Read Me
Data and Code for: "Long Run Effects of Aid: Forecasts and Evidence from Sierra Leone"
Casey, Glennerster, Miguel and Voors

All analysis was run in STATA 14.2
Additional required packages
* ssc install splitvallabels
* ssc install estout
* ssc install stripplot
* net install grc1leg2.pkg
Estimated computation time <10 minutes.

Before running the replication:
- Download and unzip the "cdd-replication" file folder

What to expect:
- The main folder contains a do-file called "cdd-replication.do" which uses the datasets in "cdd-replication/data", performs all the analysis, and outputs the tables and figures from the manuscript and appendix in their respective folders. This do-file also reproduces quantities stated in the main text of the article.
- The "cdd-replication/data" folder contains the de-identified datasets. These data files are in Stata (.dta) format.
- To run the replication, first set the directory at the top of "cdd-replication.do" to match the location where you have saved the "cdd-replication" (unzipped) folder on your device.
- The first part of the do-file formats and compiles the main variables used in the analysis. It generates the main data analysis file, entitled "SES_analysis.dta", which will appear in the "cdd-replication" folder after the program runs. It also generates several temporary intermediate datasets that are deleted at the end of the section.
- The second part of the do-file runs the analyses and outputs tables and figures in the "cdd-replication" folder

Other things to note:

- Included in the replication directory is the main analysis data file from the paper "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan" by Casey, Glennerster, and Miguel (https://doi.org/10.1093/qje/qje027). This file is titled "gbf_analysis.dta", see also: https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/21708.
- Some of the included data files correspond directly to one of the survey instruments used in data collection. These survey instruments can be found in the "Instruments" folder and are .pdf files with the questions and answer prompts given to respondents. Next to each question is a variable name corresponding to the variable name in the relevant data file. The data files correspond to survey instruments as below:

Instrument	Corresponding Data
SES – Endline Survey 2016.pdf	public_SES.dta
Manager selection – tally sheet enumerator A.pdf	
Manager selection – tally sheet enumerator B.pdf	
Managerial Capital Test.pdf	
Expert Survey.pdf	public_expert_data.dta

- Note that the dataset "public_SES.dta" contains additional data from other instruments than "SES – Endline Survey 2016" (Manager selection – tally sheet enumerator A and B, Managerial Capital Test). See the replication files for the companion paper "Skill Versus Voice in Local Development".
    - Available on Harvard Dataverse: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/C8YCXJ
    - See the ReadMe for that replication for the mapping of instruments to corresponding data.

Detailed Inventory:

"cdd-replication.do" uses the following de-identified datasets:
- public_SES.dta
- public_expert_data.dta

And the following public replication dataset:
- gbf_analysis.dta

And generates the following outputs for tables in the main text and appendix:
- Table 1: Table_1_2009.csv, Table_1_2016.csv, Table_1_decay.csv
- Table 2: Table_2.csv
- Appendix Table A1: Table_A1_2009.csv, Table_A1_2016.csv, Table_A1_decay.csv
- Appendix Table A3: Table_A3.csv
- Appendix Table A4: Table_A4.csv
- Appendix Table A5: Table_A5.csv
- Appendix Table A6: Table_A6_BOMBALI.csv, Table_A6_BONTHE.csv
- Appendix Table A7: Table_A7.csv

And the following outputs for figures in the main text and appendix:

- Figure 3: fig3.png
- Figure 4: fig4.png
- Figure 5: fig5_a.png, fig5_b.png
- Figure 6: fig6_panela.png, fig6_panelb.png, fig6_panelc.png
- Figure A1: figa1.png
- Figure A2: figa2.png
- Figure A3: figa3_panela.png, figa3_panelb.png