Replication Protocol for Push Button Replication (PBR)

Benjamin Wood, Annette Brown, Eric Djimeu, Maria Vasquez, Semi Yoon, and Jane Burke

Created: September 16, 2016. Current version: March 1, 2017

3ie's Replication Programme is conducting a PBR project. PBR studies test the ability for another researcher to use data and code to reproduce the originally published results. PBR researchers are not tasked with evaluating the quality of the original research or testing the robustness of the original results to any type of sensitivity analysis. They are also not expected to explore any original coding decisions. When conducting PBR studies, 3ie requires researchers to follow a set protocol. The steps of the protocol are listed sequentially and outlined below. Each paper subject to replication will have a unique component in the Open Science Framework (OSF) platform.

Protocol Steps:

1) PBR researchers read the article and check if data and code are publicly available.
2) If data and code are publicly available, PBR researchers email the corresponding author and at least one additional author (if applicable) to inform them that they will be conducting a PBR of their paper, include the PBR protocol, and ask for the statistical software and specific version they used and proceed to step 4.
3) If data and code are not publicly available, PBR researchers send a request to the corresponding author and at least one additional author (if applicable).
   a. PBR researchers send an email briefly explaining the PBR project and asking for any materials that are needed to replicate the tables and figures from their publication including: code, data, any replication instruction and the statistical software and specific version used in the analysis. In the same email, they will ask the date by which original authors can provide code, data and instructions if they can do so. The email will include the PBR protocol.

b. PBR researchers follow up with original authors after 2 weeks if they do not reply to the first email. Along with following up with the authors again, PBR researchers will ensure that they sent the message to the correct email. For example, by verifying the email address through calling the authors' institution.

c. If the authors ask for more than three months to provide the data, PBR researchers will inform them that projects without data and code within three months will be classified as "no access". The PBR researchers will record the PBR as "no access" if original authors do not reply or do not provide data or code within three months.

d. PBR researchers will record the PBR as "proprietary data" if original authors are unable to provide data because of confidentiality concerns, Institutional Review Board limitations, or ownership of data issues (i.e. they do not own the data and not allowed to share it). Appropriate documentation should also be provided by the authors demonstrating data-sharing limitations (DSL). These authors should still share their code. If they do not share their code nor provide appropriate DSL documentation, the PBR researchers will classify the paper as no access.

4) PBR researchers select the key result and write a brief justification for why these are the main results highlighted in the paper. The result selected should have been presented as key by the authors, by, for example, being highlighted in the abstract, introduction, or conclusion.

5) If in possession of data and code, PBR researchers will follow the PBR checklist, review any available replication instructions and replicate the results using existing code and data sets. PBR studies should include all tables and any programme-generated figures included in the published paper. Researchers do not need to reproduce tables in the appendix. An inability to replicate figures is of secondary concern and should be noted in the final report but should not generally influence replication status.

a. When available, PBR researchers should follow any replication instructions provided by the original authors. If instructions are not provided, PBR researchers are expected to follow the code provided by the original authors. This may mean PBR researchers need to install statistical software packages, rename data locations in .do files and generally adjust .do files to make them run. While PBR researchers are not expected to write code, they are expected to attempt to troubleshoot minor complications if they observe them.

b. If the PBR researcher cannot run any of the code sent by original authors or if they sent insufficient data and/or code, the PBR researcher will briefly explain the problems encountered and confirm the data and code with the original authors.

6) PBR researchers compare replication results with the original paper and report results using this typology below and the template report included in Appendix 1. This typology applies to the entire report.

    a. *Comparable replication:* identical results or very small changes (like rounding).

    b. *Minor differences:* small differences in coefficients and/or p-values. .

    c. *Major differences:* meaningful differences in reported outcomes (especially in the key results) or the code does not reproduce published results.

    d. *No access*: the original authors do not reply or decline to provide data or code.

    e. *Proprietary data*: unable to provide data but provided replication code and DSL documentation.

    f. *Incomplete:* If PBR researchers are unable to reproduce part of the publication due to missing code and/or data, they will report to the original authors on the PBR study's status and ask if the authors can provide more data or code to complete the PBR. If the PBR researchers cannot reproduce any tables after communicating with original authors, the paper will have two classifications. The paper will be classified as "incomplete" and in addition the paper will be classified based on the PBR results that were possible to run.

7) Guidelines for results comparison between PBR and the original publication:

    a. In case of p-values, we consider major differences to be changes equal to or greater than 0.1. We consider minor differences to be changes of less than 0.1 but greater than or equal to 0.05. Changes that are less than 0.05 are not considered as differences in this PBR context.

    b. When the authors only report statistical significance levels (for example, ***(<0.01), **(<0.05), *(<0.1)) instead of p-values, we will assume that the p-value is the middle of the range of the significance level. In the example above, we will assume a 0.075 p-value for a significance level of 10%, 0.03 for a significance level of 5%, and 0.005 for a significance level of 1%. PBR researchers will follow the statistical significance guidelines used in the original publication. The comparisons will be based in these assumptions. If the authors did not reported p-values nor significance levels, the assumptions cannot be made and the PBR researcher will report that the information was not included in the table.

c. For parameters, the definition of what is a major/minor difference will be based on PBR researcher's judgment. If available, the researcher should use summary statistics, like mean values, as references when assessing the level of difference between the original publication and the PBR results.

3ie Protocol Chart

Step	3ie Input	Product
1. Read article and identify all tables and programme-generated figures in the publication and check if the data and code are publicly available.	Selected article	
2. If data and code are publicly available, inform the original authors of the PBR via email.	Template for email 1	
3. Send email to authors requesting datasets, .do files and replication instructions.	Template for email 2	Replication documents (datasets, .do files, replication instructions)
* If authors do not respond within two weeks: send a second request and ensure that the messages are being sent to the correct email address.	Template for email 3	
* If the authors ask for more than three months to provide the data, inform them that if they do not provide the files within three months the project will be classified as *no access.*	Template for email 4	
* If authors do not respond within three months or do not provide the datasets or the .do file: report as *no access.*	Final report template (word and excel)	Final report
* If authors are unable to provide data because of confidentiality/ownership issues: report as *proprietary data* if authors provide codes and DSL documentation, otherwise report as *no access.*	Final report template (word and excel)	Final report
4. Select the key result and register a brief justification in the OSF platform.	Key result sample	
5. Review PBR instructions and replicate results using PBR checklist. PBR includes all tables and programme-generated figures from the paper.	PBR checklist	Results of replication: log file
* If PBR researchers are unable to reproduce part of the publication due to missing code and/or data, they will report to the original authors on the PBR study's status and ask if the authors can provide more data or code to complete the PBR.	Template for email 5	
6. Compare replication results with paper and report differences. The comparison can result in: • *Comparable replication:* identical results or minor (decimal level) differences. • *Minor differences:* similar coefficients & statistical significance. • *Major differences*: major differences on reported outcomes or code does not reproduce published results. • *Incomplete*: missing code and/or data	PBR checklist Tips document Final report template (word and excel)	Final report
7. Send the final report to 3ie. After internal review, 3ie will share it with the original authors.		

Checklist for PBR

1. Before running the code:
    -   Read the paper carefully and identify all the results presented in the text, tables and figures of the published paper.
    -   Select the paper's key result and register a brief justification in the OSF platform in the "Final report and key result" subcomponent.
    -   Make a copy of the files before working with them (following the naming convention 3ie provides). All the original files received from the authors should be uploaded in the private "Original data" OSF subcomponent.
    -   Verify that you received all documents that the authors said they were going to send. If you have not received all the documents, contact the authors.
    -   Read replication instructions if the authors sent them.
    -   Check that the files received correspond to the published paper. It is possible that several versions of the paper exist (if it is not possible, check this at the third stage).
    -   Review files:
        -   Review codebooks. It is important that you recognize the name of the main variables.
        -   Confirm that you have access to the data files.
        -   Read the .do files and understand what they did and which tables/results they produced.
        -   Check output documents (logs, excel tables) if they are provided by the authors.

2. Running the code:
    -   Change the file paths on the .do files if necessary.
    -   Ensure that all programs needed to run the code are installed on your computer.
    -   Use the same version of software that was used by the original authors.
    -   Run the code, logging the results (some codes will do this automatically).
    -   The documents that contain the output of the statistical analysis (e.g., a log file) will be uploaded in the "Analysis" OSF subcomponent
    -   Troubleshoot if the code does not run. For example: ensure all necessary packages are installed in the statistical program, or try running the code in an older version of the statistical software.
    -   If it is impossible to run most of the code, contact the original authors, explain the situation, and see if they have a suggested course of action (e.g. alternative data).
    -   If the results obtained are incomplete, contact the authors, explain the situation and ask them if they can send the complete code and data.

3. Comparison:
    -   Reproduce original tables using log/output files (to do this carefully read the description of each table). It is not important to format the PBR tables in the same manner, simply reproduce them following the final report template. If the original authors reported statistical significance levels instead of p-values, report the p-values in the replicated tables (this is the only case in which you should include information that was not reported in the original table).

- Compare the original tables with the tables replicated. Use the color-coding provided in the Final report template (excel file) to represent the differences between the replicated tables and the original publication.
- In case there are differences, check again the log/outputs files.
- If it is necessary, write the differences found in each table following the template provided.
- Select the classification for each table.
- Select the classification for the replication and write the justification.
- Complete the report and share the final report (along with appendices) with the 3ie, by uploading the report in the private "Final report and key result" subcomponent.

Final report Template (1 pager plus possible appendices)

Final Report – Journal Name – First original author full name (for example, Journal of Development Economics: Jane Smith)

Section 1: Basic information (one paragraph/header)

- Original paper citation
- Original authors and email addresses that were used to contact them
- PBR researcher
- List of materials received
- Classification
- Statistical software and version used for replication

Section 2: Replication process (one paragraph)
The PBR researcher briefly explains her/his steps trying to reproduce the original findings, including details of any necessary adjustments to the original code (e.g. renaming directories).

Section 3: PBR classification justification (one paragraph)
If the replication was classified as *comparable replication*, *minor differences*, *major differences* or *incomplete*, the researcher provides a detailed justification for the classification. If the replication fails to reproduce any of the published results, the researcher should make that clear.

Appendix: Depending on the results, the PBR report includes supporting documentation.

- Appendix 1:
  - List of tables from the paper are not subject to replication because they are not data-driven.
  - Brief description for each replication table that highlights where tables differ from the originals.
    - Classification (italicized) for each table and figure.
    - Differences in tables (other than parameters and p-values differences, including N, R-squared etc.)
    - Differences in parameters.
    - Differences in p-values.
  - Replication tables. Format: comparable estimates (no color), estimates that have minor differences (light green), estimates that have major differences (dark green). If there was no data or code to replicate a particular part of the table, the section will

be colored in orange and a diagonal line will be added. When the original table does not present p-values or significance levels, the p-values obtained in the replication will be reported and those cells will be colored in gray. The exact colors will be provided in the final report excel template.

- Appendix 2: List of exact names of all the files received by original authors.

Final reports will vary depending on the PBR classification. Here are the reporting requirements:

- *Comparable replication*: sections 1, 2 and 3, Appendix 1, Appendix 2, original data, original code, replication instructions, replication code and log file with replication results.
- *Minor differences*: sections 1, 2 and 3, Appendix 1, Appendix 2, original data, original code, replication instructions, replication code and log file with replication results.
- *Major differences:* sections 1, 2 and 3, Appendix 1, Appendix 2, original data, original code replication instructions, replication code and log file with replication results.
- *No access*: section 1 and correspondence with original authors as an appendix (if applicable).
- *Proprietary data:* section 1, replication code, DSL documentation and correspondence with original authors as an appendix.
- *Incomplete:* If the replication is also classified as incomplete (in this case there are two classifications), the correspondence with the original authors should also be included as an appendix of the final report.