Readme file for:
"Medical Worker Migration and Origin-Country Human Capital: Evidence from U.S. Visa Policy"
by Paolo Abarcar and Caroline Theoharides

Written on September 29, 2021

This readme file describes the various attached .do files and .dta files, and how to replicate the analysis in our paper. All .do files were run using Stata 17. See the online data appendix for a description of all of the datasets, variables, and sample restrictions. If you have any questions about these data or .do files, please contact Caroline Theoharides at ctheoharides@amherst.edu.

Data Sources:
We use a number of restricted access data sources from the Philippine government in our analysis. We provide details on how to request and obtain each of these datasets. Below, we discuss the do files used to input the raw data from each data source through final analysis.

1. Migration Data from the Commission on Filipinos Overseas (CFO)
    a. Send a formal letter of request to Mr. Rodrigo V. Garcia, rgarcia@cfo.gov.ph
2. Enrollment, Graduation, and Program Data from the Commission on Higher Education (CHED)
    a. Send a formal letter of request to info@ched.gov.ph
3. Nursing Licensure Data from the Philippine Regulatory Commission (PRC)
    a. Send a formal letter of request to Ms. Henrietta Narvaez, Chief of the Archives and Records Division, records@prc.gov.ph
4. 1990, 2000, 2007, and 2010 Philippine Census of Population and Labor Force Survey Data (used for calibration exercise)
    a. Send a letter requesting the appropriate datasets to info@psa.gov.ph
5. NSCB/PSCED Data/Province Population Counts:
    a. These are crosswalks of geographic codes and education codes used by the Philippine government. We compiled these in order to link our datasets. All crosswalks are included in the Data folder. We also include a raw .csv file of total province level population counts which are used as the denominator in the calculation of migration rates. The college aged population was created using Census data from PSA as outlined in #4 above.

Do Files:
The do files "_master.do" executes all cleaning and analysis do files in order.

Do Files for Data Cleaning:
1. "0 Clean Licensure Data.do", "0.0 Clean Institution Names.do", and "1 Create Pass Rates.do" read in the raw nursing licensure exam data, clean the institution names to be consistent over time, assign province codes, and create pass and examinee rates.
2. "0 Clean Raw CFO Data.do" reads in and cleans the raw CFO data.
3. "0 Clean Raw CHED Data.do" reads in and cleans the raw CHED data. "0 enrollment and grad time series.do" creates aggregate time series for graduation and enrollment by

various PSCED codes which are used for Figures 1b and 1c. “0 Create Nurse Program Counts.do” creates the number of nursing programs over time.

4. “0 Create Population Counts.do” creates annual total population and college age population counts.
5. “0 nursecounts xxxxcen.do” creates counts of the number of nurses in each Census year. This is used to create the baseline nurses per capita control variable and to do the analysis of how the number of domestic nurses has changed over time in Section 5.6. “1 nursecounts combine.do” combines these 3 datasets.

Do File to create dataset for analysis

The do file “2 create data set.do” combines all of the datasets created by the aforementioned do files in order to run the analysis.

Do File for main analysis

The do file “3 main analysis.do” creates all the main tables and figures in the paper.

Do Files for Appendix Tables:

1. “4 appendix analysis.do” creates all appendix tables and figures except those listed below.
2. “4 appendix table 3.do” creates Appendix Table 3.
3. “4 appendix figure 2.do” creates Appendix Figure 2.
4. “4 appendix figure 3.do” creates Appendix Figure 3.
5. “4 appendix figures 5 and 6.do” creates Appendix Figures 5 and 6.
6. “4 appendix figure 8.do” creates Appendix Figure 8.