

README Reproducibility Package

The State of Maintenance: Can Government and Citizens Cooperate for Improved Water Access?

by Aidan Coville, Daniel Rogger, Jerome Sansonetti, and Luca Stanus-Ghib

Version: February 2026

Corresponding author for reproducibility: Luca Stanus-Ghib, lstanusghib@worldbank.org

Overview

This reproducibility package accompanies the working paper presenting the results of the impact evaluation assessing a water maintenance intervention in rural Tanzania.

The package contains all code and data needed to replicate the tables and figures included in the paper. The datasets provided in this package are the deidentified versions of the study dataset.

To replicate the results:

- Under the "Code" subfolder, open 01_master.do.
- Update the globals at the top of the master dofile to reflect the file path where the package is saved after downloading. The code currently includes a Windows-based macro setup, that would need to be modified accordingly for Mac users.
- Run the full script. The code will draw from the datasets saved in the 'Data' folder and save output in the 'Output' folder.

Folder Structure

Maji Endelevu Paper - Reproducibility Package/

└─ Data/

|— Code/

|— Output/

List of output

- Figure 1: Water Infrastructure Functionality Rates across Africa. This is excluded from the reproducibility package as it is not generated via data analysis, rather by compiling manually extracted figures from a literature review and manually generating the graph in Excel.
- Figure 2: Survival and Functionality Rates of Water Points over Time. The code generates to separate figures (for panels A and B), which are merged into a single figure within the LaTeX code. Output file names: (1) Figure XA_Survival rate CDMT - Balanced panel Aug2018 Full_Below5yo sample combined.png; (2) Figure XB_Survival rate CDMT - Balanced panel Aug2018 Full_Below5yo sample combined.png.
- Figure 3: Distribution of Water Point Hardware Issues. Output file name: Figure X_hardware_issues_share_aggregate.png.
- Figure 4: Distribution of Challenges in Water Infrastructure Maintenance. Output file name: Fig4.png.
- Table 1: Improving Water Functionality Through Co-Production. Output file name: TableX_Distribution_Point_Functionality_binary_baseline_func_0.5_0.tex.
- Table 2: Village Maintenance Activities. Output file name: TableX_Village_Maintenance_Index_3_Panels_equal_weight_non-std_binary_baseline_func_0.5_0.tex.
- Table 3: District Support to Villages. Output file name: TableX_District Maintenance Index 3 Panels_equal_weighting_non-std_binary_baseline_func_0.5_0.tex.
- Table 4: Village-Level Investments by Sector. Output file name: TableX_Village_Investments_by_Sector_Consistent Sample.tex.
- Table 5: Coordination Channels. Output file name: TableX_Channels_binary_baseline_functionality_0.5_0.tex.
- Appendix Table B1: Spillover Analysis. Output file name: AppendixX_Spillovers.tex.
- Appendix Figure B1: Photos from Maji Endelevu Consultation.

- Appendix Table B2: Water Infrastructure Functionality Rates across African Countries. Presents list of countries and rates used to generate Figure 1. No analysis carried out.
- Appendix Figure B2: Geographic Distribution of Villages in the Study. Output file name: Figure X_Map treatment villages.png.
- Appendix Table B3: List of Outcomes. No analysis carried out.
- Appendix Table B4: District Maintenance Index - Calculation Details. No analysis carried out.
- Appendix Table B5: Village Maintenance Index - Calculation Details. No analysis carried out.
- Appendix Table B6: Baseline Balance. Output file name: AppendixX_balance_table.tex.
- Appendix Table B7: Robustness: District Support to Villages – Alternative Index Weighting. Output file name: TableX_District Maintenance Index 3 Panels_anderson.tex.
- Appendix Table B8: Robustness: Village Maintenance Activities – Alternative Index Weighting. Output file name: TableX_Village_Maintenance_Index_3_Panels_Anderson.tex.
- Appendix Table B9: Robustness: Water Point Functionality – Alternative Baseline Heterogeneity Definition. Output file name: TableX_Distribution_Point_Functionality.tex.
- Appendix Table B10: Robustness: Village Maintenance Activities – Alternative Baseline Heterogeneity Definition. Output file name: TableX_Village_Maintenance_Index_3_Panels_equal_weight_non-std.tex.
- Appendix Table B11: Robustness: District Support to Villages – Alternative Baseline Heterogeneity Definition. Output file name: TableX_District Maintenance Index 3 Panels_equal_weighting_non-std.tex.
- Appendix Table B12: Village Maintenance Index - All Variables (Raw values). Output file name: AppendixX_Village Maintenance Index Full Variables (Raw scores).tex.
- Appendix Table B13: District Maintenance Index - All Variables (Raw values). Output file name: AppendixX_District Maintenance Index Full Variables (Raw scores).tex.
- Appendix Table B14: Village-Level Investments by Sector - Baseline Functionality Interaction. Output file name: TableX_VEOInvest_3Panels_Interaction.tex.

Code

01_master.do — Master do-file that sets up the environment and runs the whole reproducibility workflow (orchestrates other do-files).

02_treatment_randomization.do — This dofile is optional and it includes the randomization process conducted at the time of the baseline for transparency. It does not generate any output that is used in subsequent dofiles or in the paper itself.

04_cleaning - water point survey.do — Cleans and standardizes the water-point survey data.

04_cleaning - village survey.do — Cleans and standardizes the village water organization survey data.

04_cleaning - district survey.do — Cleans and standardizes the raw district survey data.

04_cleaning - VEO survey.do — Cleans and standardizes the VEO (Village Executive Officer) survey data.

05_construct - water point survey.do — Builds the analysis-ready water point dataset.

05_construct - village survey.do — Builds the analysis-ready village water organization dataset.

05_construct - balance_tables.do — Constructs the baseline data for the balance tables used in the paper.

05_construct - district survey.do — Builds the analysis-ready dataset derived from the district survey.

05_construct - VEO survey.do — Builds the analysis-ready VEO survey dataset.

05_construct - district_survey_add-on_VEO_merge.do — Merges the constructed district survey data with the constructed VEO dataset.

05_construct - consultation tracking.do - Categorizes challenges discussed during Round 4 of Maji Endelevu consultations for use in Figure 4.

06_analysis - table1.do — Generates Table 1.

06_analysis - table2.do — Generates Table 2.

06_analysis - table3.do — Generates Table 3.

06_analysis - table4.do — Generates Table 4.

06_analysis - table5.do — Generates Table 5.

07_descriptives - fig2.do — Creates Figure 2.

07_descriptives - fig3.do — Creates Figure 3.

07_descriptives - fig4.do — Creates Figure 4.

06_analysis - appendix_tableB1.do — Generates Appendix Table B1.

07_descriptives - appendix_figB2.do — Creates Appendix Figure B2.

06_analysis - appendix_tableB6.do — Generates Appendix Table B6.

06_analysis - appendix_tableB7.do — Generates Appendix Table B7.

06_analysis - appendix_tableB8.do — Generates Appendix Table B8.

06_analysis - appendix_tableB9.do — Generates Appendix Table B9.

06_analysis - appendix_tableB10.do — Generates Appendix Table B10.

06_analysis - appendix_tableB11.do — Generates Appendix Table B11.

06_analysis - appendix_tableB12.do — Generates Appendix Table B12.

06_analysis - appendix_tableB13.do — Generates Appendix Table B13.

06_analysis - appendix_tableB14.do — Generates Appendix Table B14.

Software and Computational Requirements

The code was last run on a laptop carrying Windows 11 Enterprise with 32.0 GB of RAM and an Intel(R) Core(TM) Ultra 7 165U (2.10 GHz) processor. It was last run on StataNow/MP 19.5 declaring Stata 18 version. The master do-file checks for missing packages and installs them locally if necessary. Internet connection is recommended during first run, and the user can select which packages to install by editing the related global in the master dofile.

Estimated runtime: ~3 minutes

Required storage space: ~350MB

Data Availability Statement

- PbR Baseline Community -- CMO Survey.dta. Baseline village water organization survey data collected by the research team in 2018 in partnership with the Tanzanian Ministry of Water. Used in 02_treatment_randomization.do to show that the randomization process is reproducible. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Treatment Assignment - Village level.dta and Treatment Assignment - District level.dta. Village- and district-level information with treatment assignment from the baseline randomization. Used to merge information into construction of analysis datasets. Microdata catalog upload pending, access will be restricted to World Bank staff.

- dp_repeat_group.dta. Endline water infrastructure survey collected by the research team in 2023 in partnership with the Tanzanian Ministry of Water (water point level data). Used in 04_cleaning - water point survey.do. Microdata catalog upload pending, access will be restricted to World Bank staff.

- wp_repeat_group.dta. Endline water infrastructure survey collected by the research team in 2023 in partnership with the Tanzanian Ministry of Water (water outlet level data). Used in 04_cleaning - water point survey.do. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Maji Endelevu endline DP mapping_clean.dta. Endline water infrastructure survey collected by the research team in 2023 in partnership with the Tanzanian Ministry of Water (village level data, it works as a "wrapper" for the water-point and water-outlet level repeat-group datasets described above). Used to merge village-level information about the water points that was not collected as part of the repeat groups described above. The suffix _clean is used to bring it on par with the other datasets, but no cleaning tasks had to be performed on this dataset (which was just deidentified before using), hence no specific "cleaning" dofile is provided for this dataset. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Maji Endelevu endline WCG_deid.dta. Endline village water organization survey data collected by the research team in 2023 in Tanzania in partnership with the Tanzanian Ministry of Water. Used in 04_cleaning - village survey.do. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Maji Endelevu Civil Servants Endline Questionnaire__raw_corr_deidentified.dta. Endline RUWASA district water office manager survey data collected by the research team in 2023 in Tanzania in partnership with the Tanzanian Ministry of Water. Used in 04_cleaning -

district survey.do. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Maji Endelevu endline VEO_deid.dta. Endline Village Executive Officer survey data collected by the research team in 2023 in Tanzania in partnership with the Tanzanian Ministry of Water. Used in 04_cleaning - VEO survey.do. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Maji Endelevu endline VEO_deid.dta. Endline Village Executive Officer survey data collected by the research team in 2023 in Tanzania in partnership with the Tanzanian Ministry of Water. Used in 04_cleaning - VEO survey.do. Microdata catalog upload pending, access will be restricted to World Bank staff.

- PbR Baseline Community -- DP Add-On with repeats.dta. Baseline water infrastructure survey data collected by the research team in 2018 in Tanzania in partnership with the Tanzanian Ministry of Water. Merged with the endline data to be used as part of the heterogeneity analyses shown in the main tables. Microdata catalog upload pending, access will be restricted to World Bank staff.

- Maji_Endelevu_Consultation_Assessment_Interview_Round_4_Reviewer_Clean.dta. Consultation tracking data collected by the research team in 2023 in partnership with Istituto OIKOS, the implementing partner. Used to generate Figure 4. Microdata catalog upload pending, access will be restricted to World Bank staff.

- DP_all_clean_new_constructed.dta. Administrative data collected by the Government of Tanzania through their Central Data Management Team, covering periods 2019-2020. Shared for the purpose of assessing reproducibility, but it will not be archived via the microdata catalog as it is owned by the Tanzanian government. Access is thus restricted. Interested users can contact RUWASA at dg@ruwasa.go.tz to inquire about administrative data access.

- DM_village_lga_list.csv. Master list of district, village, LGA names provided by the Tanzanian government. Used to correctly assign place labels. Microdata catalog upload pending, access will be restricted to World Bank staff.

- regions folder. Contains shapefiles and dependencies for geographic boundaries, downloaded by the OCHA-provided Humanitarian Data Exchange (more details here: <https://data.humdata.org/dataset/cod-ab-tza>). Accessed March 2025. Used to generate the map of Tanzania with regional boundaries (Figure B2). Data used under Creative Commons Attribution 4.0 International license.

The authors declare that they have legitimate access to all data used in the manuscript.