

This readme file lists and describes the programs that were used to generate the results appearing in the published version of “Trade Liberalization and Mortality: Evidence from U.S. Counties,” by Justin R. Pierce and Peter K. Schott.

Instructions

Researchers should download 20180396_programs.zip and 20180396_input.zip. The former contains all STATA programs used in the analysis and the latter contains datasets called by these programs. Each of the programs and datasets in these directories is described below.

Researchers should download the files to a master directory location of their choice, which we call /master/ here. The programs in 20180396_programs.zip should be saved to /master/programs/, and the data in 20180396_input.zip should be saved to /master/input/. Researchers should also create sub-directories /master/interim/ and /master/aeri_figures/ within the master directory, and a sub-sub-directory /master/interim/limcom4x. After obtaining the confidential microdata from the National Center for Health Statistics (see below), researchers should save the raw data in /input/raw_census/. Then, the programs described below should be run, in order, to generate all results in the paper.

Programs

The following five programs were used to generate the results used in the paper. They must be run in order.

- 0_aeri_county_to_cuma_03.do – This program groups counties into CUMAS, our name for combinations of counties and the Census Bureau's Public Use Microdata Areas, as described in Section 4 of the program.
- 1_aeri_refine_mortality_microdata_07.do – This program formats the raw confidential mortality microdata and calculates county x demographic characteristic x year age-adjusted and crude death rates.
- 2_aeri_mortality_dataprep_06.do – This program creates the independent variables used in the main regression analysis.
- 3_aeri_mortality_assemble_03.do – This program creates the main dataset used in the analysis by merging mortality rates and independent variables.
- 4_aeri_estimation_05.do – This program generates the results shown in the paper's tables and figures.

Data

This paper employs confidential microdata, specifically the “Mortality – all county (microdata)” vital statistics, which were obtained through an application to the U.S. Centers for Disease Control's National Center for Health Statistics and The National Association for Public Health Statistics and Information Systems. Information on how to apply for these data is available at <https://www.naphsis.org/research-requests>. The application process is straightforward, entailing a short form requesting a description of how the data will be used. There is no charge to use the data, and they are provided to researchers on a CD. The program 1_aeri_refine_mortality_microdata_05.do reads in these data and constructs crude and age-adjusted mortality rates. Note that this program reads from a sub-directory called /raw_cdc/,

The remaining data used in the paper are provided in the /input/ directory. Sources for these data are provided in the main text of the published paper or the appendix. The sub-directories and files contained in that directory are:

In main /input/ directory:

- CA4_1969_2016__ALL_AREAS.dta – Contains county-year-level personal income data from the Bureau of Economic Analysis.
- CA35_1969_2016_ALL_AREAS.dta - Contains county-year-level transfer data from the Bureau of Economic Analysis.
- CBP_1990adj_dorn.dta – County-by-industry employment data based on David Dorn’s imputations for the 1990 County Business Patterns. Created with DORN_cbp1990_imputations.do.
- conc_sic87_naics97.dta – 1987 SIC to 1997 NAICS concordance.
- DORN_cbp1990_imputations.do – Program that performs David Dorn’s imputations on the 1990 County Business Patterns to create CBP_1990adj_dorn.dta
- hs_mfa_phase_20111208.dta – HS-level data on products covered by MFA and phase of removal.
- hs_sic_naics_imports_89_115_20160726.dta – HS to SIC and NAICS concordance.
- imp_detl_yearly_97n.dta – HS-level U.S. import data for 1997.
- la.data.64.county – Local Area Unemployment Statistics from the Bureau of Labor Statistics.
- mfa8404.dta – Product-level MFA quota fill rates.
- naics5811.dta – NAICS Industry-year level NBER-CES Manufacturing Industry Database.
- opioids_and_state_laws.csv – State-level data on opioid laws/regulation.
- pce_deflator.dta – Personal Consumption Expenditures price deflator from the Bureau of Economic Analysis.
- sic5811.dta – SIC Industry-year level NBER-CES Manufacturing Industry Database.
- survey_all_concorded_sic.dta – SIC industry-level data on Chinese production subsidies.
- tar_val.dta – HS8-year-level data on NTR and non-NTR tariff rates and the NTR gap.
- tariff_hs02_02-05 – HS-level data on Chinese import tariffs for the years 2002-2005.
- tariff_hs96_96-01 – HS-level data on Chinese import tariffs for the years 1996-2001.
- xtract_1980census_20151114.dta – 1980 decennial Census county-level data.
- xtract_1990census_20151117.dta – 1990 decennial Census county-level data.
- xtract_2000census_20151110.dta – 2000 decennial Census county-level data.
- xtract_2010census_20151111.dta – 2010 decennial Census county-level data.

In /input/raw_census/ sub-directory:

- county_to_puma_MS_download.dta – Census Bureau county to PUMA concordance.
- puma_to_county_MS_download.dta – Census Bureau PUMA to county concordance.
- us.1990_2014.singleages.adjusted – SEER population data by county, year, and demographic group.
- In /input/raw_oasdi/ sub-directory:
- oasdi_scYY_pdf_converter.dta , where YY={99,00,01,02,03,04,05,06,07,08} – County-level data on number of Social Security disability recipients. These data, for years 1999 to 2008 were generated from pdf files using a pdf converter.
- oasdi_scYY_researcher_file.dta, where YY={09,10,11,12} – County level data on number of Social Security disability recipients.