

README for Reproducibility Package

“In Tablets We Trust? Validation of the Use of Tablet-Based Tools for Land Area Measurement in Household Surveys”

Federica Petruccelli[‡], Sydney Gourlay[§], Adriana Paolantonio[§], Emanuele Clemente[§], John Ilukor[§]

Contents

1. [Overview](#)
2. [Data Availability](#)
3. [Instructions for Replicators](#)
4. [List of Exhibits](#)
5. [Requirements](#)
6. [Code Description](#)
7. [Folder Structure](#)

Overview

This reproducibility package contains code and (where access is permitted) data needed to replicate the results presented in the working paper. The analysis was carried out in StataSE (version 19).

All replication materials are contained in the folder “Replication package - Tablet”, organized into “Codes”, “Data”, and “Outputs” subfolders.

Data Availability

- ☐ All data are publicly available.
- ☒ Some data cannot be made publicly available.
- ☐ No data can be made publicly available.

The raw data underlying this analysis cannot be made publicly available. While some of the raw data can be made available following a period to allow for the several ongoing research outputs to be completed, much of the data central to this analysis cannot be made available even at a later date due to the identifying information used (i.e., geolocations of household dwellings and household land). The dataset used for analysis, which includes variables constructed from confidential data, primarily geolocation data, and is limited to the data required to replicate this work, can be made available to users upon request and is included in the replication package folder. The data was collected as part of the Uganda Climate, Land Area, and Soil Study (CLASS), which was designed and implemented by the

[‡] Department of Social Sciences and Economics, Sapienza University of Rome;

federica.petruccelli@uniroma1.it.

[§] Development Data Group, World Bank; sgourlay@worldbank.org, apaolantonio@worldbank.org, eclemente@worldbank.org, jilukor@worldbank.org.

World Bank, with a survey firm (EDI Global) as the implementing partner, and is owned by the World Bank. The Uganda CLASS is a sub-national methodological survey aimed at understanding the tradeoffs across different data collection approaches. The analysis for this paper relies heavily on georeferenced data collected directly in tablets and through handheld Garmin GPS units, which cannot be made available to users outside the study team without a duly signed non-disclosure agreement.

Data Sources

- Filename: dataset_main.dta
- Source: World Bank — Uganda Climate, Land Area, and Soil Study (CLASS).
- Access: Available upon request. The confidential data from which dataset_main.dta is constructed is not publicly available.

Statement about Rights

- ☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the LICENSE.txt file.

Instructions for Replicators

1. Ensure you have authorized access to the confidential data file (dataset_main.dta) and place it in the “Data” folder, under “Processed”.
2. Install StataSE (version 19).
3. Open the do-file “Codes/01_main_analysis.do” and update the working directory by editing the line setting the global path (i.e., `gl dr "...`) – line 52 – to point to your local copy of the folder “Replication package - Tablet”.
4. From Stata, run the main do-file “01_main_analysis.do”. This script produces all tables and figures in one run.
5. Outputs are written to the “Outputs” folder. Intermediate files created during processing are written to “Outputs/intermediate”.

List of Exhibits

The provided code reproduces:

- ☐ All numbers provided in text in the paper
- ☐ All tables and figures in the paper
- ☒ Selected tables and figures in the paper, as explained and justified below

All tables and figures listed below are reproduced by the code, except for Figure 1. Figure 1 is composed of images captured from confidential shapefiles. Given the sensitivity of these shapefiles – which identify the raw locations of household land – these shapefiles are not

included as part of the replication package. The final version of the figure is provided directly in the Outputs folder.

Exhibit	Output filename	Script	Note
Table 1	Table1.xlsx	Codes/01_main_analysis.do	Saved in the Outputs/Main/Tables
Table 2	Table2.xlsx	Codes/01_main_analysis.do	Saved in the Outputs/Main/Tables
Table 3	Table3.csv	Codes/01_main_analysis.do	Saved in the Outputs/Main/Tables
Table 4	Table4.csv	Codes/01_main_analysis.do	Saved in the Outputs/Main/Tables
Table 5	Table5.csv	Codes/01_main_analysis.do	Saved in the Outputs/Main/Tables
Figure 1	Figure 1 (provided; see Outputs folder)	Codes/01_main_analysis.do	Provided only (not generated); uses confidential shapefiles not included.
Figure 2	Figure2.png	Codes/01_main_analysis.do	Saved in the Outputs/Main/Figures
Figure 3	Figure3.png	Codes/01_main_analysis.do	Saved in the Outputs/Main/Figures
Figure 4	Figure4.png	Codes/01_main_analysis.do	Saved in the Outputs/Main/Figures
Figure 5	Figure5.png	Codes/01_main_analysis.do	Saved in the Outputs/Main/Figures
Table A1	TableA1.xlsx	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Table A2	TableA2.xlsx	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Table A3	TableA3.csv	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Table A4	TableA4.csv	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Table A5	TableA5.csv	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Table A6	TableA6.xlsx	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Table A7	TableA7.csv	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Tables
Figure A1	FigureA1.png	Codes/01_main_analysis.do	Saved in the Outputs/Annex/Figures

Requirements

Computational and Software Requirements

- Software: StataSE 19.
- Operating system: Any OS supported by StataSE 19 (Windows/macOS/Linux).
- Additional Stata packages: install any user-written commands required by `01\_main\_analysis.do` as prompted by the script or documented within the do-file header.
- Runtime and storage: Less than 10 mins; ensure sufficient write permissions and disk space for the Outputs and Outputs/intermediate folders.

Code Description

- Codes/00_dataset_creation.do
 - This do-file creates the main dataset for the analysis. The input datasets used are confidential. This code is included for information but is not run as part of the package due to confidentiality of the underlying data.
- Codes/01_main_analysis.do

- Main do-file that sets file paths (via the global `dr`), loads the analysis dataset, creates key variables for the analysis, runs the statistical analysis, and generates all tables and figures saved in the Outputs folder.

Folder Structure

```
Data
├── Raw - Confidential (suppressed for confidentiality)
└── Processed (dataset_main.dta)
Codes
├── 00_dataset_creation.do (not run if no access to confidential data)
└── 01_main_analysis.do
Outputs
├── intermediate
├── Main
│   ├── Tables
│   └── Figures
└── Annex
    ├── Tables
    └── Figures
```