Readme file for reproduction of Tables in the paper 'Nudging at scale: evidence from a government text messaging campaign during school shutdowns in Punjab, Pakistan'

Authors: Ayesha Tahir (atahir@worldbank.org), Koen Geven (kgeven@worldbank.org)

Overview

This is the readme file for the reproduction of tables and datapoints in the paper reporting on the SMS GIRL Impact Evaluation. All the data is available as Stata-data files, and the analysis is run in Stata-do files. There is a master-do file that sets paths and relevant definitions and runs all the code from the underlying do-files that produce the tables. Data were collected by the World Bank in collaboration with the Government of Punjab. Data were accessed in October 2025.

Please contact the authors for any questions about the reproducibility package.

The code provided reproduces:

- All numbers provided in text in the paper
- All tables in the paper except for Tables 1-3 which describe the school closure periods, intervention, and provide an overview of the data collected by survey round
- Note that there are no figures in the paper.

Data Availability

Data are publicly available (Licensed access) on the World Bank Microdata Catalogue.

Data Sources:

panel_data_R3_with012.dta **Access year**: 2025	Data-file that is used for the main analysis *This contains all the relevant data on the analytical sample from which the main tables with results are produced.*
panel_data_all.dta **Access year**: 2025	Data-file that is used for attrition analysis (Tables 4-5)*, which includes also relevant data from the baseline study conducted for this impact evaluation*

Requirements

The analysis is run in Stata 18 MP on Microsoft Windows.

Instructions for Replicators

New users should follow these steps to run the package successfully:

- Users must first have access to all data files if they are not included in the reproducibility package. They should go to the mentioned links, download the listed files, and place them in the data folder.
- Update the following files with your directory paths
    - replication_file_00_master_sms_girl.do
- Ensure all required software and dependencies are installed as listed in the Requirements section.
- Run the master do-file.

The replicator will need the following non-standard modules for the analysis:

- ietoolkit (for balance tables)
- outreg2 (for outputting the tables in publication format)

Statement about Rights

We certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

We certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

All data was collected with permissions of the Government of Punjab and the World Bank Pakistan Country Office.

Code and Data Description:

Data-files	
panel_data_R3_with012.dta	Data-file that is used for the main analysis. This only contains data on the analytical sample (the main sample used for analysis).
panel_data_all.dta	Data-file that is used for attrition analysis (Tables 4-5). This datafile includes also the baseline data on respondents.
Do-files	
replication_file_00_master_sms_girl.do	Stata do file that sets paths, provides definitions and runs the other main do-files to output the relevant tables in the Table

replication_file_01_tables_4_5_6_A1_A2_A3_A4.do	Stata do file that reproduction four tables in Section 3 of the paper (and four tables in the appendix): • Table 4: Sample Attrition by Survey Round • Table 5: Sample Attrition between Baseline and Analytical Sample • Table 6. Analytical Sample Descriptives • Appendix Table A.1 Attrition to Analytical sample (OLS) • Appendix Table A.2: Balance Table: Any Treatment vs. Control • Appendix Table A.3 - Balance Tests Treatment 1 vs. Treatment 2 vs. Control • Appendix Table A.4: Balance Table: Treatment 3 (Academic support) vs. Control
replication_file_02_tables_7_8_A5.do	Stata do-file that reproduces tables in section 4 of the paper (and one in the Appendix): • Table 7 Whether Child Re-Enrolled at Round 3 (Endline) • Table 8 Test Scores at Round 3 (Endline) • Appendix Table A.5. Effect of Any Treatment on (1) Re-enrollment, (2) Urdu Scores, (3) Math Scores at Endline, using all controls from balance table
replication_file_03_table_9.do	Stata do-file that reproduces a table in Section 4 of the paper: • Table 9: Heterogeneity analysis for re-enrollment at Round 3
replication_file_04_tables_10_11.do	Stata do-file that reproduces tables in Section 4 of the paper: • Table 10: Treatment effects among sample that consistently participated in all survey rounds

- Table 11: Treatment effects on time-use in the 7 days before survey administration, Round 1 and Round 2