

README

From Access to Achievement: The Primary-School-Age Impacts of an At-Scale Preschool Construction Program in Highly Deprived Communities

By Marina Bassi, Bruno Besbas, Lelys Dinarte-Diaz, Saravana Ravindran, and Ana Reynoso

May 2024

Overview

This replication package reproduces the tables in the paper "From Access to Achievement: The Primary-School-Age Impacts of an At-Scale Preschool Construction Program in Highly Deprived Communities". The numbering of the tables coincides exactly with those numbered in the paper.

The folders included are the following:

- Codes
- Data

This replication package does not produce output files (in xlsx, csv, tex or other formats), but rather displays them in the Stata console.

Requirements

Stata version 14.2 or higher required. This version of the replication package was run on Stata 17.0 MP (2-core) and Stata 18.0 MP—Parallel Edition (2-core).

Required packages are installed in the "00_Main.do" dofile (lines 25-30). Comment or uncomment according to needs.

Packages needed: wyoung, ritest, tabmiss, frmtable, st0085_1 (which includes esttab, eststo, estadd), schemepack and coefplot.

Execution time

Depending on the number of repetitions for the multiple hypothesis testing (wyoung command) and randomization inference p-values (ritest command), which were 10,000 and 2000 respectively. With these

numbers, the execution time ranges between 8 and 9 hours (for a PC with 16GB of RAM and an Intel Core i7 CPU with 6 cores and 12 logical processors).

However, you can modify the number of repetitions in the "00_Main.do" dofile, in lines 36 and 37. Enter 1 or 2 repetitions if you only need to see the coefficients and significance. In this case, the execution time is a few seconds for each table.

Folders and contents

Codes

This folder contains the code with which each table/figure was produced. The numbering of the tables coincides exactly with those numbered in the paper. The dofiles are the following:

00_Main.do

01_Table 1.do

02_Table 2.do

03_Figure 2.do

04_Figure 3.do

05_Table 3.do

06_Table 4.do

07_Table 5.do

08_Table 6.do

09_Table 7.do

10_Figure A2.do

11_Table A1.do

12_Table a3.do

13_Table a4.do

14_Table a5.do

15_Table a6.do

16_Table a7.do

17_Tables a8-a13.do

18_Tables a14-a19.do

DICIPE_VarLabels.do

Figure 1 and Figure A1.xlsx

Data

The data is anonymized and available in the "data" folder, containing:

- **data_BL.dta:** This dataset contains the baseline information of 23,943 individuals in 4687 households. It includes data on target child characteristics and measurements, caregivers, and household information at the baseline.
- **data_EL.dta:** This dataset contains follow up data of 19,373 individuals from 3765 households. It includes the outcome variables and the characteristics of target child, caregivers, and households in the follow up.
- **data_labor_wealth_EL.dta:** Contains the endline follow up data of the 19,373 individuals (the same as in 'data_EL.dta'), and includes also labor data and the control variables collected at baseline (from 'data_BL.dta')
- **dframe_indices.dta:** Includes the variables used for the Conditional Average Treatment Effect (CATE) analysis that was estimated for the three main outcomes of interest (see tables 3, 4 and 5 on the paper).
- **DICIPE Endline Escolinha.dta:** Includes data on preschools characteristics (attendance, enrollment, etc.), equipment and infrastructure, and characteristics of instructors.

Accessibility: All datasets used to produce the results presented in this manuscript are publicly available in the World Bank Microdata Catalog.

Version & Date: The baseline and endline data are unique versions produced between 2016 and 2020. They were accessed in October 2023 for the current paper.

Access Paths:

Endline public catalog link: <https://microdata.worldbank.org/index.php/catalog/4021>

Internal catalog link: <https://microdatalib.worldbank.org/index.php/catalog/12224>

Baseline public catalog link: <https://microdata.worldbank.org/index.php/catalog/3418>

Internal catalog link: microdatalib.worldbank.org/index.php/catalog/9347

Data Availability Statement

“From Access to Achievement: The Primary-School-Age Impacts of an At-Scale Preschool Construction Program in Highly Deprived Communities”

All datasets used to produce the results presented in this working paper are publicly available in the World Bank Microdata Catalog. See access paths above. We have removed all personally identifiable information (PII) in the data of this package. We are committed to ensuring the privacy and security of the data. If any replicators identify a PII variable that was missed during the deidentification process, we kindly request that they notify us immediately so that we can correct the datasets.

Marina Bassi

Program Leader

mbassi@worldbank.org

Lelys Dinarte-Diaz

Economist, DECHD

ldinartediaz@worldbank.org