

Replication Package README

When Aspirational Videos Meet Socio-Emotional Skills Training: Experimental Evidence from a Youth Employment Program in Côte d'Ivoire

Léa Rouanet, Jeannie Annan, Manil Zenaki, Clara Delavallade, Estelle Koussoubé, and David K. Evans

Replication contact: mzenaki@worldbank.org

Contents

- Overview
- Data Availability
- Instructions for Replicators
- Requirements
- Code Description
- Folder Structure
- List of Exhibits

1. Overview

This package reproduces the quantitative statistical tables in the paper using Stata. The analysis evaluates a randomized controlled trial of the Pro-Jeunes youth employment program among 6,023 eligible applicants in Abidjan and Grand-Bassam, Côte d'Ivoire. Eligible applicants were assigned to a control group or to a multi-component employment program whose socio-emotional modules emphasized either intrapersonal or interpersonal skills, and were further randomized to one of two sets of aspirational videos or no video

The single entry point is **2.Dofiles\0-master.do**. After the replicator changes the project directory on line 19, the master do-file creates the cleaned analysis data and generates 40 statistical tables in the main text and appendices. The code-generated .tex files are included in the manuscript without further manual modification.

Figures 1-2 and Tables 1-3 and A.1 are intentionally manuscript-authored contextual exhibits. They do not report estimates produced by the analysis pipeline and are not computational replication targets. The manuscript also uses qualitative interviews and focus groups to interpret mechanisms; those materials do not generate any figure or table in the exhibit map and are outside the computational reproduction scope of this package.

Expected run time: approximately 10-15 minutes on the benchmark machine described under Requirements.

2. Data Availability

Summary of availability and embargo

The de-identified baseline surveys can be shared now. They are forthcoming in the World Bank Microdata Library, where the deposit is expected shortly.

The remaining datasets (follow-up, treatment-assignment, and monitoring) remain under embargo until journal publication. During the embargo, the de-identified quantitative datasets in the replication package will be shared privately with World Bank reproducibility reviewers solely for verification. Upon journal publication, the embargo lifts and the authors will publish the remaining datasets in the World Bank Microdata Library. A study-specific catalog URL will be added when the deposit is released.

Future repository: World Bank Microdata Library, <https://microdata.worldbank.org/> (study-specific URL forthcoming).

Input datasets

Filename	Data name	Data owner	Source and description	Access/collection period
1.Data\Baseline\all_cohorts_raw_noPII.dta	Baseline data	World Bank	De-identified baseline surveys for the three study cohorts; collected with Innovations for Poverty Action.	Collected May 2020-January 2022
1.Data\Follow_up\all_cohorts_midline_raw_noPII.dta	Follow-up data	World Bank	De-identified follow-up surveys for the three study cohorts; collected with Innovations for Poverty Action.	Collected May 2022-June 2023
1.Data\Treatment\all_cohorts_assigned.dta	Treatment assignment data	World Bank	Treatment-assignment records.	Created September 2020-January 2022
1.Data\Monitoring\all_cohorts_monitoring.dta	Monitoring data	World Bank	Administrative program-monitoring records supplied by the International Rescue Committee (IRC).	Collected during program implementation 2020- 2023

URL and access status for every external input: The de-identified baseline surveys can be shared now and are forthcoming in the World Bank Microdata Library shortly. The follow-up, treatment-assignment, and monitoring datasets are not publicly accessible during the embargo and will be deposited in the World Bank Microdata Library upon journal publication.

Generated data files

Generated file (within 1.Data)	Written by
Baseline\all_cohorts_ses.dta	2-baseline_variable_SES.do
Baseline\all_cohorts_baseline_variables.dta	3-baseline_variable_creation.do
Follow_up\all_cohorts_midline_ses.dta	1a-followup_variables_SES.do
Follow_up\all_cohorts_midline_variables.dta	1-followup_variable_creation.do
Merged\data_merged_before_cleaning.dta	2-merge_midline_and_baseline.do
Merged\data_merged_after_cleaning.dta	4-data_cleaning.do

Statement about rights

- The authors certify that they have legitimate access to and permission to use the data used in the manuscript.
- The data in the private replication package are de-identified and may be provided to World Bank reproducibility reviewers during the embargo.
- The root CC0 1.0 LICENSE applies to the code and package documentation. It does not authorize redistribution of the embargoed datasets. Data released after the embargo will be governed by the terms stated in the World Bank Microdata Library catalog record.

3. Instructions for Replicators

1. Download the replication package, including the de-identified data under 1.Data.
2. Install Stata 19.5.
3. Open 2.Dofiles\0-master.do and change only line 19, global maindir, so that it points to the replication-package root on your machine.
4. If needed, install the user-written packages before running the master do-file by editing row 72 of the master.dofile from “local packages = 0” to “local packages = 1”
5. Run 2.Dofiles\0-master.do from a fresh Stata session. Expected run time is approximately 10-15 minutes.

Important: All code-generated tables are used in the manuscript as produced, with no post-export manual modification.

4. Requirements

Software requirements

- Stata 19.5 (the package uses dsregress and therefore requires a modern Stata release).

- User-written Stata packages named in the master do-file: egenmore, schemepack, swindex, listtab, and grc1leg2.
- Two project-local custom programs are supplied under 2.Dofiles\0.Custom_program: kling_noimputation_program.do and fdr_anderson_qvalues.do.

Benchmark hardware, storage, and runtime

Item	Benchmark
Operating environment	Windows laptop
Processor	AMD Ryzen 7 5800H with Radeon Graphics, 3.20 GHz
Memory	32.0 GB installed
Graphics	NVIDIA GeForce RTX 3070 Laptop GPU, 8 GB; GPU acceleration is not required by the Stata code
Package size	Approximately 555 MB before execution
Observed end-to-end runtime	Approximately 10-15 minutes

5. Code Description

Program or folder	Purpose
2.Dofiles\0-master.do	Single entry point. Runs all other dofiles. Defines paths, loads globals and custom programs, sets see , runs cleaning, main analysis, and robustness checks.
2.Dofiles\1-globals.do	Defines outcome components, indexes, descriptive variables, and regression-control lists.
2.Dofiles\0.Custom_program	Contains two custom programs: kling_noimputation_program.do: to generate Kling indices and compute fdr_anderson_qvalues.do: to compute the Anderson (2008) sharpened q-value procedure based on Benjamini, Krieger and Yekutieli (2006).
2.Dofiles\1.Cleaning	Cleans baseline and follow-up surveys, Create baseline and follow-up analysis variables, merges survey rounds, and produces the cleaned merged dataset.
2.Dofiles\2.Main_analysis\2.1.Descriptive_stats_and_checks	Produces attrition, balance, descriptive-statistics, training participation, and video-attendance tables.
2.Dofiles\2.Main_analysis\2.2.ITT_fdr	Produces the main ITT index tables, component tables, and treatment-arm equality tests with false-discovery-rate adjustments.
2.Dofiles\3.Robustness_checks	Produces robustness tables controlling for baseline imbalances, using post-double-selection Lasso, and retaining one randomly selected respondent per household.

Randomness: the master do-file sets seed 1947 before the cleaning and analysis sequence.

6. Folder Structure

```
|-- 1.Data/
|   |-- Baseline/
|   |-- Follow_up/
|   |-- Merged/
|   |-- Monitoring/
|   `-- Treatment/
|
|-- 2.Dofiles/
|   |-- 0.Custom_program/
|   |-- 1.Cleaning/
|   |   |-- Baseline/
|   |   |-- Follow_up/
|   |   `-- Merged/
|   |-- 2.Main_analysis/
|   |   |-- 2.1.Descriptive_stats_and_checks/
|   |   `-- 2.2.ITT_fdr/
|   |       |-- Components/
|   |       `-- Indexes/
|   `-- 3.Robustness_checks/
|       |-- ITT_control_imbalances/
|       |-- ITT_control_lasso/
|       `-- ITT_one_hh_member/
|
|-- 3.Outputs_main/
|   `-- 3.1.Tables/
|       |-- 3.1.a Descriptive_stats_and_checks/
|       `-- 3.2.b ITT/
|           |-- Components/
|           `-- Indexes/
|
`-- 4.Outputs_robustness_checks/
    `-- 4.1.Tables/
```

```
|-- ITT_control_imbalances/  
|-- ITT_control_lasso/  
`-- ITT_one_hh_member/
```

7. List of Exhibits

The code reproduces every statistical table in the manuscript. Figures 1-2 and Tables 1-3 and A.1 are contextual, manuscript-authored exhibits and have no generating program or standalone output. Output-directory keys used below:

- M-D = 3.Outputs_main\3.1.Tables\3.1.a Descriptive_stats_and_checks
- M-I = 3.Outputs_main\3.1.Tables\3.2.b ITT\Indexes
- M-C = 3.Outputs_main\3.1.Tables\3.2.b ITT\Components
- R-F = 4.Outputs_robustness_checks\4.1.Tables\ITT_control_imbalances
- R-G = 4.Outputs_robustness_checks\4.1.Tables\ITT_control_lasso
- R-H = 4.Outputs_robustness_checks\4.1.Tables\ITT_one_hh_member

Figure/ Table	Program	Output file	Note
Figure 1	N/A	N/A	Contextual manuscript-authored exhibit; not a computational replication target.
Figure 2	N/A	N/A	Contextual manuscript-authored exhibit; not a computational replication target.
Table 1	N/A	N/A	Contextual manuscript-authored exhibit; not a computational replication target.
Table 2	N/A	N/A	Contextual manuscript-authored exhibit; not a computational replication target.
Table 3	N/A	N/A	Contextual manuscript-authored exhibit; not a computational replication target.
Table 4	Attrition.do	Attrition.tex	M-D; inserted as-is.
Table 5	balance_check.do	balance_check.tex	M-D; inserted as-is.
Table 6	descriptive_stats_bygender.do	descriptive_statistics.tex	M-D; inserted as-is.
Table 7	modules_completed_bygender_new.do	takeup_modules_completed_bygender_new.tex	M-D; inserted as-is.
Table 8	takeup_video.do	takeup_video.tex	M-D; inserted as-is.
Table 9	ITT_indexes_main_ses_fdr.do	itt_indexes_main_ses_fdr.tex	M-I; inserted as-is.
Table 10	ITT_indexes_asp_job_savings_finance_fdr.do	itt_indexes_asp_job_savings_finance_fdr.tex	M-I; inserted as-is.
Table 11	ITT_indexes_labor_business_fdr.do	itt_indexes_labor_business_fdr.tex	M-I; inserted as-is.
Table 12	ITT_indexes_satisf_dm_ga_norms_fdr.do	itt_indexes_satisf_dm_ga_norms_fdr.tex	M-I; inserted as-is.

Figure/ Table	Program	Output file	Note
Table 13	ITT_indexes_srh_gbv_fdr.do	itt_indexes_srh_gbv_fdr.tex	M-I; inserted as-is.
Table A.1	N/A	N/A	Contextual manuscript-authored exhibit; not a computational replication target.
Table B.1	Attrition_gender.do	Attrition_gender.tex	M-D; inserted as-is.
Table B.2	balance_check_gender.do	balance_check_gender.tex	M-D; inserted as-is.
Table C.1	ITT_modules_completed_bygender.do	itt_takeup_modules_completed_bygender.tex	M-D; inserted as-is.
Table D.1	ITT_test_treatment_arms_equality.do	itt_test_treatment_arms_equality.tex	M-I; inserted as-is.
Table E.1	itt_components_ses_intra_sr_fdr.do	itt_components_ses_intra_sr_main_fdr.tex	M-C; inserted as-is.
Table E.2	itt_components_ses_inter_sr_fdr.do	itt_components_ses_inter_sr_main_fdr.tex	M-C; inserted as-is.
Table E.3	itt_components_ses_sjt_fdr.do	itt_components_ses_sjt_fdr.tex	M-C; inserted as-is.
Table E.4	itt_components_ga_fdr.do	itt_components_gender_attitudes_fdr.tex	M-C; inserted as-is.
Table E.5	itt_components_labor_bus_fdr.do	itt_components_labor_business_fdr.tex	M-C; inserted as-is.
Table E.6	ITT_components_norms_fdr.do	itt_components_norms_fdr.tex	M-C; inserted as-is.
Table E.7	ITT_components_jobsearch_fdr.do	itt_components_jobsearch_fdr.tex	M-C; inserted as-is.
Table E.8	ITT_components_network_fdr.do	itt_components_network_fdr.tex	M-C; inserted as-is.
Table E.9	ITT_components_decisions_particip_fdr.do	itt_components_particip_decisions_fdr.tex	M-C; inserted as-is.
Table E.10	ITT_components_perceived_autonomy_decisions_fdr.do	itt_components_perceived_autonomy_decisions_fdr.tex	M-C; inserted as-is.
Table E.11	itt_components_srh_fdr.do	itt_components_srh_fdr.tex	M-C; inserted as-is.
Table F.1	ITT_indexes_main_ses_imbalances.do	itt_indexes_main_ses_imbalances.tex	R-F; inserted as-is.
Table F.2	ITT_indexes_asp_job_savings_finance_imbalances.do	itt_indexes_asp_job_savings_finance_imbalances.tex	R-F; inserted as-is.

Figure/ Table	Program	Output file	Note
Table F.3	ITT_indexes_labor_business_imbalances.do	itt_indexes_labor_business_imbalances.tex	R-F; inserted as-is.
Table F.4	ITT_indexes_satisf_dm_ga_norms_imbalances.do	itt_indexes_satisf_dm_ga_norms_imbalances.tex	R-F; inserted as-is.
Table F.5	ITT_indexes_srh_gbv_imbalances.do	itt_indexes_srh_gbv_imbalances.tex	R-F; inserted as-is.
Table G.1	ITT_main_ses_control_lasso.do	itt_indexes_main_ses_control_lasso.tex	R-G; inserted as-is.
Table G.2	ITT_asp_job_finance_control_lasso.do	itt_indexes_asp_job_savings_finance_control_lasso.tex	R-G; inserted as-is.
Table G.3	ITT_labor_business_control_lasso.do	itt_indexes_labor_business_control_lasso.tex	R-G; inserted as-is.
Table G.4	ITT_satisf_dm_ga_norms_control_lasso.do	itt_indexes_satisf_dm_ga_norms_control_lasso.tex	R-G; inserted as-is.
Table G.5	ITT_srh_gbv_control_lasso.do	itt_indexes_srh_gbv_control_lasso.tex	R-G; inserted as-is.
Table H.1	ITT_indexes_main_ses_one_hh_member.do	itt_indexes_main_ses_one_hh_member.tex	R-H; inserted as-is.
Table H.2	ITT_indexes_asp_job_savings_finance_one_hh_member.do	itt_indexes_asp_job_savings_finance_one_hh_member.tex	R-H; inserted as-is.
Table H.3	ITT_indexes_labor_business_one_hh_member.do	itt_indexes_labor_business_one_hh_member.tex	R-H; inserted as-is.
Table H.4	ITT_indexes_satisf_dm_ga_norms_one_hh_member.do	itt_indexes_satisf_dm_ga_norms_one_hh_member.tex	R-H; inserted as-is.
Table H.5	ITT_indexes_srh_gbv_one_hh_member.do	itt_indexes_srh_gbv_one_hh_member.tex	R-H; inserted as-is.