

Reproducibility package for: M'ap Li Net Ale — Improving Early Grade Reading in Haiti

This reproducibility package regenerates every statistical table (in LaTeX format) and every data-derived figure in the manuscript from the raw survey data, following the World Bank [Reproducible Research Repository](#) guidelines.

Overview

The paper evaluates the *M'ap Li Net Ale* early grade reading program in Haiti using a school-level randomized controlled trial with three arms (Full treatment, Light treatment, Control). Student reading skills were measured with the Early Grade Reading Assessment (EGRA) at baseline, midline, and endline, together with head-teacher/school-director questionnaires at baseline and both follow-ups.

Running `main.do` executes the entire pipeline in about 5 minutes: raw data → score construction → estimation → LaTeX tables and PNG figures.

Data Availability Statement

All data is not yet publicly available but is expected to be made available through the World Bank Microdata Library in the future.

Rights and permissions

- ☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☒ I certify that the author(s) of the manuscript have permission to redistribute/publish the data contained within this replication package.

Data sources

All data were collected for the impact evaluation of the *M'ap Li Net Ale* program (Tout Timoun Ap Li — ToTAL / EPT project) by the Haitian Ministry of National Education and Vocational Training (MENFP) with World Bank support, between 2014 and 2016.

#	Dataset	File (data/raw/)	Round	Unit	N
1	Student EGRA + student questionnaire, baseline	EGRA_LB_MLNA.dta	2014-15	student	6,052
2	Head-teacher questionnaire, baseline	EGRA_LB_MLNA_Head_Teacher.dta	2014-15	school	360

#	Dataset	File (data/raw/)	Round	Unit	N
3	Student EGRA, midline	el_ves_egra_final_mars_n_126260bs.dta	Mar 2016	student	12,626
4	School-director questionnaire, midline	base_directeurs_final_ept.dta	Mar 2016	school	446
5	Student EGRA, endline	final_egra_me2016_endline.dta	May-Jun 2016	student	12,520
6	School-director questionnaire, endline	db_direkte_lekol_la_me_2016_ept_final29juillet.dta	May-Jun 2016	school	434

De-identification

The packaged datasets are de-identified: they contain no names, addresses, GPS coordinates, or other direct identifiers. Schools are identified by pseudonymous codes, which are retained because they are required for clustering, merging, and the by-department sample tables.

Computational requirements

Software

- **Stata 15 or later** (developed and verified with Stata 18 MP on macOS).
- Community-contributed package **blindschemes** (for the plotplain graph scheme used in the figure). `code/00_setup.do` installs it automatically from SSC if it is not already installed (requires internet access on first run only).
- No other dependencies. No Python/R required.

Memory and runtime

- Runs on a standard laptop (< 1 GB memory).
- Full runtime: approximately 5 minutes.

Randomness

No random processes are used; all results are exactly stable across runs (verified: output files are byte-identical across consecutive executions).

Instructions for replicators

1. Copy the entire ReproducibilityPackage folder to your machine.

2. Open Stata, change the working directory to the package folder (`cd ".../ReproducibilityPackage"`), and run:

```
do main.do
```

Alternatively, edit the `global root` line near the top of `main.do` to the absolute path of the package folder and run the file from anywhere. This is the only line a replicator should ever need to change (“one-button run”).

3. Outputs are written to:

- `output/tables/` — one `.tex` file per manuscript table (same file names as the manuscript’s `tables/` folder, ready for `\input{}`),
- `output/figures/` — `dist_scores_b1.png`,
- `output/logs/main.log` — full run log.

Code structure

Script	Purpose
<code>main.do</code>	One-button runner: sets the root path, runs all scripts in order,
<code>code/00_setup.do</code>	Folder globals, graph scheme, shared table-writing programs (<code>texrow6</code> , <code>texrow6_robust</code>)
<code>code/01_baseline_egra.do</code>	Baseline EGRA score construction; levels/zero/fluency/standardized tables; density figure; department counts table
<code>code/02_baseline_students.do</code>	Baseline student characteristics (balance and pooled means)
<code>code/03_baseline_schools.do</code>	Baseline school characteristics and management/governance tables
<code>code/04_midline_egra.do</code>	Midline EGRA score construction and results tables
<code>code/05_endline_egra.do</code>	Endline EGRA score construction and results tables; midline/endline sample counts table
<code>code/06_implementation.do</code>	Implementation status tables (director questionnaires merged onto student files)

Score construction replicates the original analysis do-files (`DescriptivesBaseLine.do`, `DescriptivesMidLine.do`, `Endline_Data_Analysis.do`, `BaselineDataAnalysis102015.do`); the LaTeX tables are generated directly from the estimation results (group means and differences from OLS — or logit margins for zero-score shares — with standard errors clustered at the school/class level, starred at $p < 0.10/0.05/0.01$).

List of exhibits and where they are produced

Output file (output/tables figures/)	Manuscript exhibit	Produced by
school_means_bl.tex	Schools and students by department at baseline	01_baseline_egra.do
students_bl.tex	Student characteristics at baseline	02_baseline_students.do
EGRA_bl.tex	EGRA results at baseline (% correct)	01_baseline_egra.do
EGRA_ml.tex	EGRA results at midline (% correct)	04_midline_egra.do
EGRA_ml_zero.tex	Midline: % zero scores	04_midline_egra.do
EGRA_ml_fl.tex	Midline: fluency scores	04_midline_egra.do
EGRA_el_fl.tex	Endline: fluency scores	05_endline_egra.do
EGRA_el.tex	EGRA results at endline (% correct, 2 panels)	05_endline_egra.do
EGRA_el_zero.tex	Endline: % zero scores	05_endline_egra.do
socioeconomic_bl.tex	Socioeconomic characteristics at baseline	02_baseline_students.do
schools_students_ml_el.tex	Schools/students by department, midline & endline	05_endline_egra.do
EGRA_bl_fl.tex	Baseline: fluency scores	01_baseline_egra.do
EGRA_bl_zero.tex	Baseline: % zero scores	01_baseline_egra.do
schools_char_bl.tex	School characteristics at baseline	03_baseline_schools.do
school_char_mngmt_bl.tex	School management & governance at baseline	03_baseline_schools.do
EGRA_bl_std.tex	Baseline: standardized results	01_baseline_egra.do
EGRA_ml_std.tex	Midline: standardized results	04_midline_egra.do
EGRA_el_std.tex	Endline: standardized results	05_endline_egra.do
implementation_ml.tex	Implementation status at midline	06_implementation.do
implementation_el.tex	Implementation status at endline	06_implementation.do

Output file (output/tables figures/)	Manuscript exhibit	Produced by
dist_scores_bl.png	Density of baseline composite score by group	01_baseline_egra.do
Timeline.png, Timeline_detailed.png	Program timeline diagrams	manually created (copied from data/manual_figures/; not data-derived, excluded from verification)

License

The code in this package is licensed under a MIT license. See LICENSE for details.

Contact

Juan Baron — jbaron@worldbank.org