## Title

Factor Market Failures and the Adoption of Irrigation in Rwanda

Authors: Maria Jones, Florence Kondylis, John Loeser, and Jeremy Magruder

Data Repository: https://www.openicpsr.org/openicpsr/project/159061

## Overview

The code in this replication package (executed by 00master.R) constructs data frames used for analysis from 13 raw data files (listed below) using R (01construct.R) and then produces analysis exported in 24 tables and 12 figures used in the paper (02analysis.R).

## Data Availability and Provenance Statements

The data are from household surveys collected by enumerators (including plot boundaries mapped using handheld GPS devices), and from Landsat. Landsat data are from Landsat 7 Level 2, Collection 2, Tier 1 and were extracted from Google Earth Engine using the rgee and rgeeExtra packages in R; Landsat data are public domain.

### Statement about Rights

- [X] I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

### Summary of Availability

- [X] All data **are** publicly available.

### Details on each Data Source

All data files are .rds and can be read using R. All data files contain data from household surveys, with the exception of rds/landsatdf.rds, which contains data from Landsat with identifiers to link to data from household surveys. A data dictionary is provided in DATADICTIONARY.txt.

### Dataset list

| Data file | Notes | Provided |
|---|---|---|
| rds/areas.rds | Areas at the plot level | Yes |
| rds/distm.rds | Centroid-to-boundary distances between plots | Yes |
| rds/h.rds | Panel dataset at the household level | Yes |
| rds/hbaseline.rds | Baseline household characteristics | Yes |
| rds/hpbaseline.rds | Baseline plot characteristics | Yes |
| rds/hps.rds | Panel dataset at the plot-by-season level | Yes |

| Data file | Notes | Provided |
|---|---|---|
| rds/hpsc.rds | Panel dataset at the plot-by-season-by-crop level | Yes |
| rds/hpssamp.rds | Panel dataset at the plot-by-season level on attrition | Yes |
| rds/hs.rds | Panel dataset at the season level | Yes |
| rds/landsatdf.rds | NDVI at the plot-by-satellite image level | Yes |
| rds/nearwug.rds | Nearest water user group at the plot level | Yes |
| rds/prices.rds | Price panel dataset at the district-by-season level | Yes |
| rds/trackps.rds | Tracking dataset at the plot-by-season level | Yes |

Additional details on each data source are available in DATADICTIONARY.txt

## Computational requirements

### Software Requirements

R (code was last run with version 3.6.3).

The following comma separated list of packages was also used: ggtext_0.1.1, viridis_0.6.2, viridisLite_0.4.0, tikzDevice_0.12.3, scales_1.1.0, ggmap_3.0.0, rgeos_0.5-2, gridExtra_2.3, lfe_2.8-5, Matrix_1.2-18, rgdal_1.5-23, sp_1.4-5, readstata13_0.9.2, forcats_0.4.0, stringr_1.4.0, dplyr_1.0.7, purrr_0.3.3, readr_1.3.1, tidyr_1.0.0, tibble_3.1.5, ggplot2_3.3.5, tidyverse_1.3.0, magrittr_1.5, foreign_0.8-75, matrixStats_0.55.0, httr_1.4.1, jsonlite_1.6, modelr_0.1.5, Formula_1.2-3, assertthat_0.2.1, cellranger_1.1.0, pillar_1.6.3, backports_1.1.5, lattice_0.20-40, glue_1.4.2, gridtext_0.1.4, rvest_0.3.5, colorspace_1.4-1, sandwich_2.5-1, plyr_1.8.5, pkgconfig_2.0.3, broom_0.7.9, haven_2.2.0, xtable_1.8-4, jpeg_0.1-8.1, generics_0.0.2, ellipsis_0.3.2, withr_2.1.2, cli_2.0.1, crayon_1.3.4, readxl_1.3.1, fs_1.3.1, fansi_0.4.1, xml2_1.2.2, tools_3.6.3, hms_0.5.3, RgoogleMaps_1.4.5.2, lifecycle_1.0.1, munsell_0.5.0, reprex_0.3.0, compiler_3.6.3, rlang_0.4.11, rstudioapi_0.10, rjson_0.2.20, filehash_2.4-2, bitops_1.0-6, gtable_0.3.0, DBI_1.1.0, R6_2.4.1, zoo_1.8-7, lubridate_1.7.4, utf8_1.1.4, stringi_1.5.3, parallel_3.6.3, Rcpp_1.0.3, vctrs_0.3.8, png_0.1-7, dbplyr_1.4.2, tidyselect_1.1.1

### Memory and Runtime Requirements

The code was last run on a 12-core Intel-based laptop with Ubuntu version 20.04 with 16 GB of RAM. Computation took approximately 8 hours.

## Description of programs/code

The code in this replication package (executed by 00master.R) constructs data frames used for analysis from 13 raw data files (listed below) using R (01construct.R) and then produces analysis exported in 24 tables and 12 figures used in the paper (02analysis.R).

00master.R runs all code. 01construct.R constructs data frames used for analysis from raw data, and 02analysis.R produces analysis and exports that analysis in tables and figures.

## Instructions to Replicators

- Download all files in the repository in a folder somewhere on your computer. Name the folder 'Replication'.
- Edit the first line of 00master.R to "replicationfolder <- 'Replication' "
- Ensure that all libraries listed in "Software" have been installed and R version 3.6.3 or more recent has been installed.
- Run 00master.R to run all steps in sequence.

## List of tables and programs

| Figure/Table # | Output file |
| --- | --- |
| Table 1 | tabfig/tabs/croptab.tex |
| Table 2 | tabfig/tabs/reg_balplotrdall_small.tex |
| Table 3 | tabfig/tabs/reg_balhrd_small.tex |
| Table 4a | tabfig/tabs/reg_plotrd_smalldry.tex |
| Table 4b | tabfig/tabs/reg_plotrd_smallrainy.tex |
| Table 6 | tabfig/tabs/reg_balmip_all_small.tex |
| Table 7 | tabfig/tabs/reg_plotmip_smallnoca.tex |
| Table 8 | tabfig/tabs/reg_plotmip_small.tex |
| Table 9 | tabfig/tabs/reg_plotmiphet_small.tex |
| Table A1 | tabfig/tabs/reg_landsat.tex |
| Table A2 | tabfig/tabs/reg_landsatmip.tex |
| Table A3 | tabfig/tabs/reg_balplotrdother_small.tex |
| Table A4a | tabfig/tabs/reg_BL_plotrd_smalldry.tex |
| Table A4b | tabfig/tabs/reg_BL_plotrd_smallrainy.tex |
| Table A5 | tabfig/tabs/reg_plotrdattrit.tex |
| Table A6a | tabfig/tabs/altsp_sp.tex |
| Table A6b | tabfig/tabs/altsp_mip.tex |
| Table A7 | tabfig/tabs/reg_BL_plotmip_smallnocadry.tex |
| Table A8 | tabfig/tabs/reg_BL_plotmip_small.tex |
| Table A9 | tabfig/tabs/reg_plotmip_smallnocarainy.tex |
| Table A10 | tabfig/tabs/reg_plotmip_smallrainy.tex |
| Table A11 | tabfig/tabs/reg_hrd_small.tex |
| Table A12 | tabfig/tabs/reg_xseccontrol.tex |
| Table A13 | tabfig/tabs/reg_exp.tex |
| Figure 1 | tabfig/figs/timeline.pdf |
| Figure 2b | tabfig/figs/alphardd_irr.pdf |
| Figure 3a | tabfig/figs/yield.pdf |
| Figure 3b | tabfig/figs/hhlab_profit.pdf |
| Figure 4 | tabfig/figs/irr_subst.pdf |

| Figure/Table # | Output file |
|---|---|
| Figure 5 | tabfig/figs/expfiga.pdf |
| Figure A2 | tabfig/figs/adoptiondynamicsshort.pdf |
| Figure A3 | tabfig/figs/rdfig.pdf |
| Figure A4 | tabfig/figs/mipfig_tall.pdf |
| Figure A5 | tabfig/figs/wages.pdf |
| Figure A6a | tabfig/figs/prices_karongi.pdf |
| Figure A6b | tabfig/figs/prices_nyanza.pdf |

## References

## Acknowledgements