README for the Reproducibility Package for
Do Psychosocial Stimulation, Parental Distress and Early Childhood Education Enrollment Show Different Associations with Early Childhood Development Outcomes for Boys and Girls? Findings from a Phone Survey in Pakistan

Overview

This paper analyzes data from a nationally representative phone survey in Pakistan conducted from December 2021 to February 2022 with caregivers with at least one child aged 72 months or younger. The code of this reproducibility package uses Stata to clean the data from raw files, run multiple imputation, and conduct all analysis as presented in the tables and figures in the paper. "Master do file" is the .do file to be run first. The replicator should expect all code to run in about 15 minutes.

Data Availability and Provenance Statements

The data used to support the findings of the present study have been deposited in the World Bank Reproducible Research Repository. The data were collected by the World Bank and will be made available via the microdatacatalog under a licensed terms of use (https://microdata.worldbank.org/index.php/terms-of-use).

Statement about Rights

- **X** I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- **X** I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are available at: https://microdata.worldbank.org/index.php/terms-of-use.

Summary of Availability

- ☐ All data **are** publicly available.
- X Some data **cannot be made** publicly available.
- ☐ **No data can be made** publicly available.

Details on Data Source

All analyses were conducted using the following data source:

Data.Name	Data.Files	Location	Provided	Citation
"Pakistan Early Childhood Development Phone Survey 2022"	ECE Survey Data Instalment 6_original data.dta	rawdata/	TRUE	World Bank (2022)

Computational requirements

The code was last run with Stata 17 on a Macbook. All files are in Stata format.

Controlled randomness

- **X** Random seed is set at lines 29 and 48 of program '2_imputation.do'.

Description of files

- The 'Master do file' runs all programs.
- 'dofiles' folder: contains all the Stata programs needed to clean and impute the raw data and produce the tables and figures in the manuscript. '1_clean.do' cleans the raw data. '2_imputation.do' creates the imputed datasets. '3_tables and figures.do' runs all analyses presented in the tables and figures in the main paper and appendix.
- 'rawdata' folder: contains the raw data file referenced above.
- 'cleandata' folder: where cleaned and imputed datasets will be saved after running the corresponding '1_clean.do' and '2_imputation.do' programs.
- 'output' folder: where all output and figures will be saved after running the '3_tables and figures.do' program. Each file will be labeled with its corresponding figure or table number from the paper.

List of Exhibits

Overview of output and connection to figures and tables in draft. We link the tables and figures to the lines in the "3_tables and figures" do file and output from the replication package here.

- **Table 1:**
  - Lines 19-165 calculate the descriptive statistics and bivariate associations presented for child development outcomes, risk and protective factors, and socio-demographics.
  - Corresponding output is captured in *$dataWorkFolder/output/Table1.log*.
- **Table 2:**
  - Lines 172-185 run Model 1 and Model 2 for children aged 0-35 months. Lines 189-201 run Model 1 and Model 2 for children aged 36-72 months.
  - Corresponding output is captured in *$dataWorkFolder/output/Table2.log.*
- **Figure 1:**
  - Lines 213-227 produce the figure and marginal effect estimates for children aged 0-35 months.
  - Lines 231-245 produce the figure and marginal effect estimates for children aged 36-72 months.
  - Corresponding figures are saved as *$dataWorkFolder/output/mimrgnsplot_distz_younger.gph* (for children 0-35 months), and *$dataWorkFolder/output/mimrgnsplot_distz_older.gph* (for children 36-72 months)
  - Corresponding output for calculating marginal effects is found in *$dataWorkFolder/output/Figures 1-3_Marginal effects.log.*
- **Figure 2:**
  - Lines 254-268 produce the figure and marginal effect estimates for children aged 0-35 months
  - Lines 272-286 produce the figure and marginal effect estimates for children aged 36-72 months.

    - Corresponding figures are saved as *$dataWorkFolder/output/mimrgnsplot_psz_younger.gph* (for children 0-35 months) *and $dataWorkFolder/output/mimrgnsplot_psz_older.gph* (for children 36-72 months)
    - Corresponding output for calculating marginal effects is found in *$dataWorkFolder/output/Figures 1-3_Marginal effects.log.*
- **Figure 3:**
    - Lines 295-309 produce the figure and marginal effect estimates for children aged 36-72 months.
    - Corresponding figure is saved as *$dataWorkFolder/output/mimrgnsplot_ece.gph*
    - Corresponding output for calculating marginal effects is found in *$dataWorkFolder/output/Figures 1-3_Marginal effects.log.*
- **Appendix Table I:**
    - For children aged 0-35 months, lines 326-332 run Model 1 and Model 2 examining associations with parental distress; and lines 336-342 run Model 1 and Model 2 examining associations with psychosocial stimulation.
    - For children aged 36-72 months, lines 346-356 run Model 1 and Model 2 examining associations with parental distress; lines 360-366 run Model 1 and Model 2 examining associations with psychosocial stimulation; and lines 370-378 run Model 1 and Model 2 examining associations with early childhood education enrollment.
    - Corresponding output is found in *$dataWorkFolder/output/Appendix Table I.log*
- **Appendix Table II**:
    - Lines 385-400 run Model 1 and Model 2 for children aged 0-35 months
    - Lines 405-413 run Model 1 and Model 2 for children aged 36-72 months
    - Corresponding output is found in *$dataWorkFolder/output/Appendix Table II.log*