README file for replication archive
"The Long-Run Effects of Recessions on Education and Income"
American Economic Journal Applied Economics
Bryan A. Stuart
July 2021

Repository Project ID: openicpsr-122001

Instructions and Notes

The analysis in this paper uses publicly available data, along with restricted-access data from the U.S. Census Bureau and the Panel Study of Income Dynamics (PSID). I describe the programs and data sets used in the analysis below. Restricted-access data from the Census Bureau and PSID are not included, but instructions on accessing the data are. Data from ICPSR are not included in the replication files but can be downloaded after registration with this site. To gain access to the confidential Census microdata, follow the directions here on how to write a proposal for access to the data via a Federal Statistical Research Data Center: https://www.census.gov/ces/rdcresearch/howtoapply.html. You must request the 2000 Census, 2001-2013 American Community Surveys, and NUMIDENT files in your proposal.

The do-file master_main.do references all of the do-files that can be run on publicly available data and results. See the programs/main subfolder. As noted in master_main.do, some of the do-files take as inputs the results from restricted-access Census and PSID data. To replicate the analysis, one can run the do-files between cr_cpi.do and cr_ipums_summstats.do, then complete the analysis of restricted-access data (described below), and then run the remaining files in master_main.do.

The do-file master_psid.do references the do-files that are run on restricted-access PSID data. See the programs/psid subfolder. The analysis uses information on individuals' county of residence, which is not available in public-use PSID data. The specific restricted-access files necessary are the "county identifiers file" and the "county where born/grew up." Researchers can request access to these data as described on the PSID website: https://simba.isr.umich.edu/restricted/ProcessReq.aspx. To run master_psid.do, one needs to import temp/cr_combine_psid_april2019.dta into the secure computing environment and put it in the data subfolder in the PSID space. This data file is created in main/cr_combine.do (called in master_main.do). One also needs to manually add the outreg2 ado file to the PSID space.

The do-file master_rdc.do references the do-files that are run on restricted-access Census data. See the programs/rdc subfolder. Before running these do-files, one needs to run the SAS programs convert_allvars_2000_v2.sas, convert_allvars_acs_v2.sas, and convert_cnum_2012.sas, which convert datasets into Stata format. Be sure to modify the directories in the SAS programs as necessary (the general structure is given in master_rdc.do).

To run master_rdc.do, one also needs to import the following files into the RDC:
- convert_cbp_rdc_v3.dta
  - Put this in the data folder in the RDC.

- This file is created in programs/main/cr_bartik.do and is available in the temp folder in the replication package.
- pov_thresholds_1980_2013.dta
  - Put this in the datap folder in the RDC.
  - This file is available in the data folder of the replication package and described below.
- cr_cz_xwalk.dta
  - Put this in the data folder in the RDC.
  - This file is created by programs/main/cr_cz_xwalk.do and is available in the temp folder in the replication package.
- cr_statepop.dta
  - Put this in the data folder in the RDC.
  - This file is created by programs/main/cr_statepop.do and is available in the temp folder in the replication package.
- us2010trf.dta
  - Put this in the datap folder in the RDC.
  - This file is created by programs/main/cr_us2010trf.do and is available in the temp folder in the replication package.
- 07736-0001-Data.dta
  - Put this in the data folder in the RDC.
  - This file is the 1947-1977 County Data Book, available from ICPSR (see below).
- 08256-0001-Data.dta
  - Put this in the data folder in the RDC.
  - This file is the 1983 County Data Book, available from ICPSR (see below).
- 36603-0001-Data.dta
  - Put this in the data folder in the RDC.
  - This file is the county-level natality and mortality data available ICPSR (see below).
- cr_pop.dta
  - Put this in the data folder in the RDC.
  - This file is created by programs/main/cr_pop.do and is available in the temp folder in the replication package.
- cr_bea.dta
  - Put this in the data folder in the RDC.
  - This file is created by programs/main/cr_bea.do and is available in the temp folder in the replication package.
- cr_coudb.dta
  - Put this in the data folder in the RDC.
  - This file is created by programs/main/cr_coudb.do and is available in the temp folder in the replication package.
- cr_nhgis_demo.dta
  - Put this in the data folder in the RDC.
  - This file is created by programs/main/cr_nhgis_demo.do and is available in the temp folder in the replication package.
- cr_cpi.dta
  - Put this in the data folder in the RDC.

- o This file is created by programs/main/cr_cpi.do and is available in the temp folder in the replication package.
- batistichbond_cz_ipw_variables.dta
  - o Put this in the data folder in the RDC.
  - o This file comes from Mary Kate Batistich (see Batistich and Bond 2019) and is available in the data folder in the replication package.

Within the RDC, make sure that num_match_1900.dta is in the datam folder. This is the output of the FIPS matching algorithm described below.

Disclosing results from the restricted-access Census and PSID data requires additional formatting beyond what is done by these do-files. For purposes of replication, I have included the output from the Census and PSID analysis in the rdcresults and psidresults folders.

The only code not provided in the replication package is the code which processes the NUMIDENT file to identify individuals' county of birth in the Census and ACS (see Appendix A of the paper). This code contains information that cannot be publicly disclosed. However, individuals can obtain the code in the secure FSRDC computing environment, as long as they have requested access to the same data sets (the 2000 Census, 2001-2013 ACS, and NUMIDENT should be sufficient). When requesting the code, researchers should reference the following technical memo, which is available in the FSRDC computing environment and provides additional details:

Evan J. Taylor, Bryan A. Stuart, Martha J. Bailey, 2016. "Summary of Procedure to Match NUMIDENT Place of Birth County to GNIS Places," CES Technical Notes Series 16-01, Center for Economic Studies, U.S. Census Bureau.

The code was developed under Census Project #1284 (home institution: University of Michigan). Note that the code uses a data file (referred to as reed_hand.dta) which contains hand-checked matches from Isen, Rossin-Slater, and Walker (2017). I am unable to provide access to these data file, but researchers may request access to it by contacting the authors. In practice, hand-checked matches account for a small share of matches (see Appendix Table A.1), so this additional data file is of modest importance outside of pure replication.

Finally, the MATLAB file lognormal.m estimate the GMM model described in Appendix C. See the programs/main subfolder. Running this file is not necessary for replication, as its output is already incorporated into main/cr_cbp_emp.do.

This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1284 (CBDRB-FY16-P1284-R5220, CBDRB-FY20-P1284-R8650, CBDRB-FY20-P1284-R8657).

For any further questions, contact Bryan Stuart (bryanastuart@gmail.com).

Data

1. CPI for All Urban Consumers (CPI-U), All items in U.S. city average, Not seasonally adjusted, 1960-2014 (CUUR0000SA0)

   These data can be downloaded from the U.S. Bureau of Labor Statistics (https://www.bls.gov/cpi/data.htm). The series ID is CUUR0000SA0.

   The included file cpi_19602015.xlsx contains only annual averages.

2. Commuting Zones and Labor Market Areas

   These data can be downloaded from the U.S. Department of Agriculture Economic Research Service (https://catalog.data.gov/dataset/commuting-zones-and-labor-market-areas).

   cz00_eqv_v1.xls

3. [ICPSR] County Business Patterns, 1976, 1977, 1979, 1982, 1992

   1976 (county files)

   United States. Bureau of the Census. County Business Patterns, 1970-1976 [United States]: U.S. Summary, State, and County Data. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-02-05. https://doi.org/10.3886/ICPSR24722.v1

   CBP76TAB2UDIV[1-9].txt

   1977 (county and nationwide files)

   United States. Bureau of the Census. County Business Patterns, 1977 [United States]: U.S. Summary, State, and County Data. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-01-12. https://doi.org/10.3886/ICPSR08464.v1

   08464-00[11-19]-Data.txt

   1979 (county files)

   United States. Bureau of the Census. County Business Patterns, 1979 [United States]: U.S. Summary, State, and County Data. [distributor], 2006-01-12. https://doi.org/10.3886/ICPSR08442.v1

   08442-00[04-12]-Data.txt

1982 (county files)

United States. Bureau of the Census. County Business Patterns, 1982 [United States]: U.S. Summary, State, and County Data. [distributor], 2006-01-18. https://doi.org/10.3886/ICPSR08360.v1

08360-00[03-11]-Data.txt

1992 (county files)

United States. Bureau of the Census. County Business Patterns, 1992 [United States]: U.S. Summary, State, and County Data. [distributor], 2006-01-18. https://doi.org/10.3886/ICPSR06488.v1

06488-00[03-11]-Data.txt

4. County Population Estimates, U.S. Census Bureau

   Data are downloaded from https://www.census.gov/programs-surveys/popest/data/data-sets.html

   PopEst1970.xlsx
   PopEst1980.xlsx
   STCH_icen[1990-1999].txt
   CO-EST00INT-ALLDATA-[1-56].csv

5. State Population Estimates, U.S. Census Bureau

   Data are downloaded from https://www2.census.gov/programs-surveys/popest/tables/

   e7080sta.txt
   stiag78[0-9].txt
   st-99-[12-19].txt

6. Bureau of Economic Analysis Regional Economic Accounts

   These data downloaded in March 2015 from https://apps.bea.gov/regional/downloadzip.cfm. The BEA occasionally updates the methodology used to generate these data, and all files are updated when the methodology is updated, so data downloaded at later dates might differ somewhat.

   CA1: Personal Income Summary: Personal Income, Population, Per Capita Personal Income

   CA1_1969_2013_ALL.csv

CA5: Personal Income by Major Component and Earnings by SIC Industry

CA5_1969_2000_ALL.csv

CA25: Total Full-Time and Part-Time Employment by SIC Industry

CA25_1969_2000_ALL.csv

CA25N: Total Full-Time and Part-Time Employment by NAICS Industry

CA25N_2001_2013_ALL.csv

CA35: Personal Current Transfer Receipts

CA35_1969_2013_ALL.csv

7. Census County Data Books, 1947-1977, 1983, 1994, 2000

1947-1977

United States. Bureau of the Census. County and City Data Book [United States] Consolidated File: County Data, 1947-1977. Inter-university Consortium for Political and Social Research [distributor], 2012-09-18. https://doi.org/10.3886/ICPSR07736.v2

07736-0001-Data.dta

1983

United States. Bureau of the Census. County and City Data Book [United States], 1983. Inter-university Consortium for Political and Social Research [distributor], 2008-06-18. https://doi.org/10.3886/ICPSR08256.v1

Note: Running the included do-file programs/main/08256-0001-Setup.do is necessary.

08256-0001-Data.dta

1994

Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. https://doi.org/10.3886/ICPSR02896.v3

DS80 1994 County Data Book (County and State)

02896-0080-Data.dta

2000

Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. https://doi.org/10.3886/ICPSR02896.v3

DS81 2000 County Data Book (County and State)

02896-0081-Data.dta

8. Annual Survey of Governments

    http://data.nber.org/asg/ASG_release/County_City/Data/

    local_asg.dta

9. Data from NHGIS

    nhgis0035_ds84_1950_county.csv
    nhgis0036_ds83_1950_county.csv
    nhgis0037_ds84_1950_county.csv

    nhgis0036_ds89_1960_county.csv

    nhgis0033_ds94_1970_county.csv
    nhgis0033_ds95_1970_county.csv
    nhgis0033_ds98_1970_county.csv
    nhgis0036_ds94_1970_county.csv
    nhgis0037_ds98_1970_county.csv

    nhgis0030_ds104_1980_county.csv
    nhgis0030_ds107_1980_county.csv
    nhgis0036_ds104_1980_county.csv
    nhgis0037_ds107_1980_county.csv

    nhgis0030_ds120_1990_county.csv
    nhgis0030_ds123_1990_county.csv
    nhgis0035_ds120_1990_county.csv
    nhgis0036_ds120_1990_county.csv
    nhgis0037_ds123_1990_county.csv

    nhgis0035_ds146_2000_county.csv
    nhgis0036_ds146_2000_county.csv
    nhgis0037_ds151_2000_county.csv

nhgis0039_ds151_2000_county.csv

nhgis0060_ds176_20105_2010_county.csv
nhgis0061_ds206_20145_2014_county.csv

In addition, the following shapefiles come from NHGIS: US_county_2000.dbf, US_county_2000.shp, US_state_2000.dbf, US_state_2000.shp.

Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. 2021. http://doi.org/10.18128/D050.V16.0

10. Bureau of Labor Statistics Local Area Unemployment Statistics

1976-1989

These data were obtained via email from Brian Hannon of the BLS.

BLS_all_archived_counties7689aa (UNOFFICIAL).xlsx

I have made these data available at:

United States Department of Labor. Bureau of Labor Statistics. Bureau of Labor Statistics Local Area Unemployment Statistics, 1976-1989. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-07-29. https://doi.org/10.3886/E146501V1

1990-2013

These data were downloaded in July 2016 from the BLS https://download.bls.gov/pub/time.series/la/

la.data.64.County.txt

11. NCHS Vital Statistics Natality Birth Data, 1970-1986

National Center for Health Statistics. Data File Documentations, Natality, 1970-1986 (machine readable data file and documentation), National Center for Health Statistics, Hyattsville, Maryland. Accessed via http://data.nber.org/data/vital-statistics-natality-data.html

natl[1970-1986].dta

12. Data on FIPS state codes, NCHS state codes, Haines state codes, and state names

state_codes.dta

13. County-level natality and mortality data.

    Bailey, Martha, Clay, Karen, Fishback, Price, Haines, Michael R., Kantor, Shawn, Severnini, Edson, and Wentz, Anna. U.S. County-Level Natality and Mortality Data, 1915-2007. Inter-university Consortium for Political and Social Research [distributor], 2018-05-02. https://doi.org/10.3886/ICPSR36603.v2

    36603-0001-Data.dta

14. Chambers. M. M. Appropriations of State Tax Funds for Operating Expenses of Higher Education. 1972-1973. PDF downloaded from https://education.illinoisstate.edu/grapevine/historical/ and entered into a spreadsheet.

    Grapevine_7273.xlsx

15. Poverty thresholds by size of family and number of children, 1980-2013.

    These data were downloaded from https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html, and have been converted into Stata format.

    pov_thresholds_1980_2013.dta

16. IPUMS data from 2000 Census and 2001-2013 ACS

    Samples: 2000 Census (5%), 2001-2013 ACS (single-year), all individuals

    Variables: year, datanum, serial, hhwt, hhtype, statefip, gq, gqtype, gqtyped, pernum, perwt, famsize, famunit, momloc, poploc, sploc, subfam, sftype, sfrelate, relate, related, sex, age, marst, birthyr, race, raced, bpl, bpld, educ, educd, empstat, empstatd, labforce, classwkr, classwkrd, wkswork1, wkswork2, uhrswork, inctot, ftotinc, incwage, incbus00, incss, incwelfr, incinvst, incretir, incsupp, incother, incearn, poverty, migrate5, migrate5d, migrate1, migrate1d, qage, qmarst, qrelate, qsex, qbpl, qrace, qeduc, qclasswk, qempstat, qocc, quhrswor, qwkswork, qincearn, qincbus, qincinvs, qincothe, qincreti, qincss, qincsupp, qinctot, qftotinc, qincwage, qincwelf, qmigrat1, qmigrat5

    This is named ipums20002013.dta.

    Steven Ruggles, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler and Matthew Sobek. IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D010.V11.0

17. NLSY79 data. These data were downloaded in September 2016 via the NLS Investigator, available here: https://www.nlsinfo.org/content/getting-started/accessing-data.

The data are included as nls_11sept2016.csv. In addition, there is a custom weight file (obtained from https://www.nlsinfo.org/weights/nlsy79) which is included as customweight_nlsy79_57d600e718f92.dat.

18. Japanese import competition data

These data were obtained via email from Mary Kate Batistich. They are described in Batistich and Bond (2019)

batistichbond_cz_ipw_variables.dta

19. 2010 Tract Relationship File, U.S. Census Bureau

Data downloaded from https://www2.census.gov/geo/docs/maps-data/data/rel/trf_txt/.

us2010trf.txt