Maternal Depression Data and Code

This readmefile explains how to replicate the analysis in Baranov, Bhalotra, Biroli, Maselko (2019) "Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Control Trial"

Authors: Victoria Baranov & Pietro Biroli

Date: October 6, 2019

Data and Code can be downloaded here: https://github.com/pietrobiroli/maternalDepression/

How to replicate the results

STEP 1:

- Download and unpack the zip file into a preferred location.
- Folder should contain the following elements

    - folders:

        - dataClean/
        - figures/
        - figures/coefficientPlot/
        - figures/preExisting
        - logfiles/
        - tables/

    - execution files:

        - README.md
        - README.pdf
        - 00_runall.do
        - THP_analysis.do
        - THP_cleandata.do
        - THP_mergedata.do
        - THP_label_variables.do
        - _gweightave.ado
        - pstar.ado
        - randcmd.ado
        - randcmd.sthlp
        - stepdownB.ado

- stepdownrandcmd.ado
- figures/coefficientPlot/coefplot_all.do

- data files:

  - dataClean/THP_clean.csv
  - dataClean/THP_clean.dta

DATA ACCESS:

The replication files contain a publically available, cleaned, and anonymized dataset, THP_clean.dta, which can be used to run the replication code in THP_analysis.do. Files THP_mergedata.do and THP_cleandata.do merge and clean raw data to produce THP_clean.dta.

The raw data for this project are confidential and not publically available, but may be made accessible with Data Use Agreements with the Human Development Research Foundation http://hdrfoundation.org/. Researchers interested in access to the data may contact Victoria Baranov at victoria.baranov@unimelb.edu.au. It can take some months to negotiate data use agreements and gain access to the data. Completion of human subject research training may also be required (https://citiprogram.org). The author will assist with any reasonable replication attempts for two years following publication.

STEP 2:

- Open 00_runall.do in Stata
- Change the "global maindir" location to the path on your computer where you downloaded the data
- run 00_runall.do: this will run the data analysis, probably over serval days, and produce all of the output presented in the paper.
- To obtain only a subset of the output, open THP_analysis.do and set the switches to 1 or 0 accordingly
- To make the code run faster, open THP_analysis.do and change "global iterations" to a lower number (e.g. 10). This change will impact the p-values calculated with randomization inference and/or the stepdown procedure reported in the tables and in Figure 2. It might also create errors in estimating the quantile treatment effects.

Description of the code

NOT FOR REPLICATION:

1. THP_mergedata.do – This file merges the raw data and ensures that the publically available data contains no confidential information. The raw datasets are not included as they all contain identifying information on respondents and health workers.

2. THP_cleandata.do – This file starts from the merged data and cleans it for the analysis

- This file will use the following inputs

  - /dataRaw/THP_merge.dta

- This file will produce the following outputs

  - /dataClean/THP_clean.dta
  - /dataClean/THP_clean.csv

- commands needed:

  - _gweightave (from https://github.com/PrincetonBPL/ado-gallery)
  - mat2txt
  - xtgraph
  - zanthro (from https://www.stata-journal.com/article.html?article=dm0004_1)

FOR REPLICATION:

1. THP_analysis.do – This file takes the clean data (from THP_cleandata.do) and runs the analysis for the paper. NB: The code can take several days to run to reproduce all the tables in the paper because of the randomization inference and stepdown procedures. p-values in Tables 2-4-6-9 and confidence interval in Figure 2 do not fully replicate beacuse of inherent randomness in both randcmd and the stepdown procedure.

   - This file will use the following inputs

     - /dataClean/THP_clean.dta

   - This file will produce the following outputs

     - all Tables and Figures in the manuscript and online appendix.

   - commands needed:

     - pstar (from https://github.com/PrincetonBPL/ado-gallery)
     - leebounds (from https://github.com/PrincetonBPL/ado-gallery)
     - randcmd (from A Young's website http://personal.lse.ac.uk/YoungA/)
     - stepdownB (adapted from https://github.com/PrincetonBPL/ado-gallery)
     - stepdownrandcmd (adapted from https://github.com/PrincetonBPL/ado-gallery)
     - mat2txt
     - estout
     - moremata
     - xtgraph
     - kdens
     - coefplot

- blindschemes
- grc1leg (from http://www.stata.com/users/vwiggins/grc1leg/grc1leg.ado)
- leebounds (from https://www.stata-journal.com/article.html?article=st0364)
- figures/coefficientPlot/coefplot_all.do

- Latex Packages needed to compile the output:

  - siunitx

    ```
      % siunitx
    \usepackage{siunitx}
    \sisetup{
      detect-mode,
      group-digits           = false,
      input-symbols          = ( ) [ ] - +,
      table-align-text-post  = false,
      input-signs            = ,
    }
    ```

  - Add the following to the preamble

    ```
    % Allow line breaks with \\ in specialcells
    \newcommand{\specialcell}[2][c]{%
    \begin{tabular}[#1]{@{}c@{}}#2\end{tabular}
    }
    ```

Correspondence between code output and paper tables and figures

Here below a crosswalk between the tables and figures in the final version of the paper and the section of the code that create them.

	Code section	Output name
Figure 2	`itt_figure'	figures/coefplot_all.pdf
Figure 3	`dep_trends'	figures/dep_trends.pdf
Table 1	`balance_tables'	tables/baseline_balance.tex
Table 2	`depression_trajectory'	tables/depression_mainvars.tex
Table 3	`depression_trajectory'	tables/depression_mainvars.tex
Table 4	`main_tables'	tables/c_main_motherdecisions.tex
Table 5	`dep_nondep'	tables/c_dep_nondep_mothergap.tex
Table 6	`main_tables'	tables/c_main_childoutcomes.tex
Table 7	`dep_nondep'	tables/c_dep_nondep_childoutcomes.tex
Table 8	`main_attrition_ipw'	tables/c_ipw_main_allindices.tex
Table 9	`main_tables'	tables/c_main_mediators.tex
Appendix Table A1	`balance_tables'	tables/attrition_balance.tex
Appendix Table A2	`balance_tables'	tables/baseline_balance _bygender.tex

Note 1: Table 2 and 3 are created from the same file, and then manually separated Note 2: p-values in Tables 2-4-6-9 and confidence interval in Figure 2 do not fully replicate beacuse of inherent randomness in both randcmd and the stepdown procedure. (Setting the seed is not enough to ensure deterministic results) Note 3: In order to compile these tables in Latex the `\specialcell{}` command is used in the headings, and we also use the packaage siunitx as a special column delimiter to indicate alignment at the decimal.

Correspondence for the Online Appendix Tables/Figures Referenced in Paper

	Code section	**Output name**
Online App Table B2-B5	`sumtab_by_index'	tables/c_sumstats*.tex
Online App Table B6-B7	`correlation_tables'	tables/correlates_index*.tex
Online App Table D11	`sr_diffs_bysmpl'	tables/c_treatmenteffect_diffs_byattrition.tex
Online App Table D12	`attrition_bygender'	tables/c_attrition_bygender.tex
Online App Table E13-E14	`sensitivity_controls'	tables/c_control_sensitivity.tex
Online App Table E15	`main_tables'	tables/c_main_motherdecisions_7y.tex
Online App Table E16	`misc'	tables/c_parenting_childdev.tex
Online App Table E17	`misc'	tables/c_dep_empowered.tex
Online App Table E18	`dd_tables'	tables/c_dd_allindices_lhwFE.tex
Online App Table E19	`magnitude'	tables/c_magnitude.tex
Online App Table G22-G25	`het_tables'	tables/c_*_het1.tex
Online App Table H26-H47	`within_index_tables'	tables/c_within_*.tex
Online App Figure I3	`density_graphs'	figures/density_*.pdf
Online App Figure I4	`qte_graphs'	figures/qte_*.pdf