

README — Replication Package

Texting Parents about Early Child Development: Behavioral Changes and Unintended Social Effects

Authors: Oscar Barrera-Rodríguez (University College Dublin); Karen Macours (Paris School of Economics, INRAE); Patrick Premand (DEC Development Impact, World Bank); Renos Vakis (DEC Development Impact, World Bank)

Journal: *Journal of Development Economics* (forthcoming)

Package prepared by: Daniel Corredor Vallejo and Oscar Barrera-Rodríguez

This version: 2026

1. Overview

This package reproduces every exhibit in the paper and its online appendix: main-text Tables 1–8, online-appendix Tables A2–A38, main-text Figure 1, and appendix Figures A3–A6. (Figures A1–A2, Table A1, and Table A39 are not produced by code — see the List of Exhibits.)

All code is written in **Stata**. A single master do-file (`master.do`) sets folder paths, selects which exhibits to produce, and runs two analysis do-files. The complete run produces `.tex` table fragments (in `/tables`) and `.png` figures (in `/figures`), and writes a plain-text log to the package root.

Estimated run time: approximately 15–30 minutes on a standard laptop for the full set of exhibits.

2. Data Availability and Provenance Statement

2.1 Statement of rights and summary of availability

The data used in this paper are primary survey data collected by the authors for the study, in the context of technical support to Nicaragua's Ministry of the Family (Ministerio de la Familia, Mifan). The authors have the legal right to redistribute the **de-identified** versions of these data, which are included in this replication package. All personal and direct identifiers have been removed.

All input data needed to reproduce every exhibit are included in this package (folder `/data`). A third party therefore does **not** need to obtain any data from an external source to run the package (see Section 6).

The de-identified data and accompanying documentation will additionally be archived in the **World Bank Microdata Library (MDL)**. The MDL catalog entry and DOI are *forthcoming* and will be added here once issued.

[AUTHORS — TO COMPLETE BEFORE PUBLICATION]

- MDL catalog URL:
<https://microdata.worldbank.org/index.php/catalog/XXXXX>
- Data DOI: <https://doi.org/XXXXXXX>
- Confirm the MDL access category (e.g., Public Use File / Licensed / Direct Access) and reflect it in the "License" rows below.

2.2 License

The de-identified data are distributed for the purpose of scientific reproduction. The code in this package is released under a permissive license (e.g., MIT/BSD) at the authors' discretion.

2.3 Details for each input dataset

The package uses five input datasets, all in `/data`. Each is de-identified. Because they are the authors' primary data and are provided directly, "access" for a replicator is simply the included file.

Dataset 1 — `final_nicaSMS.dta`

- **File name:** `data/final_nicaSMS.dta`
- **Brief description:** Main analysis file. One observation per target child/household (3,013 observations; 3,576 variables). Contains the follow-up (endline) outcomes — child development (Denver, TVIP, memory, self-control), socio-emotional measures, intermediate parenting indices — together with treatment-assignment variables (`treat`, `lead_treat`, `treatgen`), stratification and control variables, and the household identifier `hogarid`. Used for all impact regressions (Tables 3–8, A7, A9–A38) and the recipient assignment in Figure A6.
- **Citation:** Barrera-Rodríguez, O., Macours, K., Premand, P., and Vakis, R. (2026). *Replication data for: Texting Parents about Early Child Development*. World Bank Microdata Library. [DOI forthcoming]
- **URL / access instructions:** Included in this package (`/data`). To be archived in the World Bank Microdata Library [URL forthcoming].
- **Access / download date:** Provided by the authors; included with this package on [DATE — to complete].
- **Download steps (point-and-click):** Not applicable — file is included. After MDL archiving: open the MDL catalog page, request/accept the access terms, and download the Stata file.
- **License / access restrictions:** De-identified; redistributable for reproduction. (*Confirm final MDL access category.*)

Dataset 2 — `baseline_nica_nonames_balance.dta`

- **File name:** `data/baseline_nica_nonames_balance.dta`

- **Brief description:** Baseline survey file used for balance and attrition exhibits (16,291 observations; 583 variables). Contains baseline child, caregiver, and household characteristics, treatment-assignment and stratification variables, and identifiers (`hogar id`, `hogar id2`). The "nonames" suffix indicates names and direct identifiers have been removed. Used for Tables 1, 2, A4, A5, A6, A7, A8, and the education/leader merges in Figures A3–A4.
- **Citation:** As Dataset 1 (same data deposit).
- **URL / access instructions:** Included in this package (`/data`); MDL archiving forthcoming.
- **Access / download date:** [DATE – to complete].
- **Download steps (point-and-click):** Not applicable — file is included.
- **License / access restrictions:** De-identified; redistributable for reproduction.

Dataset 3 — `traffic_bycontent.dta`

- **File name:** `data/traffic_bycontent.dta`
- **Brief description:** Weekly counts of quiz responses, aggregated by message content arm and quiz number (`quiznb`, `triviadate`, `totalquizanswer1/2/3`). Derived from the text-messaging platform administrative logs. Used for Figures 1, A3, A4, and A5.
- **Citation:** As Dataset 1.
- **URL / access instructions:** Included in this package (`/data`); MDL archiving forthcoming.
- **Access / download date:** [DATE – to complete].
- **Download steps (point-and-click):** Not applicable — file is included.
- **License / access restrictions:** De-identified administrative aggregates; redistributable for reproduction.

Dataset 4 — `trafico_agregado.dta`

- **File name:** `data/trafico_agregado.dta`
- **Brief description:** Message-level log of the text-messaging/quiz platform (one row per message event), including message text (`mensaje`), timestamp (`fecha`), message type (`nature`), and device id (`dn`). Derived from the platform provider's administrative records. Used to build weekly quiz-participation series for Figures A3, A4, and A6.
- **Citation:** As Dataset 1.
- **URL / access instructions:** Included in this package (`/data`); MDL archiving forthcoming.
- **Access / download date:** [DATE – to complete].
- **Download steps (point-and-click):** Not applicable — file is included.

- **License / access restrictions:** De-identified administrative log; redistributable for reproduction.

Dataset 5 — Assignment.dta

- **File name:** data/Assignment.dta
- **Brief description:** Crosswalk linking the platform device identifier (dn) to the survey household identifier (hogarid/hogarida). Used to merge platform logs to survey assignment for Figures A3 and A4.
- **Citation:** As Dataset 1.
- **URL / access instructions:** Included in this package (/data); MDL archiving forthcoming.
- **Access / download date:** [DATE – to complete].
- **Download steps (point-and-click):** Not applicable — file is included.
- **License / access restrictions:** De-identified crosswalk; redistributable for reproduction.

2.4 Provenance notes

The original study was conducted as technical support to Mifan (Nicaragua). Survey data collection was carried out by the Centro de Investigación de Estudios Rurales y Urbanos de Nicaragua (CIERUNIC). Text-message delivery and quiz logs were provided by the platform operator (Cronos Ltd.; mobile service via Telefonica/Movistar). Baseline was collected in Fall 2014 and the follow-up in July–August 2015, across 97 rural villages in four municipalities (Totogalpa, Telpaneca, Yalagüina, Palacagüina). Financial support: Government of the Netherlands (BNPP) and EUR grant ANR-17-EURE-0001.

3. Dataset List

Data file	Source	Provided	Used in
data/ final_nicaSMS.d ta	Authors' survey (endline)	Yes	Tables 3–8, A7, A9– A38; Figure A6
data/ baseline_nica_n onames_balance. dta	Authors' survey (baseline)	Yes	Tables 1, 2, A4–A8; Figures A3–A4
data/ traffic_byconten t.dta	Platform logs (weekly aggregates)	Yes	Figures 1, A3, A4, A5
data/ trafico_aggrega do.dta	Platform logs (message level)	Yes	Figures A3, A4, A6
data/	Device–household	Yes	Figures A3, A4

Data file	Source	Provided	Used in
Assignment.dta	crosswalk		

4. Computational Requirements

Software

- **Stata version 15.1 or higher** (the master enforces version compatibility via `ieboilstart, versionnumber(15.1)`).
- **User-written Stata commands** (the master attempts to install these automatically from SSC; if the verification machine has no internet, install them beforehand):
 - `ietoolkit` — provides `ieboilstart` (settings standardization).
 - `estout` — provides `esttab` and `eststo` (table export). Used throughout.
 - `rangestat` — used once, for the share-of-treated-neighbors spillover table (Table A29).
- Built-in commands used include `areg`, `factor`, `pca/rotate`, `ipolate`, `reshape`, `collapse`, `merge`. No additional installation needed for these.

Operating system

- Platform-independent (Windows, macOS, Linux). All paths in the code use forward slashes `/`.

Memory and storage

- The input datasets total roughly 90 MB. `final_nicaSMS.dta` has 3,576 variables, so allow a few hundred MB of working memory; any modern machine (≥ 4 GB RAM) is sufficient.
- Output is small (text `.tex` fragments and `.png` images).

Run time

- Approximately **15–30 minutes** for the full package on a standard laptop. (*Authors: replace with the measured time.*)

Randomness

- The analysis is deterministic (no simulation/bootstrap that requires a seed). Results are fully reproducible across runs.
-

5. Description of Programs / Code

All code is in `/do`. Run **master.do**; it calls the others.

File	Role
<code>do/master.do</code>	Entry point. Installs packages, standardizes settings, sets folder paths (the only file a replicator edits), defines on/off switches for every exhibit, and runs the two analysis do-files. Writes <code>reproducibility_run.log</code> to the package root.
<code>do/global_setup.do</code>	Defines reusable globals: model specifications (<code>\$specification1</code> , <code>\$specification2</code> , <code>\$balancesp</code> , <code>\$attrition</code>) and outcome variable lists (<code>\$ecd</code> , <code>\$intermediary</code> , <code>\$syd_old</code> , <code>\$syd_young</code> , <code>\$interlead</code> , <code>\$interother</code>). Included automatically by <code>master.do</code> .
<code>do/main_paper_tables.do</code>	Produces all tables (main-text 1–8 and appendix A2–A38). Each table is wrapped in an <code>if \$tableN { ... }</code> block, in paper order. Outputs <code>.tex</code> fragments to <code>/tables</code> .
<code>do/main_paper_figures.do</code>	Produces all figures (1, A3–A6) natively in Stata. Each figure is wrapped in an <code>if \$figureN { ... }</code> block. Outputs <code>.png</code> files to <code>/figures</code> .

Notes:

- The table and figure do-files include a small "standalone safety" block so they can also be run on their own for spot-checks; that block never overrides the switches set in `master.do`.
- Tables A2 and A3 (principal-components loadings) are produced as Stata `factor/pca` output to the results window/log; the typeset appendix tables are built from those loadings (no `.tex` fragment is written for these two).

6. Instructions to Replicators

A third party can reproduce the entire package: all input data are included, so no external downloads are required.

The only modification a replicator must make is to set one folder path:

1. **Unzip** the package. Confirm it contains these sub-folders:

```
<package root>/
├── data/      (the 5 input .dta files listed above)
├── do/        (master.do, global_setup.do, main_paper_tables.do,
```

```
main_paper_figures.do)
└─ tables/      (output; may be empty before running)
└─ figures/     (output; may be empty before running)
```

2. **Open `do/master.do`** and edit the single line in PART 1:

```
global root "C:/path/to/JDE_replication_package" // <-- EDIT THIS LINE ONLY
```

Set it to the full path of the unzipped package root, using forward slashes /. No other edits are required.

3. *(Optional, offline machines only)* Pre-install the user-written commands: `ssc install ietoolkit, ssc install estout, ssc install rangestat`.
4. **Run** `master.do` (in Stata: `do "<root>/do/master.do"`).
5. **Outputs:** table fragments appear in `/tables (.tex)`, figures in `/figures (.png)`, and a full log (`reproducibility_run.log`) is written to the package root.

To reproduce a **single** exhibit, set its switch to 1 and the others to 0 in PART 2 of `master.do`.

7. List of Exhibits

`master.do` switch in **bold**. Output files are written to `/tables` (tables) or `/figures` (figures). Table fragments combine into the typeset exhibit (panels are written as `_PA/_PB`, or `A/B` stubs).

Main text

Exhibit	Switch	Output file(s)	Input data
Table 1	table1	t_1_table_base_new.tex	baseline
Table 2	table2	t_2_table_village_base.tex	baseline
Table 3	table3	t_3A_table_ITT_final.tex, t_3B_table_educ.tex	final
Table 4	table4	t_4A_table_interm.tex, t_4B_table_educ.tex	final
Table 5	table5	t_5A_table_itt_spill.tex, t_5B_table_itt_spill_educ.tex	final
Table 6	table6	t_6A_table_distance_leader.tex, t_6B_table_trea	final

Exhibit	Switch	Output file(s)	Input data
		tment_stat.tex	
Table 7	table7	t_7A_table_spill_interm.tex, t_7B_table_spill_interm_educ.tex	final
Table 8	table8	t_8_interlead.tex	final
Figure 1	figure1	f_1_treat_total.png	traffic_bycontent

Online appendix — Tables

Exhibit	Switch	Output file(s)	Input data
Table A2	tableA2	factor loadings in Stata log (no .tex)	final
Table A3	tableA3	factor loadings in Stata log (no .tex)	final
Table A4	tableA4	t_A4_table_base_cont.tex	baseline
Table A5	tableA5	t_A5_table_base_gender.tex	baseline
Table A6	tableA6	t_A6_base_leadersvsnoleaders.tex	baseline
Table A7	tableA7	t_A7A_attrition_ok_eq1.tex, t_A7B_attrition_ok_eq2.tex, t_A7C_attrition_ok_eq3.tex	baseline, final
Table A8	tableA8	t_A8_attrition.tex	baseline
Table A9	tableA9	t_A9_general_outcomes.tex	final
Table A10	tableA10	t_A10A_table_IT_theme.tex, t_A10B_table_IT_target.tex	final
Table A11	tableA11	t_A11_syd_old_PA.tex, t_A11_syd_old_PB.tex	final

Exhibit	Switch	Output file(s)	Input data
Table A12	tableA12	t_A12_syd_young_PA.tex, t_A12_syd_young_PB.tex	final
Table A13	tableA13	t_A13_syd_old_content_recip_PA.tex, _PB.tex	final
Table A14	tableA14	t_A14_syd_young_content_recip_PA.tex, _PB.tex	final
Table A15	tableA15	t_A15_interm_by_content.tex	final
Table A16	tableA16	t_A16_general_outcomes_inter.tex	final
Table A17	tableA17	t_A17_general_outcomes_inter1.tex	final
Table A18	tableA18	t_A18_general_outcomes_inter2.tex	final
Table A19	tableA19	t_A19_leader_outcomes.tex	final
Table A20	tableA20	t_A20_syd_old_PA.tex, t_A20_syd_old_PB.tex	final
Table A21	tableA21	t_A21_syd_young_PA.tex, t_A21_syd_young_PB.tex	final
Table A22	tableA22	t_A22_sp2_leadership.tex	final
Table A23	tableA23	t_A23_distance_network.tex	final
Table A24	tableA24	t_A24_leader_ecd_bytarget.tex	final
Table A25	tableA25	t_A25_leaders_interm_bytarget.tex	final
Table A26	tableA26	t_A26_leader_social_interactions.tex	final
Table A27	tableA27	t_A27_liders_in	final

Exhibit	Switch	Output file(s)	Input data
		term_baseline.tex	
Table A28	tableA28	t_A28_PA_untreated_leaders.tex, t_A28_PB_untreated_leaders.tex	final
Table A29	tableA29	t_A29_share_treated_spillovers.tex	final
Table A30	tableA30	t_A30_leader_no_leader.tex	final
Table A31	tableA31	t_A31_leader_no_leader_interm.tex	final
Table A32	tableA32	t_A32_leader_Attitudes.tex	final
Table A33	tableA33	t_A33_interm_Attitudes.tex	final
Table A34	tableA34	t_A34_commitment.tex	final
Table A35	tableA35	t_A35_commitment_interm.tex	final
Table A36	tableA36	t_A36_commitment_predicted.tex	final
Table A37	tableA37	t_A37_commitment_predicted_interm.tex	final
Table A38	tableA38	t_A38_interother.tex	final

Online appendix — Figures

Exhibit	Switch	Output file	Input data
Figure A3	figureA3	figure_participation_educ_caregiver.png	trafico_agregado, trafico_bycontent, Assignment, baseline, final
Figure A4	figureA4	figure_participation_leader.png	trafico_agregado, trafico_bycontent, Assignment, baseline, final
Figure A5	figureA5	f_A1_treatind.png	trafico_bycontent

Exhibit	Switch	Output file	Input data
Figure A6	figureA6	f_A6_responses_by_treatgen_weekly.png	trafico_agregado, final

Exhibits NOT produced by code

These are typeset/drawn directly in the manuscript and have no associated program:

- **Figure A1** — Household-level randomization diagram (schematic).
- **Figure A2** — Village-level randomization diagram (schematic).
- **Table A1** — Example text messages, by content (illustrative text).
- **Table A39** — Minimum Detectable Effects / statistical power (values typeset in the manuscript; the underlying sample sizes are reported in the paper text).

8. Notes for the Reproducibility Reviewer

- Two figures apply a documented presentational smoothing of known data-transmission gaps in the platform logs (the late-April-2015 weeks are excluded across the participation figures; the early-August-2015 transmission gap is linearly interpolated in Figure A6). These are flagged in code comments and match the smoothing noted in the figure captions.
- Figure file names retain their working stems (e.g., f_A1_treatind.png for Figure A5); the manuscript's \includegraphics calls reference these names. The mapping is given in the List of Exhibits above.

9. References

Barrera-Rodríguez, O., Macours, K., Premand, P., & Vakis, R. (forthcoming). *Texting Parents about Early Child Development: Behavioral Changes and Unintended Social Effects*. Journal of Development Economics.

Barrera-Rodríguez, O., Macours, K., Premand, P., & Vakis, R. (2026). *Replication data and code for: Texting Parents about Early Child Development*. World Bank Microdata Library / [repository]. [DOI forthcoming]