****

****

**README DO-FILES AND DATASETS USED IN THE ANALYSIS
**Small Firm Death in Developing Countries
**March 24, 2018
**David McKenzie (dmckenzie@worldbank.org) and Anna Luisa Paffhausen (apaffhausen@worldbank.org)

**This readme explains the files and steps needed to replicate the analysis for the following paper: “Small Firm Death in Developing Countries” by David McKenzie and Anna Luisa Paffhausen, revised version (March 2018)
**The analysis was performed with Stata, version 14.2
****

****

****

The folder “Do-files and readme - Revised version” contains datasets and do-files used in the analysis presented in the paper “Small Firm Death in Developing Countries” by David McKenzie and Anna Luisa Paffhausen, revised version (March 2018). These can be used to replicate the analyses and follow the construction of the underlying dataset. All analyses, data cleaning and manipulation were conducted with Stata, using version 14.2. When replicating the do-files, the analysis might make use of Stata packages that are not yet installed on your computer (e.g. mat2txt, outreg2, coefplot). Please install them to run the do-file without interruptions.

Note: You may rename this folder, which is currently named “Do-files and readme”, but have to leave its structure, names of files and folders unchanged to ensure that all paths referred to in the do-files will work.

***FOR (REPLICATION OF) FIGURES AND TABLES SHOWN IN THE PAPER
1) Open the do-file FiguresAndTablesRevisedVersion.do
2) Change the directory to this folder (currently named “Do-files and readme - Revised version”) on your computer (This is noted in the do-file and an example is given).
3) Run the do-file
The do-file will create the figures and replicate the tables shown in the paper, including appendix figures and tables. They will be stored in the folders “Do-files and readme - Revised version/Figures”, “Do-files and readme - Revised version/Tables”, and “Do-files and readme - Revised version/Appendix Tables and Figures”. It uses CombinedMaster_RH.dta and data stored in the folder “Preparation of data for Figures/ Data for figures”.

***FOR (REPLICATION OF) PREPARATION OF THE FIRM PANEL DATASET FOR THE FIGURES SHOWN IN THE PAPER

There are two do-files, which prepare the data for the figures shown in the paper: PrepDeathRatesForFigures.do and Recode_opnewfirm.do. You will find them in the folder “Do-files and readme - Revised version/Preparation of Data for Figures”. Recode_opnewfirm.do is included for informational purposes only, but cannot be used for replication as the underlying raw data is not made available. PrepDeathRatesForFigures.do can be replicated before running FiguresAndTablesRevisedVersion.do, but there is no need to do so, as all necessary files are already available in the folder “Do-files and readme - Revised version/ Preparation of Data for Figures/ Data for figures”.

1) PrepDeathRatesForFigures.do
Note: You can run this before running FiguresAndTablesRevisedVersion.do but there is no need to do so, as the folder “Do-files and readme - Revised version/Preparation of Data for Figures/ Data for figures” already contains the necessary files.
1 - 1) Go to the folder “Do-files and readme - Revised version/Preparation of Data for Figures“
1 - 2) Open the do-file PrepDeathRatesForFigures.do
1 - 3) Change the directory to this folder (currently named “Do-files and readme - Revised version”) on your computer (This is noted in the do-file and an example is given).
1 - 4) Run the do-file
The do-file will prepare the firm panel dataset for the figures shown in the paper and store the created datasets in the folder “Do-files and readme - Revised version/ Preparation of Data for Figures/Data for figures”.

2) Recode_opnewfirm.do
This do-file prepares the firm panel dataset for appendix figure A.2. It is included for informational purposes only, but cannot be used for replication as the underlying raw data is not made available

***FOR (REPLICATION OF) CONSTRUCTION AND PREPARATION OF MASTER DATASET FOR ANALYSIS
Notes:
1) There is no need to replicate the construction and preparation of the master dataset before running the replication do-files PrepDeathRatesForFigures.do and FiguresAndTables.do. The folder “Do-files and readme” already contains the necessary files.
2)The folder “Do-files and readme - Revised version” does not include the raw data for the construction of the firm panel dataset coming from the individual surveys. So the do-files preparing the data from each survey for combination into the master dataset do only have informational character. They cannot be replicated. Each do-file, however indicates where the corresponding raw data can be downloaded from, in case it is publicly available, and notes when it is not (yet) available. If you would like to replicate the construction with one or more of the publicly available survey data, you have to download the data from the source given and modify the paths in the do-files accordingly.

1) In case you want to replicate one of the do-files preparing the data from each survey for combination:
1 - 1) Go to the folder “Do-files and readme - Revised version/Construction of Dataset“
1 - 2) Open the do-file :
For Benin - Cotonou Informal Enterprise Survey: BENINFORM.do
For Egypt - Macroinsurance for Microenterprises Survey: EgyptMacroins.do
For Ghana - Microenterprises Survey: GhanaFLYP.do
For Indonesia - Family Life Survey: IFLS.do
For Kenya - GET Ahead Business Training Program Impact Evaluation Survey: KENYAGETAHEAD.do, FirmDeathExtractionKenya.do
For Malawi - Business Registration Impact Evaluation Survey: MALAWIFORM.do
For Mexico - Family Life Survey: MXFLS.do
For Nigeria - General Household Survey: NGLSMSISA.do
For Nigeria - YouWIN! National Business Plan Competition Impact Evaluation Survey: NGYOUWIN.do
For Sri Lanka - Impact Evaluation Survey of Business Training for Women: SLKFEMBUSTRAINING.do
For Sri Lanka - Informal Enterprises Survey: SLKINFORMALITY.do
For Sri Lanka - Longitudinal Survey of Enterprises: SLLSE.do, ExtractSLLSEforfirmdeath.do
For Sri Lanka - Microenterprise Survey: SLMS.do
For Togo - Lomé Informal Enterprise Survey: TOGOINF.do
For Thailand - Townsend Thai Survey: TTHAI.do
For Uganda - Women’s Income Generating Support Impact Evaluation Survey: UG_WINGS.do
1 - 3) Download all rounds of data needed from the source given in the do-file
1 - 4) Change the directory (This is noted in the do-file).
1 - 5) Change the paths to the data you are using (This is noted in the do-file).
1 - 6) Make sure the final dataset is saved in the folder “Do-files and readme/ Construction of Dataset/Data for combination”. You can do this by either changing the directory accordingly or the path for saving the final dataset.
1 - 7) Run the do-file
The do-file will use and prepare the data from the individual survey for combination into the master dataset and store it in the folder “Do-files and readme - Revised version/ Construction of Dataset/Data for combination”.

2) To combine the prepared data from the individual surveys into the master dataset. (Note that we included all necessary prepared datasets into the folder “Do-files and readme - Revised version/Construction of Dataset/Data for combination”, so you can replicate the combination without having replicated the preparation of the data from the individual surveys before.)
1) Go to the folder “Construction of Dataset“
2) Open the do-file CombinedMaster.do
3) Change the directory to this folder (currently named “Do-files and readme - Revised version”) on your computer (This is noted in the do-file and an example is given)
4) Run the do-file

The do-file will use the data currently stored in the folder “Do-files and readme - Revised version/Construction of Dataset/Data for combination”, create the dataset CombinedMaster_RH.dta and store it in the folder “Do-files and readme”.

****
OVERVIEW OF CONTENT AND STRUCTURE OF THIS FOLDER NAMED “Do-files and readme”

The folder contains:
1) This readme-file

2) The master dataset CombinedMaster_RH.dta, used in the analysis

3) The replication do-file FiguresAndTablesRevisedVersion.do:
It creates the figures and replicates the tables shown in the paper, including appendix figures and tables, and stores them in the folders “Do-files and readme - Revised version/Figures”, “Do-files and readme - Revised version/Tables”, and “Do-files and readme - Revised version - Revised version/Appendix Tables and Figures”. The do-file uses CombinedMaster_RH.dta and data stored in the folder “Do-files and readme/ Preparation of data for Figures/Data for figures”.

4) The folder “Figures”:
Contains the figures shown in the paper.

5) The folder “Tables”:
Contains the tables shown in the paper.

6) The folder “Appendix Tables and Figures”:
Contains the tables and figures shown in the appendix

7) The folder “Construction of Dataset”:
The folder contains:
7 - 1) The do-files that prepare the data from each survey for combination into the master dataset.

7 - 2) The do-file that combines the prepared data from each survey and constructs the master dataset:
CombinedMaster.do (uses data stored in the folder “Data for combination” and creates the master dataset CombinedMaster_RH.dta, which is stored in the folder “Do-files and readme - Revised version”)

7 - 3) The folder “Data for combination”:
Contains the survey specific prepared datasets for combination, as well as other data used in the construction of the master panel dataset, such as GDP per capita and CPIs

8) The folder “Preparation of Data for Figures”:
The folder contains:

8 - 1) The do-file PrepDeathRatesForFigures.do:
Prepares the firm panel dataset for the figures shown in the paper. Data is stored in the folder “Data for figures”.

8 - 2) The do-file Recode_opnewfirm.do:
Prepares the firm panel dataset for appendix figure A.2. This do-file has only informational purpose but cannot be used for replication, as the underlying raw data is not made available.

8 - 3) The folder “Data for figures”: Contains the datasets created in PrepDeathRatesForFigures.do, Recode_opnewfirm.do, and others used for the construction of figures shown in the paper.