Replication Data and Code for "Age at Immigrant Arrival and Career Mobility: Evidence from Vietnamese Refugee Migration and the Amerasian Homecoming Act"

JPE Microeconomics

Public Use Data Component

Sari Pekkala Kerr, Wellesley College

William Kerr, Harvard Business School

Kendall Smith, London Business School

*** OVERVIEW

This file set contains the materials to construct the analysis and prepare the public data-based figures and tables in the paper.

*** DATA AVAILABILITY AND PROVENANCE

To gain access to the public-use data used in this paper, request the following datasets from IPUMS:

1990 and 2000 Decennial Censuses

2005-2009, 2010-2014, and 2015-2019 American Community Surveys

*** COMPUTATIONAL REQUIREMENTS

We recommend running the empirical exercises on a UNIX server that includes STATA 18 or similar environment. The initial data build can take ~45GB of memory space.

*** DESCRIPTION OF CODE

To execute, please take the following steps:

- Update the initial “cd” command on the second line of each of the two programs to be the root directory of the new environment.
- Place in the root directory the “data-input” folder and its contents.
    - This folder contains base data (e.g., combined Decennial Censuses and American Community Survey data) that are called upon by our programs.

    - The raw data for these programs are from Decennial Censuses from 1990, 2000, and the 2005-2009, 2010-2014, and 2015-2019 American Community Surveys.
    - The data on cluster site locations comes from "The Amerasians from Vietnam: A California Study" by Chung Hoang Chuong and Le Van (1994), pages 83-83. The data on the official count of Amerasian arrivals is from government documents from Office of Refugee Resettlement Reports and Secretary of State Refugee admissions reports. The mapping files in the data-input folder were created from Decennial Census and American Community Survey data, as well as specific mapping instructions on historical changes detailed by IPUMS USA. Some locations were then manually mapped to criteria like being one of the cluster site locations mentioned above.
- Create a "temp" folder and an "output" folder. The programs will populate the "temp" folder with working files and the "output" folder with the outputs of most analyses. You will need about 45GB of memory space.
- Execute the programs in the following sequence:
    - aha-prep-1-do: set-up file that prepares data for analysis.
    - aha-analysis-2.do: core file that runs data analyses.
- The core output will populate in files and sub-folders within the output folder.
    - The following exceptions are manually entered from the aha-analysis-2.log file:
        - Figure 1: Figure created using values output in the log file.
        - Table 3: Parts B-D are calculated in the log file.
        - Table 4: Calculated in the log file.