

README: Replication Files for

Food Accounting Salience and Dietary Diversity: A Randomized Impact Evaluation of an Intervention to Track and Improve Young Children Food Consumption in Bihar, India

Overview

This reproducibility package contains the files required to replicate the cleaning, randomization checks, balance assessment, and analysis for the randomized controlled trial evaluating food tracking intervention related to Infant and Young Child Feeding (IYCF) practices.

The package includes:

- Raw survey data files
- Stata do-files for cleaning, construction, analysis, tables, and figures
- Intermediate and cleaned datasets generated during execution
- Output tables and figures used in the analysis

The replication package is designed to run sequentially through a single master setup do-file.

Software Requirements

Required Software

- Stata (tested on Stata 18.5)

Required Stata Packages

The following Stata packages are required:

- ietoolkit
- outreg2
- randtreat
- schemepack
- winsor2

If packages are not already installed, they can be installed using:

```
ssc install ietoolkit
```

ssc install outreg2

ssc install randtreat

ssc install schemepack

ssc install winsor2

Memory and Runtime Requirements

The full replication package can be executed on a standard modern computer.

Approximate requirements:

- RAM: 8GB or higher recommended
- Runtime: approximately 5–15 minutes depending on system specifications and graph rendering

Folder Descriptions

Folder	Description
raw/	Contains original raw survey and administrative datasets
working/	Creates and stores intermediate cleaned and constructed datasets
output/	Creates and stores generated tables, figures, and exported outputs
do/	Contains all Stata do-files

Raw Data Files

The following raw data files are required for execution.

Data File	Description
hh_baseline survey.xlsx	Baseline household survey data
hh_endline survey.xlsx	Endline household survey data
hh_midline.xlsx	Midline household survey data
aww_training.xlsx	Anganwadi Worker (AWW) training assessment data
aww_endline survey.xlsx	Endline Anganwadi Worker (AWW) survey data
listing data.xlsx	Household and beneficiary listing and roster data
who_muac_age_gender_zscores.dta	WHO anthropometric reference dataset used for z-score construction
hh_baseline survey metadata.xlsx	Baseline survey administration meta data

Instructions for Replicators

Step 1: Set Repository Path

Open the main setup do-file:

1. initial setup.do

Modify the global root directory:

```
global root "[YOUR LOCAL PATH]"
```

The remaining macros (global) are automatically constructed:

```
global raw "$root/raw"
```

```
global working "$root/working"
```

```
global output "$root/output"
```

```
global do "$root/do"
```

Step 2: Execute the Replication Package

Run:

1. initial setup.do

This master script sequentially executes all remaining do-files in the correct order.

The execution order is indicated by the numerical prefix to each do-file.

Description of Dofiles

Do-file	Main Tasks	Primary Outputs
1. initial setup.do	Initializes Stata environment, defines folder globals, creates working/output folders, removes previous outputs, and sequentially executes all remaining do-files.	Complete replication workflow execution; refreshed output folder containing all final tables and figures.
2. randomization.do	Cleans child listing and roster data, standardizes identifiers and dates, prepares treatment assignment variables, and constructs randomization datasets.	Cleaned randomization datasets and child roster files stored in working/.

Do-file	Main Tasks	Primary Outputs
3. aww_training.do	Cleans and harmonizes AWW training datasets, creates training knowledge indicators and scores, and prepares pre/post training analysis files.	aww training pre post.dta and cleaned AWW training datasets in working/.
4. hh_baseline.do	Imports and cleans baseline household survey data, harmonizes variable formats, generates baseline indicators and derived variables, and constructs analysis-ready datasets.	Cleaned baseline household datasets such as baseline_prelim_cleaned.dta in working/.
5. balance_tables.do	Produces treatment-control balance checks, generates descriptive statistics, merges household and AWW information, and constructs baseline comparability tables.	Baseline balance tables and descriptive outputs exported to output/ (e.g., Excel balance tables).
6. hh_endline.do	Imports and cleans endline household survey data, harmonizes variables with baseline, constructs endline outcomes and indicators, and prepares analysis datasets.	Cleaned endline household datasets in working/.
7. hh_midline.do	Imports and processes midline household survey data, harmonizes variables across rounds, and prepares datasets for longitudinal and monitoring analysis.	Cleaned midline datasets and intermediate merged files in working/.
8. aww_endline.do	Cleans and processes AWW endline survey data, generates implementation and compliance	Cleaned AWW endline and monitoring datasets in working/.

Do-file	Main Tasks	Primary Outputs
	indicators, and prepares monitoring datasets.	
9. output_annex.do	Generates supplementary and robustness analyses, constructs appendix tables, produces additional descriptive statistics and regression outputs.	Appendix tables, robustness outputs, and supplementary results exported to output/.
10. output_body.do	Produces the main manuscript tables and figures, estimates treatment effects, generates regression outputs and publication-ready figures.	Main body tables, regression outputs, and figures exported to output/ in .xls, .xlsx, .txt, and .png formats.

Outputs

All exported tables and figures are stored in: *output/*

Outputs include:

- Regression tables
- Balance tables
- Appendix tables
- PNG figures
- Excel exports used for manuscript preparation

Notes on Reproducibility

- Figures 1 and 2 within Appendix 7 of the paper are ***not included*** in the reproducibility package. These figures reflect upon an auxiliary component of the evaluation and the data for them was govt. administrative data that was provided to us under special access permission. Those datasets are not a part of our data repository on MDL as well.¹

¹ We collect data from AWCs that have administrative data digitally stored using the POSHAN Tracker application (the Poshan Tracker is a mobile application that was launched in 2021. It digitalizes data

- All do-files can be executed independently provided the folder globals are correctly defined.
- The recommended approach is to execute the package through 1. initial setup.do
- Several temporary datasets are generated during execution.
- Existing output files may be overwritten during reruns.

Data Availability and Permissions

Users are responsible for ensuring compliance with all applicable data-sharing agreements and ethical approvals.

Data available at – (submitted for approval to microdata library)

<https://microdata.lib.worldbank.org/index.php/datadeposit/summary/3182>

License

This replication package is intended solely for internal World Bank Group use. Code and materials included herein may be shared and reused for internal research and operational purposes with appropriate attribution. Access is restricted to World Bank Group staff and authorized personnel.

Users should verify any additional restrictions associated with the underlying datasets before use. Redistribution or use outside the World Bank Group is not permitted without prior written approval.

Contact

For questions related to the replication package, data construction, or analysis code, please contact the study authors.

collection and monitoring processes for nutrition outcomes of women and children. We were also able to collect information from physical record for an additional set of centers that did not use the application