

This paper misses the forest for the trees. It concentrates its criticism on a paper commissioned by the AER in 2011 that examined the data and programs that several authors of published papers had submitted to the AER as part of the AER Data Policy. The paper selected a nonrandom set of 39 published papers, examined the data and programs that had submitted, and attempted replication for a nonrandom set of those, and only attempted to partially replicate the results even of those. It was not a thorough replication project nor was it intended to be, as the author made clear from the beginning. The reason that it was not a thorough replication project was because the AER has never attempted to replicate the articles it has published. It only requires, for published papers, that “the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication” (p.5 of Glandon). The policy explicitly does not say that papers will be published only if it has been shown that they are replicable.

McCullough’s fundamental misunderstanding is to assume that the role of journals should be to ensure that all papers they publish are replicable: “The purpose of a journal’s data/code archive is to ensure that the journal’s published results are reproducible” (p.3 of McCullough). That is not the purpose of any data archive of any leading journal in economics; the purpose is only to provide data and programs so that other economists can make a replication attempt.

Consider what would be involved if every economics journal had to reproduce the results of every accepted paper. First, it would have to hire (for pay: economists do not work for free) an army of economists who work in different areas of economics and who are familiar with different data sets and estimation methods, and pay them to do the hard work of replication of every paper. There would be back-and-forth with the authors. The monetary cost would be enormous. It is conceivable that the AEA is wealthy enough to do that, but there are few other journals or professional associations that have anything close to the resources to accomplish that.

It would also add months to the publication time of every paper, which authors would not be happy about, and it would delay the communication of research results to the economics profession as a whole.

Even more seriously, it would be a waste of social resources to replicate every paper published. Most papers, even those in the AER, have a short citation half-life and have little impact on the future economic research. Spending vast social resources on replicating papers that have no influence in the profession would be a waste.

How, then, to have a mechanism in the discipline whereby important papers are checked for replicability and less important ones are not, which would be the efficient solution? The answer is to let the profession form the demand for replication of individual articles. Individual economists should choose which papers they think are important enough to replicate, and they should attempt replication using the data and programs that have been posted by the journals. There is still a strong role for the journals, for if the data and programs supplied by the authors do not permit replication, even though they tell the journal that they are, then the journal has the responsibility to ask the authors to provide correct data and programs to the replicator and to serve as the adjudicator of disputes.

But an even larger problem in the discipline is that even when data and programs are put on a journal's website, very few researchers actually bother to download them and replicate them. The reason for that is that it is very difficult to get replication papers published, and they don't add much to an individual's c.v., at least compared to an original paper. The first problem may be a problem with the journals, but the second problem is a problem with the profession.

It would have been better if this paper had focused on the larger issues of how replication should take place in the profession than by attacking a paper and a journal for not doing something that they never intended to do.