaCour, he was able to publish a paper in *Science*, supposedly the top journal across all disciplines for academic research, in 2016 which claimed to show that contact with a homosexual individual improved one's support for gay marriage proposals.¹ This was a blockbuster finding picked up by many news outlets. It was quite humiliating to many involved when it was later discovered that the data were faked. Diedrik Stapel is a repeat offender on this issue as he was able to publish many different studies in high quality journals on the basis of faked data.² There are also other types of poor quality research that are fraudulent despite using real data which show up in journals as well. Among the more notorious offenders here would be Brian Wansink, the former head of a large research center at Cornell, who was also forced to retract many articles once the methods behind those articles were revealed.³ In his case, the data existed but he engaged in methods to achieve his results which involved, to quote Cornell's Provost at the time Michael Kotlikoff, "misreporting of research data, problematic statistical techniques, failure to properly document and preserve research results, and inappropriate authorship."⁴ Many of the results from these papers had also been picked up in the popular press and so the findings of research misconduct here were quite public and embarrassing to all of the research community that allowed this work to publish. Many more examples of these problems can be found on <https://retractionwatch.com/> and indeed the fact that such a website exists is a testament to the fact that far too much problematic research somehow makes its way to the pages of scientific journals.

We clearly need to do better and requiring more transparency in empirical work at journals is a good start. Facing requirements to provide the all of the underlying data, explicit details on methods for data collection and code for conducting the regressions would undoubtedly deter most of the cases discussed above and many other besides. This is because being required to produce the data and make it visible to others would often unmask the underlying fraud quickly and easily. There would also be a clear public record one could check to determine legitimacy of the work. Knowing it will be harder to pass through, one hopes fewer would

¹Note this is in latin.

²Note this is in latin.

³<https://retractionwatch.com/2022/05/31/cornell-food-marketing-researcher-who-retired-after-misconduct-finding-is-publishing-again/>

⁴<https://statements.cornell.edu/2018/20180920-statement-provost-michael-kotlikoff.cfm>

try and when those few still try, it should be easier to uncover the problems and deal with them as necessary. Further, not only should these requirements reduce these egregious cases of fraud, which thankfully are not that wide spread, but they will force all authors to think very carefully through their empirical processes knowing that they will be publicly viewable. This increased scrutiny should hopefully improve the quality of all research published in our journals. Preventing cases of fraud while making the details of high quality research transparently available should be a substantial boost to the legitimacy of all of our work.

It is also important that journals have policies about data availability because the ability for future researchers to reproduce existing work is necessary for the advancement of science. In many cases, one research group may wish to build upon the work already published in a journal. A first step in that process is often reproducing the initial work so that the researchers can start from there and build up. Unfortunately, if these data availability policies are not in place it is often quite difficult for a set of researchers to back out exactly what others did from a published paper alone. In one case at my own journal, a paper was submitted which was attempting to do exactly this of building off of a previous paper published at the journal. The new paper's goal was to improve on the estimation process of the previous one. The problem is that the new researchers could not reproduce the original results and so their "replication" estimation generated a result not just quantitatively different from the original authors but qualitatively different in a very meaningful way as well. This makes it then difficult to evaluate whether their improvement to the original estimation approach yielded an improvement as it is unclear that they replicated the original one correctly. That is a problem for the researchers who previously published their work as it is harder for others to build on it and it is certainly frustrating for the later researchers who cannot replicate the prior work. Having replication packages accompanying published papers can resolve this problem quickly as researchers who wish to build off of the work of others can see exactly what they did to get those results without guessing and potentially failing to identify exactly what they did.

A great example of the reason that replicating the work of others is often difficult is contained in Huntington-Klein et al (2021). This study examines the problem of replication at a deeper level than what journals usually engage in. The authors of this paper asked several teams of researchers to take the same raw data as two published papers and try to provide an answer to the same research question posed in those papers. This meant that the new researchers had to take the initial data, make all of the choices empirical researchers have to make about processing that data and specify a final regression to examine the issue. The results were that the original results often did not replicate. In some cases, the replication studies found a different sign on the key effect in question while in others, the magnitude and standard error of the effect were quite different. Importantly, in all cases, the final number of data points considered differed despite all studies starting with the same raw data with the same number of observations. The discrepancy in the final results may have been due to the fact that different research teams often made very different choices along the way to the final specification. And thus, to really know how a team of researchers arrived at a set of results, one really needs to know more than just what was the nature of the regression conducted but you need to know

all the small steps along the way to get there. Without this detailed level of information, it can be impossible to really understand how two different studies arrived at different outcomes.

It is important at this point to distinguish between two very different, though related, goals of the data availability policies of journals and how data archives may be vetted by journals. The most commonly discussed check that journals may wish to perform about a replication archive is whether one can use the archive to reproduce the results in the paper. Such a certification verifies that indeed when code is run that the results of that code reproduce what is in the paper. This verification is valuable, but a certification that the authors can re-produce their own results is not really all that useful on its own. What the paper just discussed points out is that we also need the replication sets to provide all of the details regarding how the empirical analysis was performed so that future researchers can know exactly what the authors did. With this knowledge, future researchers can begin from more robust baselines regarding published work. Without this information, we run the risk of having many parallel research programs generating seemingly conflicting results with no way to clearly determine if the conflict is due to regression specifications, different choices in data processing, errors in data processing or something else along the research chain. When designing data availability policies, we need to keep both of these goals in mind and when certifying archives as being of high quality, we need to ensure that both of these goals can be achieved.

4.1.2 Why Shouldn't Journals Require Replication Packages?

While I find the arguments above convincing regarding why journals should require replication packages, when I was contemplating putting one in place for EI, I did talk to many people who were of the opinion that journals should not be putting these requirements in place. It is worth examining their arguments against these policies to determine how convincing they are.

The first concern many would suggest about these archives is that if authors are required to post their data and their code for conducting their analysis, then others would be able to copy their work. Their concern is that the authors may have spent a great deal of time figuring out how to find the data involved, merge multiple data sets and clean them so that they work together. It may have also taken a great deal of time to implement the empirical methodology for the model in the paper. Many researchers may wish to keep that work for themselves so that they may continue to exploit that in future publications and do not want to allow others to make use of their efforts. At face value, this argument may seem somewhat convincing. While I had my own response to this, I have to say that the most convincing counter-argument against this line of thinking came from Guido Imbens in our panel discussion on this topic. He pointed out that allowing empirical researchers to hide their methods like this is similar to allowing theorists to publish theorems while keeping the proofs hidden. A theorist could mount the same argument that the proof may have taken a long time to work out, perhaps requiring the development of special techniques in the process and they may wish to be the only ones exploiting their methods in future work. We do not allow theorists to avoid providing proofs

because we need to see verification that the theorems are valid. We do not simply trust them blindly. Yet empiricists who wish to hide their methods are asking journals to blindly trust them. That should not be how publishing works. Also, while yes, making your methods and data transparent may allow others to "copy" your work, the proper way to see that is that it allows others to build off of your work. Your work can now form the foundation of the work of others and have greater impact. I would argue that the possibility that it allows others to learn more from your work is in fact one of the main reasons why journals should be requiring these packages. It is not a downside.

Another common concern about journals requiring replication packages is the suggestion that these requirements place an undue burden on authors. This can be of particular concern to certain journals as putting such requirements in place could potentially decrease the number of submissions to the journal as authors decide to submit to peer journals without such requirements. Journals likely do need to weigh this concern when considering how stringent to make their data availability policy. It is worth noting that as more and more journals adopt these policies, authors will have fewer places to submit where they can avoid these requirements and so over time concerns over this issue should diminish. It is also worth considering as a journal editor whether you want to be among the last journals not enforcing these requirements. If you are, this will mean that all those people who do not want to engage in transparent research practices will submit to your journal. As an editor, do you want to be the recipient of those submissions? Perhaps not though that decision may depend on the peer journal group for a specific journal. For journals whose peers are not yet putting these policies in place, then even high quality authors might wish to avoid the burden if they have good alternatives. For journals whose peers mostly have these requirements, then being one of the few that do not poses significant risk to the journal of receiving work for which there is a reason the authors wish to avoid transparency. Different journal editors may examine this issue and come to different conclusions on the right policy for their journal at a specific point in time. For EI, we have had the policy in place for a little less than one year and based on the current data our total submissions are slightly lower than the previous few years at this point in the cycle.⁵ There are a few other possible explanations so it is not clearly attributable to this policy but the decrease is not at a problematic level even if the data policy is responsible for the entire decline.

My other view on the issue of a replication package being a burden on authors is that this is only true if authors wait until the end of the publication process to think about the reproducibility of their work. If authors have engaged in their work in a haphazard way prior to acceptance, then it can indeed be a substantial burden to go back and document all of the data manipulation that was done and script all of the regressions performed. If, however, authors begin thinking about these issues when they begin their research, there is no real burden and in fact I would argue that engaging in your research in a way from the beginning which will make the work replicable will actually save the authors time and allow them to do higher quality work. In my

⁵I note that in the discussion I think I said our submissions were more or less unchanged. I re-checked the data after and with the most up to date numbers we have had a small but noticeable drop.

own work, I admit that early in my career I did much of my data work by hand. Then when I got a referee comment suggesting a different way to conduct a regression I would have to engage in some forensic econometrics to first back out what I actually did to get the prior result. This was wasted time and not the best way to do research. Now that I have all the analysis scripted, making changes like this is much faster and I do not have to wonder exactly how I created a variable or exactly which observations may have been dropped or why. All of that is in the scripting files from the beginning. As authors begin to expect to face these requirements and learn how to take this into account from the beginning of their analysis, the burden of providing a replication package upon acceptance of the paper diminishes substantially. I expect that these practices should be becoming more common in the profession and so the concern over this element should diminish with time. We can further diminish them by making sure that replicable research is brought into Ph.D. training programs.

A final notion that some suggested to me is that there is no need for journals to require replication packages. Individuals who want to provide their data can do so on their own sites and if there are professional incentives to do so perhaps in the form of these packages being seen as signals of high quality, everyone will do this anyway. Perhaps this could be true but most do not currently publicly archive replication files absent journal requirements. Were that to start, then it could be seen as a high-quality signal when someone does it which would mean that as journal editors we should be taking it into account in our decisions whether someone provides the data archives. If we do that, it is just a backdoor way to require replication archives but with a serious downside. If we make an accept decision based on an author saying that they will post an archive, after the paper is published authors could quickly pull that archive. Essentially, this approach is not an effective way of accomplishing the goals of research transparency. In order to ensure that the data remains available, it is best that journals maintain the archives for integrity of the process so that authors cannot manipulate the archive after the paper is published.

4.2 Conclusion

I believe quite strongly in the need for transparency in research. In order to preserve and maintain the integrity of all of academic research, we need to push for ever greater transparency in how research is done. That way when there are questions about the legitimacy of a claim, those questions can be quickly and easily addressed. This level of legitimacy is a benefit to us all. The main "cost" (if one sees it that way) would be that greater transparency limits the ability of people to publish ill-founded results. It is true that greater transparency does place greater requirements on researchers to engage in more careful and rigorous work which can survive the greater scrutiny possible with the increase to transparency. I see this as a clear benefit rather than as a cost.

Of course, the path to this greater transparency norm will not be direct and not all journals will adopt the same standards at the same time. There are some journals leading in this direction,

some following and some lagging behind. There are good reasons for different journals to be in each of those stages. As journals collectively move along this path it is important that we do so in ways that are not unduly burdensome on authors. This means that while requiring replicable archives is a valuable thing, it does not make sense for different journals to impose very different and idiosyncratic requirements about file structures and things like that such that when authors prepare a replication archive they must do a great deal of work to change it from a format suitable to one journal versus another. As a journal editor I appreciate the work done by others to establish clear guidelines on these issues which other journals can adopt as well to try to harmonize these requirements where we can.

4.3 References

Huntington-Klein, N., Arenas, A., Beam, E., Bertoni, M., Bloem, J.R., Burli, P. et al. (2021) The influence of hidden researcher decisions in applied microeconomics. *Economic Inquiry*, 59: 944– 960. <https://doi.org/10.1111/ecin.12992>

Part II

Session 2 - Replication and IRB

Lorem ipsum dolor sit amet. Vel aliquid autem et deleniti deleniti qui labore vitae est soluta exercitationem est laborum culpa At numquam voluptas. Ab inventore sunt et ullam excepturi vel quod aspernatur aut consequatur voluptatum in modi perspiciatis. Eos repudiandae quod aut quia accusamus qui sunt deserunt sit officia ipsa eum soluta nesciunt qui voluptas velit ut quidem mollitia. Ea nostrum excepturi aut perspiciatis iusto qui porro adipisci.

Molestiae facere aut voluptas beatae qui repudiandae iusto eos quibusdam culpa. Et nulla dolor in minima illo ad mollitia consequuntur eos odit ratione ut voluptatem eligendi ea unde ullam. Sed illo fuga eos doloremque deleniti ea mollitia voluptatibus nam officiis sint. Eos sunt molestias vel molestias alias id recusandae incidunt aut omnis culpa?

Ab minus consectetur ab eius quia non atque magnam est neque expedita aut voluptate asperiores in error consequatur et minima libero. Est deleniti voluptatem eos ratione reiciendis qui consequatur sunt et molestiae corrupti. Sit vitae quia ut suscipit necessitatibus et numquam nobis in nemo deserunt vel eius consequatur.

5 Webinar 1 Write-up (Or put title here)

Aleksandr Michuda

Lorem ipsum dolor sit amet. Eum accusamus expedita id totam laborum ea debitis natus est molestiae accusamus et alias quis sed repellendus omnis ad blanditiis incidunt. Et corrupti omnis qui autem fugit id natus exercitationem non nobis sunt aut reiciendis dolorem ad laboriosam sint. Cum eaque doloremque aut voluptatem quam ex molestiae aliquam.

Ex deserunt exercitationem sit reprehenderit blanditiis rem molestiae eveniet facilis quam non dolorum provident hic dolorem officia ex tempora sint. Est dolorum Quis aut esse modi aut quod molestias et illum quia provident omnis ut sunt dolor. Qui impedit totam qui nihil error et voluptate sint eum velit alias est commodi esse hic repudiandae error id aperiam officia.

5.1 Background

Et officia dolorem aut deserunt odio est eligendi corporis eum nobis debitis id dolorum nihil ea maiores nostrum aut quas iusto.¹ Est quidem voluptas adipisci consectetur et animi autem et similique voluptatem ab expedita error qui repudiandae eligendi quo magnam distinction (Alesina et al. 2003; Posner 2004).

5.2 Main Thoughts

Et alias beatae eum omnis rerum aut eligendi eligendi qui dolor voluptate a debitis recusandae! Id galisum facere et mollitia enim et fugiat fugiat et dicta dolorem est atque nemo nam sunt natus id repellendus dolorum (Alesina et al. 2003).

¹Note this is in latin.

5.3 Conclusion

Et officia dolorem aut deserunt odio est eligendi corporis eum nobis debitis id dolorum nihil ea maiores nostrum aut quas iusto.² Est quidem voluptas adipisci consectetur et animi autem et similique voluptatem ab expedita error qui repudiandae eligendi quo magnam distinction (Alesina et al. 2003; Posner 2004).

5.4 References

Alesina, Alberto et al. 2003. “Fractionalization.” *Journal of Economic growth* 8(2): 155–94.

Posner, Daniel N. 2004. “Measuring Ethnic Fractionalization in Africa.” *American journal of political science* 48(4): 849–63.

²Note this is in latin.

6 Webinar 1 Write-up (Or put title here)

Aleksandr Michuda

Lorem ipsum dolor sit amet. Eum accusamus expedita id totam laborum ea debitis natus est molestiae accusamus et alias quis sed repellendus omnis ad blanditiis incidunt. Et corrupti omnis qui autem fugit id natus exercitationem non nobis sunt aut reiciendis dolorem ad laboriosam sint. Cum eaque doloremque aut voluptatem quam ex molestiae aliquam.

Ex deserunt exercitationem sit reprehenderit blanditiis rem molestiae eveniet facilis quam non dolorum provident hic dolorem officia ex tempora sint. Est dolorum Quis aut esse modi aut quod molestias et illum quia provident omnis ut sunt dolor. Qui impedit totam qui nihil error et voluptate sint eum velit alias est commodi esse hic repudiandae error id aperiam officia.

6.1 Background

Et officia dolorem aut deserunt odio est eligendi corporis eum nobis debitis id dolorum nihil ea maiores nostrum aut quas iusto.¹ Est quidem voluptas adipisci consectetur et animi autem et similique voluptatem ab expedita error qui repudiandae eligendi quo magnam distinction (Alesina et al. 2003; Posner 2004).

6.2 Main Thoughts

Et alias beatae eum omnis rerum aut eligendi eligendi qui dolor voluptate a debitis recusandae! Id galisum facere et mollitia enim et fugiat fugiat et dicta dolorem est atque nemo nam sunt natus id repellendus dolorum (Alesina et al. 2003).

¹Note this is in latin.

6.3 Conclusion

Et officia dolorem aut deserunt odio est eligendi corporis eum nobis debitis id dolorum nihil ea maiores nostrum aut quas iusto.² Est quidem voluptas adipisci consectetur et animi autem et similique voluptatem ab expedita error qui repudiandae eligendi quo magnam distinction (Alesina et al. 2003; Posner 2004).

6.4 References

Alesina, Alberto et al. 2003. “Fractionalization.” *Journal of Economic growth* 8(2): 155–94.

Posner, Daniel N. 2004. “Measuring Ethnic Fractionalization in Africa.” *American journal of political science* 48(4): 849–63.

²Note this is in latin.