

# Replication for Jedwab and Vollrath

## Calibration and simulation

The calibration of the model in the paper requires some data work to produce the targets for various runs of the calibration, and then a Matlab program that performs the calibrations themselves.

The short recipe to replication is:

1. Copy this repository to a master folder on your own computer. Leave the folder structure intact.
2. Edit the `jvmaster.do` script in the Code folder to set the working directory to your master folder.
3. Execute `jvmaster.do` in Stata
4. Edit the `mcmaster.m` script in the Code folder to set the working directory to your master folder.
5. Execute `mcmaster.m` in Matlab

The output is all written to the Drafts folder. All that follows from here is detail about the programs called.

## Data and code preparation

The Stata scripts pull summary information on urbanization and population size from a dataset of developing countries. The data files are:

1. `jv_flags.dta`: this is a set of indicators for which countries were ex-communist, etc.. that we use to control which countries belong in which sub-samples
2. `cbdcdrerni.dta`: this contains country/year information on crude birth rates, crude death rates, and crude rates of natural increase
3. `jv_data_fit.dta`: this is a set of country/year information that includes urbanization rates, population size, and various measures of slum shares

These Stata scripts write text files with summary information that is read by Matlab to perform calibrations. To set up all the data and code necessary for the calibration, you should **edit “`jvmaster.do`” to have the correct path to your master folder** and then **run `jvmaster.do`**. This script calls the following:

1. `jvextract_prepare.do`: this merges the three data files, makes a few overrides to data values, calculates the maximum slum share observed, and then reshapes the data to “wide” format - each row is a single country
2. `jvextract_robust.do`: this calculates average urbanization and slum rates, and calculates the parameters for the fertility function, for given samples.
3. `jvextract_ind.do`: this takes a given file of country data, and writes a text file of country-specific urbanization and slum shares that can be read by Matlab to perform individual-level calibrations

The output of all of these files is written to the “Work” folder, where the Matlab code will look for it.

## Adding or changing samples included in the analysis

If you want to add or alter a given sample we use in the calibration, then you have to edit the **jvextract\_robust.do** file. In this file, adding a sample requires that you call the program “calc\_pop”. For example, the line in the code

```
calc_pop if urbrate1950<20 & country~="China" & country~="India", name("ex-China and India")
```

calls the calc\_pop program for countries that have a 1950 urbanization rate less than 20%, and which are not named China or India. After the comma, the program takes an argument for how to label the sample in the output file. The only restriction here is that this “name” command **cannot include commas**.

To add a new sample, simply add a new line calling calc\_pop, with an “if” statement that determines which countries you want to include. Important variables you can select on are: 1. urbrateXXXX: urbanization rate by year XXXX 2. slumXXXX: slum share in year XXXX - this is often missing 3. slummax: this is the maximum slum share observed over all years 3. oecd: this is 0/1 for membership of the OECD 4. cdrXXXX: this is the crude death rate in year XXXX

Once you’ve done that, run the “jvmaster.do” script again, and you will get new output files in the Work folder, and those will include the data from your new sample so that it can be read by Matlab.

## Running the calibration

The master script for reproducing all results is “mcmaster.m”. You should **edit this to point to your master folder** before running it. This sets the initial values of all parameters, and fits the model to find the calibrated values of the congestion elasticities in formal and informal areas. It then calls subroutines to write results to tables, and to perform various robustness checks. It should only be necessary to call “mcmaster.m” from Matlab to reproduce all the results. All output tables are written to the “Drafts” folder.

Running the full script takes a while, depending on your machine. It is straightforward to comment out specific function calls in “mcmaster.m” to avoid producing certain tables to save time.

Remaining Matlab code in the Code folder have various functions

1. mctable2.m, mctable3.m, mctable5.m, mcrobust.m, mchistory.m: these take the baseline calibration and write out results to text files that can be used as input to tables in the paper.
2. mcindividual.m: this runs the calibration separately for each country, as well as simulating the model using the mean of the individual elasticities, the median, etc.. This writes a text file with rows formatted to be included in the same table as the robustness results from mcrobust.m
3. mceval.m, mcfix.m, mcrobusti.m, mcrobustj.m: these are subroutines called by the other scripts to produce or write results
4. mctable4.m: deprecated