Readme for the replication package of **Entrepreneurship education and teacher training in Rwanda**

Moussa P. Blimpo and Todd Pugatch

Updated: June 14, 2024

Abstract

This is the replication package for "Entrepreneurship education and teacher training in Rwanda," by Moussa P. Blimpo and Todd Pugatch. This folder contains all the data and code necessary for replicating the tables in the paper. The replication included in this package is for cleaning intermediate datasets and producing the analysis data and results. De-identified raw data is also available, but is not included in the replication flow. The data files are in Stata format and the replication code was written in Stata. The replication time is about 5 minutes.

Contents

1 Top-folder structure **2**

2 Replication instructions **2**

3 Dataset information **3**

1 Top-folder structure

- *master.do* – Runs cleaning do-files for intermediate datasets producing final analysis data, and runs do-files for each table in the paper
- **01_data** – Contains raw intermediate and final datasets
  - *codebook.xlsx* – Contains variable-level metadata for raw intermediate and final datasets, including variable labels when available
- **02_dofiles** – Contains cleaning and analysis do-files
- **03_results** – Output folder for analysis – contains the replicated results from running *master.do*
- **04_output** – Temporary folder for replication logs
- **05_surveyinstruments** – Contains baseline and endline survey instruments for teacher, headteacher and student surveys
- **06_rawdata_materials** – Contains de-identified raw data and do-files for creating intermediate datasets (not included in the replication flow)
  - *tracking_raw_data.xlsx* – This file contains information on how the intermediate datasets included in the replication flow were created from the raw data included in **06_rawdata_materials**. It also contains do-files (where available) that were used to create the intermediate datasets. However, the data and do-files listed here are not included in the replication flow.

2 Replication instructions

1. Open *master.do* and change the directory on line 36 to where the master.do file lives
2. Run *master.do*

3 Dataset information

This package contains data from an RCT evaluating how a comprehensive teacher training program affects the delivery of a major entrepreneurship curriculum reform in Rwanda. The baseline and endline survey data included in the package was collected from schools from teachers, students, and classroom observations. Baseline survey data occurred prior to the intervention at the beginning of the 2016 academic year. The endline survey was conducted between July and September 2018. The package also includes data from the 2012 Rwanda census.

The data is organized in the folder in the following way:

1. **01_data:**
   (a) **01_raw** contains raw intermediate datasets included in the replication flow
   (b) **02_clean** contains clean analysis datasets included in the replication flow
2. **06_rawdata_materials:**
   (a) **01_rawdata** contains raw baseline and endline survey data from students, teachers and headteachers.
      i. **01_baseline**: *Master_HeadTeacher - Copy.dta*, *Master_Teacher - Copy.dta*, and *Master_Student - Copy.dta* are the de-identified raw datafiles from each respondent population.
      ii. **02_endline**: *Student Survey Final.dta*, *Endline Head Teacher Survey Final.dta*, and *Endline Teacher Survey Final.dta* are the de-identified raw datafiles from each respondent population.