

Reply to the Comments of Referee 3 on “Quis custodiet ipsos custodes?: Despite evidence to the contrary, the American Economic Review concluded that all was well with its archive”

I thank the referee for his considered remarks. That he misinterpreted my paper is wholly my fault, and I shall have to rewrite it so that such a misinterpretation is impossible.

I deny categorically that I missed the forest for the trees, and I misunderstand nothing. The referee imputes to me the belief that the AER should reproduce the articles it publishes prior to publication. Never have I written or implied such a thing, and I don't know how the referee drew this errant conclusion. I shall have to specifically disclaim this notion when I rewrite the paper. I will not bother deconstructing the referee's argument on this point, as my rewrite will preclude such an inference by the reader.

I have many times documented that failure of archives to ensure reproducible research, because authors do not submit good data and code. What is missing is an enforcement mechanism. The enforcement mechanism that I have recommended is that the journal publish failed replications. To date, no economics journal enforces its archive, and no economics journal publishes reproducible research as a matter of course.

I note that some political science journals currently operate under this pre-publication replication model. The recent advertisement by the AER for a data editor suggests that the AER is considering this model; the advertisement states that the duties of the data editor will be:

- Design, in collaboration with the AEA journal editors and the AEA Executive Committee, a comprehensive strategy for archiving and promoting the curation of data and code that guarantees to the extent possible reproducibility of research and addresses the challenges above.
- Determine the staff and computing resources necessary to implement the strategy.
- Oversee the hiring of staff and implementation of the new policy.

If the AER is actually implements such a model, it would be a horrendous mistake not only for the reasons the referee mentions, but for two additional reasons.

First, placing the onus on the journal relieves the researcher of the responsibility to produce data and code that reproduce the published results, and is an incentive to sloppiness; our culture should not encourage this.

Second, and much more importantly, only a few journals have the resources to do this. It would be wrong for the flagship journal to do *as a first resort* that which most journals cannot do at all. The AER should be showing the not-so-well-endowed journals that it is possible to publish reproducible research without a large capital investment in software and people to do the checking. In particular, simply by being willing to publish failed replications, a journal can enforce the discipline necessary to ensure that authors submit data and code that reproduce the published results. To date, no journal in the profession does this. Only if the AER attempt this and it does not work should the AER resort to doing the replications itself.

I need to add these two reasons, and the referee's reasons, to my paper. I agree with the referee wholeheartedly that the inability to publish replications is the real problem: and this can be laid at the feet of the journal editors themselves, who are not willing to "waste" article space to ensure that the articles they do publish actually are backed by data and code. Hence, my article criticizes 30 years of AER editors failing to do anything about the fact that they habitually publish articles that are not reproducible. Current archives are all just so much kabuki theatre orchestrated by the editors.