

# Taken by Storm: Hurricanes, Migrant Networks, and U.S. Immigration DATA GUIDE

Parag Mahajan and Dean Yang

June 24, 2019

## Overview and Data Access

This document describes how the results can be reproduced. Tables 3 through 8 can only be replicated with access to non-public-use demographic data files, accessed through a Census Research Data Center (RDC). Specifically, researchers need to request the restricted-use versions of the 1980 Long Form Census, the 2000 Long Form Census, and the 2005 through 2015 American Community Surveys. Researchers interested in conducting a replication must be approved for access by the U.S. Census Bureau. For details, see, e.g., <https://mrdc.isr.umich.edu/using-mrdc/>.

## Setup

Each program in this .zip file references a global external and a global internal directory. The external directory (`$external_dir`) should be the folder that this .zip file has been extracted into—it should contain a `/Data/` folder, an `/Output/` folder, and a `/Programs/` folder. Replications are enabled by the programs in the `/Programs/` folder. The internal directory (`$internal_dir`) refers to a researcher's directory in the RDC. A researcher with access to RDC data will have to create the following folders in `$internal_dir` in order to use the programs in this .zip file to replicate our results:

- `$internal_dir/data/`, within which create the subfolders
  - `$internal_dir/data/Census 1980/`
  - `$internal_dir/data/Census 2000/`
  - `$internal_dir/data/ACS 'year'/` (with 'year' ranging from 2005 through 2015—this is 11 folders in total)
- `$internal_dir/external_data/`: researcher places the files located in `$external_dir/Data/external_data/` of the .zip file here
- `$internal_dir/output/`: regression model output for Tables 3-8 will be stored here

## Tables 1 and 2

Replication of Tables 1 and 2 can be completed without access to RDC data. The program `tables1_2_external.do` in the `$external_dir/Programs/` folder completes their replication, generating and placing each table in Excel format in the `$external_dir/Output/` folder.

## Tables 3 through 8

Replication of Tables 3 through 8 require the following steps, once access to RDC data is obtained:

- Generate the folders within `$internal_dir` as described above
- Use StatTransfer to convert original .sas files to .dta files for

- Each separate 1980 Census Long Form state file, placing the .dta files into `$internal_dir/data/Census 1980/`
  - Each separate 2000 Census Long Form state file, placing the .dta files into `$internal_dir/data/Census 2000/`
  - Each separate ACS file for ‘year’, placing the .dta file into `$internal_dir/data/ACS ‘year’/`, with ‘year’ ranging from 2005 through 2015
- Run, in order
    1. `1_census1980_stock.do`
    2. `2_census2000_flows.do`
    3. `3_acs_flows.do`
    4. `4_combine_flows.do`
    5. `5_finalize_panel.do`
      - Note: due to restrictions, the user will have to create the variable “country\_str” at the beginning of this program (as described in the code comments). This involves using the provided numeric country codes in the RDC documentation to generate a string variable that reflects a given country’s name. Country names will need to match those provided in this .zip file to ensure a perfect merge with external data.
  - At this point, the 159 country by 25 year panel has been created and stored as `$internal_dir/data/panel.dta`
  - Then,
    - `table3.do`: generates Table 3
    - `table4.do`: generates Table 4
    - `table5.do`: generates Table 5
    - `tables6_7.do`: generates Tables 6 and 7
    - `table8.do`: generates Tables 8
  - Output will be stored in `$external_dir/output/`