

Guide to replication files for “Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia”

Nur Cahyadi, Rema Hanna, Benjamin A. Olken, Rizal Adi Prima, Elan Satriawan, Ekki Syamsulhakim

I. DATA

The folders “Wave I,” “Wave III,” and “Wave IV” in “data/raw” contain the original survey data. “Wave I” contains data from the baseline survey, while “Wave III” and “Wave IV” contain data from the 2-year (2009) and 6-year (2013) follow-ups, respectively. The folder “Wave I” also contains the original sub-district-level treatment assignments in **pkhben2013_use.dta**.

Note that the raw survey datasets we have published have been modified slightly from the original raw data in order to remove PII. In particular, we have removed the recorded date of birth for all survey respondents (year and month of birth are still included). As a result, because many of our outcomes rely on age in days, we have included coded age variables in the “raw” survey data. We have also included a few variants of the raw file R_AR_01.dta in each survey wave.

All intermediate datasets used for tables are stored in the folder **data/coded**. Note that we are unable to provide the unaltered raw survey data for the intermediate dataset **mothers_fertility_timing.dta**. Therefore, this intermediate dataset is already included in the zipped replication files attached here.

II. CODE

Please update the directory path found at the top of the .do file **code/MASTER.do** to reflect the location of the directory containing these files on your computer. This do file can be run to generate all intermediate datasets and tables used in the paper. The only figures included in the paper that are not generated by code in **MASTER.do** are Appendix Figures 1 and 3, which are entirely text-based and do not rely on data.

F-statistic in Table 1, Column 2

Note that for some users, slight numerical discrepancies (on the order of 0.1) from the reported F-statistic in Table 1, Column 2 can be expected.

WHO “igrowup” anthropometrics module

The World Health Organization’s “igrowup” Stata module is included in these replication files in order to calculate child anthropometric z-scores. The module has been only lightly modified in order to allow for file paths friendly to Mac and Linux operating systems. This module is called in **code/coding/code_master_child0to36mo.do**. Note that in lines 1163 and 1166 of this do file, a string variable is generated for storing the file path containing the location of the “igrowup” module and its output. The string length is set at 90 characters in the .do file. If the file path of this directory on your computer is longer than 90 characters, either change the variable type to a string longer than 90 characters, or modify the name of the directory on your local computer.

Calculation of Romano-Wolf p -values

Note that a key element of Tables 2-7 are Romano-Wolf (2005, 2016) p -values to adjust for multiple hypothesis testing using family-wise error adjustments. The code that calculates these p -values is run before the code that generates the tables themselves and takes an extremely long time to run (over 24 hours on some computers). Therefore, you have the option to toggle off the Romano-Wolf p -value code using the local “do_rwolf” at the top of **MASTER.do**. The Excel files containing the adjusted p -values for each outcome are located in the **output** folder.

Note that the Romano-Wolf procedure relies on two .ado packages customized for this project, **rwolf_pkh.ado** and **rwolf_pkh_hh.ado**, which are located in the **ado** folder.

III. OUTPUT

All output from do files called by **MASTER.do** are stored in the **output** folder. Tables and most Appendix figures are stored as .tex files and graphs are stored as PNG and EPS files.