

Reproducibility package for Spillovers in ICT Adoption from Formal to Informal Firms: Evidence from Zambia

2024-04-19

Running Instructions

To run the code from the start the user first get access to the sources listed below and not included in the package (not included), follow the naming conventions and put them in the folder Data/.

Then you will need to change the directory in line 13 of the “Do_File.do” file.

If replicators want to proceed from the intermediate data (line 225), you can do so, by only changing the file path in the Do file, and by running the code. This will use the intermediate dataset Data.dta and produces all the results in the paper.

Data Statement

This research utilizes data from public sources. The final dataset combines the following three datasets (all accessed in July 2022): While all data used in this paper is publicly available, some of it is not republishable due to restrictions and sensitivities. Specifically, the GPS coordinates and the enumeration forms from the WBMES and WBISES have restricted access due to confidentiality of the respondents. Disclosing this information could reveal the identity of the business. The GPS coordinates are used only for the calculation of distances and densities. The data entries in this package outline all sources, including how to access the public but non-reshareable data, along with instructions for their download and placement in the specified directories for full replication.

- a. Zambia 2019 World Bank Enterprise Survey (WBES) (link: <https://microdata.worldbank.org/index.php/catalog/3763/>) file should be named Zambia-2019-full data.dta and saved into the “Data” folder (not included)
- b. Zambia 2019 World Bank Micro Enterprise Survey (WBMES) (link: <https://microdata.worldbank.org/index.php/catalog/3746/>) file should be named ZambiaMicro-2019-full data-long form.dta and saved into the “Data” folder (not included)
- c. Zambia 2019 Informal Sector Enterprise Survey (WBISES) (link: <https://login.enterprisesurveys.org/content/sites/financeandprivatesector/en/library/library-detail.html/content/dam/wbgassetshare/enterprisesurveys/economy/zambia/ZambiaInformal-2019-full-data-long-form.dta>) file should be named ZambiaInformal and Micro-2019-full.dta and saved into the “Data” folder/ (not included)

*Note that you will need to login as a WB staff or create an account to be able to download these data files.

In addition some auxiliary files are used in the analysis.

- a. Productivity.dta, SquareDensity.dta, squareid.dta and Dist_Closest(ES/Formal/Formal_ALL_Weak/Formal_Web). These data files were manually created by the authors using restricted GPS coordinates from the Enterprise Survey, therefore the initial data is not shared. However, these dataframes are included in the package. (included)

The datasets containing distances and the productivity of the closest formal firm were created using sensitive GPS location data. For this reason, we only provide the final measures without the source data.

- b. NTL.dta Available here <https://www.arcgis.com/home/item.html?id=edabcbb5407547f5bc883018eb6e7986>. This dataframe is available in the package with some manual modifications that included adding the firms GPS location. (included)
- c. Altitude.dta. This dataset is a product of the Zambia Informal and Sector Enterprise Survey and cannot be shared. During the data collection, which was implemented by a private firm hired on behalf of the World Bank Enterprise Analysis Unit (DECEA), each business was enumerated and their GPS coordinates were captured. This includes the latitude, longitude, and altitude of the location of each business. To protect the identity of the respondents as a promise, and as a legal commitment of the WB, the Enterprise Analysis unit is NOT allowed to publish this information. (not included)

To request access for the restricted GPS coordinates data (Altitude and the GPS the authors used for some of the auxiliary datasets), please contact enterprisesurveys@worldbank.org.