Reproducible Research Repository

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

First Submission: RR_MOZ_2024_138
María Reyes Retana
reproducibility@worldbank.org
June 10th, 2024

This review verifies the reproducibility of the exhibits included in the paper *"From Access to Achievement: The Primary-School-Age Impacts of an At-Scale Preschool Construction Program in Highly Deprived Communities"*.

Contents in this review:

1. Main findings

2. List of exhibits and reproducibility status

3. Reproduction Environment

Main findings

- The code was successfully executed on a new computer after:

  1. Changing the main do file paths and running it.

- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.

- The code has two run options:

  1. Option 1: The current setup takes approximately 10 minutes to run. This will produce nearly all the results in the paper, except for the Westfall and Young (1993) p-values, which are displayed in square brackets.

  2. - Option 2: To obtain these p-values using 10,000 replications, follow the instructions in the README file to change the number of replications. However, note that this will increase the run time to around 9 hours.

- We conducted our reproducibility analysis based on the paper shared by the authors in their reproducibility package.

- Every exhibit has been reproduced accurately.

- **Reproducibility Summary**:

  - **Data**: All data sources are publicly available but not all are included in the reproducibility package.
  - **Code**: All code files (from cleaning to analysis) are included in the reproducibility package.
  - **Outputs**: All outputs are generated by code included in the reproducibility package.

– **Reproducibility verification**: Reviewers had access to the same materials in the public reproducibility package.

– **Dependencies environment**: The reviewers created a new environment for dependencies using the latest versions available for each dependency at the moment of the review.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

- **Figure 1** Does not show analysis results.

- **Table 1** Reproduced.

- **Table 2** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments are necessary. This process is detailed in the main do file.

- **Figure 2** Reproduced.

- **Figure 3** Reproduced.

- **Table 3** Reproduced.

- **Table 4** Reproduced.

- **Table 5** Reproduced.

- **Table 6** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file. The p-value with the repetitions in the code shows 0.005 while the value in the paper is 0.006, but this does not jeopardize reproducibility.

- **Table 7** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

Results in the Annex

- **Figure A.1** Does not show analysis results.

- **Figure A.2** Reproduced.

- **Table A.1** Reproduced.

- **Table A.2** Does not show analysis results.

- **Table A.3** Reproduced.

- **Table A.4** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

- **Table A.5** Reproduced.

- **Table A.6** Reproduced.

- **Table A.7** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

- **Table A.8** Reproduced. The title in the code says A.7, but in the paper, this table is titled A.8. However, it is clear that the table matches.

- **Table A.9** Reproduced. The title in the code says A.8, but in the paper, this table is titled A.9. However, it is clear that the table matches.

- **Table A.10** Reproduced. The title in the code says A.9, but in the paper is shown as A.10. However, it is clear that the table matches. There is a minor difference in the number of observations in the last two columns (3764 vs 3765), but this does not jeopardize the reproducibility of the findings.

- **Table A.11** Reproduced.

- **Table A.12** Reproduced.

- **Table A.13** Reproduced.

- **Table A.14** Reproduced.

- **Table A.15** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

- **Table A.16** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file. There is a minor difference in the number of observations in the last two columns (3764 vs 3765), but this does not jeopardize the reproducibility of the findings.

- **Table A.17** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

- **Table A.18** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

- **Table A.19** Reproduced. All numbers in the table have been successfully reproduced. However, to match the exact values of the p-values in brackets, which are based on 10,000/2,000 repetitions, manual adjustments to the code are necessary. This process is detailed in the main do file.

- **Table A.20** Does not show analysis results.

Reproduction Environment

- Paper exhibits were reproduced in a computer with the following specifications:

  - OS: Windows 10 Enterprise

  - Processor: Intel(R) Xeon(R) Gold 6226R CPU @ 2.90GHz, 2900 Mhz, 4 Core(s), 4 Logical Processor(s)

  - Memory available: 16.2 GB

  - Software version: Stata version 18