

Replication data and do-files for
Simon Heß, Matthias Schündeln, and Dany Jaimovich:
Development Projects and Economic Networks:
Lessons From Rural Gambia

Review of Economic Studies

Contact:

Simon Heß, hess@econ.uni-frankfurt.de
Matthias Schündeln, schuendeln@wiwi.uni-frankfurt.de
Dany Jaimovich, dany.jaimovich@utalca.cl

June 5, 2020

1 Content

This folder contains the following objects.

1. Stata do-files:

- `replicate main tables.do`, iteratively calls all do files necessary to replicate tables from the main body of the text and the appendix.
- `replicate online supplementary tables.do`, iteratively calls all do files necessary to replicate tables from the supplementary online appendix.
- `table1.do-table9.do`, produce the nine tables of the main body of the table.
- `tableA1.do-tableA3.do`, produce the three tables of the appendix.
- `supplement_table1.do-supplement_table20.do`, produce the twenty tables (and subtables) for the supplementary online appendix.
- `rerandomize.ado` contains code that does not have to be called actively. Randomization inference relies on this code to simulate the original random treatment assignment.

2. Stata data-files:

- `dyads.dta`, contains the data at the dyad level.
- `households.dta`, contains the data at the household level.
- `villages.dta`, contains data at the village level.
- `census.dta`, household-level data from then census 2013.
- `census03.dta`, individual-level data from the census 2003.
- `households09.dta`, contains household level data from the 2009 networks data set.
- `diddyads.dta`, contains merged dyad-level data from the 2009 networks data set and the 2014 networks.
- `projects.dta`, contains data at the project-report level (reports on project statuses were made by households, enumerators and a village focus group).

3. R files:

- `random forests.R`, contains the code that is used to replicate the random forest prediction. (The relevant Stata data files already contain the predicted VDC scores that are outputted from this R file. This file is only provided to show how those VDC scores are created.)

2 Instructions

To replicate the tables from the main body of the paper and the appendix, run the “`replicate main tables.do`” file. To replicate the tables in the supplementary appendix, run the “`replicate online supplementary tables.do`” file.

Prior to executing either of these do-files the following should be done within the respective do-file: (1) Specify the base folder in which all do-files are located. This folder

should also contain the included subfolder “data”. (2) Adjust the number of iterations used for randomization inference. The default setting is 100. All tables in the paper are based on 10,000 draws, but reasonable approximations of the randomization inference p -values can be expected already for 100 draws. Computation time is roughly linear in the number of draws used, so adjusting the number of draws can lead to excessive run times for the code.

Note: The do files will first install the Stata packages ‘ritest’, ‘estout’, ‘reghdfe’, ‘pdslasso’, ‘lassopack’, and ‘unique’, which are required for the replication. The do files will also run the file “**rerandomize.ado**”, which defines a program that is needed for randomization inference (which requires repeated rerandomization following the same procedure as the original treatment assignment).