

Comments on: *A replication recipe: list your ingredients before you start cooking* by Andrew C. Chang, Economics Discussion Papers, No 2017-74

As I am the last reviewer to submit comments I leave out some points that others have already made.

The call for the special issue asked to include four points in the paper, and I will follow them for the structure of my review.

- (i) a general discussion of principles about how one should do a replication
- (ii) an explanation of why the "candidate" paper was selected for replication
- (iii) a replication plan that applies these principles to the "candidate" article, and
- (iv) a discussion of how to interpret the results of the replication (e.g., how does one know when the replication study "replicates" the original study).

Regarding point (i):

The two main contributions I see in this paper both come in this section and are that it points out the role pre-analysis plans can and should play in replications and that it explains how different contexts of replications should lead to different approaches.

When the different contexts are described I miss an explanation of how "a verification of the original paper for the archival record" is meant. Why should one keep such a record of others' papers? For the journal editor? Why are online appendices irrelevant?

A main aim of a pre-analysis plan usually is to publicly tie the researchers' hands to prevent data mining practices like p-hacking and HARKing to add credibility to results. Just setting a flow-time and a budget does not help much in this respect. A definition of what would be considered a successful replication somewhat does but it would be more important to describe which steps will actually be taken in the analysis and the general plan here is very short in this respect. The example later shows what kind of detail can be given in a pre-analysis plan. The problem with pre-analysis plans in empirical economic research when data is available is that there are many fast and easy ways to engage in data mining that would not leave any traces and could easily have been done before the pre-analysis plan was published. The value of a pre-analysis plan then depends much on the question whether the researcher could already have had the data when the plan was published. Otherwise a pre-analysis plan becomes mainly a tool to commit the researcher to his own plan to not waste more resources than originally planned. Given the new information the researcher assembles in the process of the replication it might become advisable not to adhere to the commitment. The discussed drawback of pre-analysis plans that they discount unexpected but extraordinary findings does not seem such an issue to me because as long as a researcher can convincingly explain how an unexpected finding came about and why plans changed readers can come to their own judgment.

Regarding point (ii):

While it is specifically mentioned that pre-analysis plans should help to ground our estimates in statistical theory I don't see any theoretical reasoning for the replication. What hypothesis does the author want to test? Just "the published results can be replicated"? Why is this important to the author? Replications seem better motivated when an author has an alternative hypothesis like testing whether "the results still hold when I use alternative approach x that I regard as more appropriate because of y". The author mentions an extension but does not specify in what way he plans to extend the original study. I would regard this as crucial for the interpretation of the replication. Is a replication "successful" if the original results can be replicated but an extension shows that these results are of questionable value to answer the research question?

One google scholar citation is a questionable measurement of influence. Two working papers of the same author that cite each other automatically have a google scholar citation each, and that does not even mean anyone has ever read them. An influential paper should be one that actually influenced other researchers to build on it or to challenge their previous findings. I would like to see a short explanation of what contribution of the original study the author regards as worthy to replicate. It can be interesting to not only take into account the influence a paper had in research. In the case of the "Growth in a Time of Debt" replication the original study was not even the newest and most important one of the original researchers on the topic, but it had gotten much attention from politicians because it was easy to take a simple message from it that was used in discussions of economically important policies. When these were challenged by a graduate student who found errors and it was seen what simple methods had been used and how easy it was to come to different conclusions it shed a light on the reliability of empirical economic research as a whole. I would thus argue to especially consider papers that had an influence in the public discussion as could be measured by citations in the media and in policy advise.

Checking the ReplicationWiki is not sufficient to determine whether a study has already been replicated. A literature research is necessary. Here google scholar can be quite helpful as many papers electronically available that cite a study are listed and a replication would most usually cite an original study.

Regarding point (iii):

For the replication plan I would suggest to mention which software the author plans to use. As the discussion of the pre-analysis plan describes a plan for working hours I would expect an estimation of how much time would be required to write the code. I agree with the referee who wrote the amount of time for the original authors to respond to inquiries for their material and the prespecified number of attempts that the author would try to contact the original authors should be specified.

Step 7 should be placed after 8 and 9 as data is needed before replicating anything and point 9 describes a precondition for the actual replication.

Regarding point (iv), the interpretation, I already mentioned that I think the paper would benefit from a description of the extension. If the real reason for the replication is to make an extension that does not have a potential to question the original study's findings that is fine, but it should be explained.

Finally, I would like to comment on the first referee report that the journal's call was only mentioned once, and it quite obviously refers to the call of the Economics e-journal for the special issue on replication, which should have been read by the referee. While I agree that there is relevant literature in medicine and psychology on pre-analysis plans that could have been cited I would think it should be taken for granted to at least name one example and an explanation of how it would benefit the paper if one sees this as a substantial criticism.