READ ME

This archive contains the program files that need to be run to reproduce the results contained in the paper "Health Shocks, Health Insurance, Human Capital, and the Dynamics of Earnings and Health" by Capatina and Keane.

There are two main folders in the replication package:

1. Data: Contains the Stata DO files used to obtain all statistics from publicly available data sources: MEPS, CPS, and HRS, and process and analyse the simulated data. This READ ME file explains its contents.

2. Model: Contains the FORTRAN files used to calibrate and simulate the model. – See "READ ME – MODEL"

Data File Structure:

The data folder needs to be organized as follows. The replication package already contains the correct structure, with 1 exception: The folder "Simulated data" needs to be placed under Data/. (This exception exists due to uploading restrictions, so the folder "Simulated data" needed to be uploaded separately, in the main directory.)

Data/

- MEPS/
    - Do Files/
    - MEPS Data Files/
- CPS/
    - Do Files/
    - CPS Data Files/
- HRS/
    - Do Files/
    - HRS Data Files/
- Simulated data/
    - Calibration/
    - Calibration Minorities/
    - Experiments/
    - Experiments Minorities/
- Parameters for Model/ - Folder where we export parameters that are externally estimated, directly from the data. These are to be copied and pasted into the corresponding folders in "Model/Fortran code/.../External_Parameters/." Sub-

folders “Blacks” and “Hispanics” contain the estimated parameters that differ for these groups.
- Tables and Figures/ - Folder where we export tables and figures.
- Configure stata.do – Do-file for configuring STATA 18.

Public Data Sources

The paper uses data that is publicly available from the following sources. There is a sub-folder for each one of these under “Data.”

1. Medical Expenditure Panel Survey (MEPS) from the Agency for Healthcare Research and Quality (AHRQ).

   Specifically, we use two of the Household Component Full-Year files:

   1. Full-Year Consolidated Data files
   2. Medical Conditions files

   We downloaded the data for each year from 2000-2014 and combined all years into one file.

2. Current Population Survey (CPS)

   Citation: *Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry. IPUMS CPS: Version 12.0 [dataset]. Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0*

3. Health and Retirement Study (HRS)

   Citation: *"Health and Retirement Study, randhrs1992_2016v1 public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (2018)."*

MEPS/

The folder MEPS contains the following sub-folders:

1. Do Files
   - Contains all do files, including “DO ALL” that executes all sub-files.
2. MEPS Data Files

i. Raw MEPS Data files/ –yearly consolidated data files downloaded from the MEPS website are to be saved here
ii. Intermediate Data Files/ – this folder contains exported STATA data files, including intermediate data files, and the two main data files that are used for analysis:
    1. “MEPS_merged_Health_Whites.dta" and
    2. “MEPS_merged_Health_all_races.dta"

The file “DO ALL.do” runs the following sub-files:

1. “0.1 Combine Consolidated Data Files for all years.do”
    - takes the raw data files downloaded from the MEPS website, makes changes to each year so that variables are consistently named over the years, and then combines the Consolidated data files for all years into one data file “Consolidated_data_merged.dta.”

2. "0.2 Clean MEPS Consolidated Data Files.do"
    - Reshapes the data so that we only have one line per ID.
    - Keeps the variables we need and recodes some of them to make them easier to use.
    - Creates some new variables such as employment/labor force participation/insurance that are at the annual level and takes into account the length of the round.
    - We identify the spouses and we create new variables for each individual with the characteristics of their spouses. These variables start with “spouse_”.
    - Saves the file “Consolidated_data_reshaped.dta” at the end.
3. "1.1 MEPS Merge with data on H - All Races.do"
    - Merges the above file with the file that contains the health shocks, H and R for each ID
    - Constructs income quintiles.
    - Saves the file “MEPS_merged_Health_all_races.dta" which is the file used for much of the analysis on racial minorities.
4. "1.2 MEPS Merge with data on H.do"
    - Does the same as the file above, but only keeping whites.
    - Exports the income thresholds that determine the income quintiles – to be used in the code (parameters_inc_ts.txt).
    - Saves the file “MEPS_merged_Health_Whites.dta" which is the file used for most of the analysis on whites.
5. "2. Marital Status and Spouse Stats.do"

- Estimates external parameters used in the model related to spouses:
    i. Marriage0.txt – initial distribution of married individuals
    ii. Marriage1.txt – probability of getting married
    iii. Marriage2.txt - probability of separation
    iv. Prob_spouse_work.txt - probability of spouse working
    v. Income_spouse1.txt – spousal income
    vi. OOP_spouse.txt – spousal OOP medical costs at working ages
    vii. OOP_spouse_old.txt - spousal OOP medical costs at ages 65+
- Produces tables 51 and 89-91 for the Appendix.

6. "3. Medical expenditures.do"
    - Generates the medical expenditures used in model. Constructs Appendix Tables 26-32. Exports the estimates to files to be read into model:
        i. ME_ins_1.txt and ME_ins_2.txt for those insured, with catastrophic and non-catastrophic costs, working ages.
        ii. ME_no_ins_1.txt and ME_no_ins_2.txt for those uninsured, working ages.
        iii. ME_ret_1.txt and ME_ret_2.txt for age 65+.
    - Generates the statistics for Table 35 in appendix.
    - Saves numbers that will be used later for Appendix Fig 12 – concentration of OOP.
7. "4. Medical Charges and Treatment.do"
    - Tables 37 – 40 with Treatment and Payment rates.
    - Saves data for Figures 10 and 11 on Medical charges by age and ESHI.
8. "5. H and R.do"
    - Figure 1 – H, R, and health shocks over time
    - Tables 60, 61 – Fraction with ESHI by H.
    - Saves data for Figures 5 and 6 – H transitions to compare with model data.
    - Table 47 – Income gradient by H
    - Saves data for Table 1 in paper: H Ordered logit
    - Table 22 – Ordered logit, latent class
    - Tables 23 and 24 – H transitions by insurance
    - Saves data for Figure 7 (H by age and education), and Figures 14, 15, 16 (H by age and ESHI, by education)
    - Saves data for Figure 4 (R by age)
    - Tables 15, 16, 17 – R regressions
    - Produces Figure 3
    - Exports R transition probabilities to file "parameters_health_R.txt"
    - Distribution of H and R in data at age 25
9. "6. Health Shocks.do"

- Produces Tables 9-12 in Appendix with health shock regressions and frequencies.
- Figure 2 with health shocks by age
- Exports shock probabilities to “parameters_health_shocks.txt"

10. “7. Labor Statistics.do”
    - Produces statistics for Appendix D on wages, earnings, employment by health: Tables 57, 58, 63, 64, 69, 81 and saves data for Fig 24.
11. “8. Sick Days.do”
    - Exports the sick days to files:
        i. sick_days.txt - sick days for workers (also presented in Table 33)
        ii. sick_days_nw.txt – sick days for non-workers (also presented in Table 34)
12. “9. Health Stats Minorities.do”
    - This generates all the statistics for Appendix K on Racial disparities: Tables 112 – 117: racial distribution, health regressions, medical charges regressions.
    - Table 127 (left panel) (Appendix L)
    - Saves statistics from data for Appendix L, Figure 33, 35, 36, 37, 38 (H and R by various sub-groups for minorities)
    - Tables 128 and 129 with treatment and payment rates (targets in calibration)
13. “10. Marital Status and Spouse Stats Minorities.do”
    - These are stored in the Blacks and Hispanics sub-folders in “Parameters for Model”.
        i. Marriage0.txt , Marriage1.txt, Marriage2.txt - initial distribution of married, probabilities of getting married and separated
        ii. Prob_spouse_work.txt - probabilities the spouses work
        iii. Income_spouse1.txt - Spousal income
        iv. OOP_spouse.txt – Medical costs for spouses
    - Produces Appendix Table 120

CPS/

Contains the following sub-folders:

1. Do Files/
    - Stats from CPS.do
    - Stats from CPS Races.do
2. CPS Data Files/

i. Raw CPS Data files/ – downloaded CPS raw data extract is to be saved here
ii. Intermediate Data Files/ – intermediate files saved here.

"Stats from CPS.do"

- Imports data extracted from http://cps.ipums.org/ and performs basic data cleaning. Saves file CPS.dta for future use to obtain statistics.
- Family size
  - Figure 26 – family size by age
  - parameters_fam_size.txt – exports family size parameters for model
- Gov transfers
  - Tables 70, 71, 72 and Figure 25
- Wage statistics
  - Table 65 – PT/FT ratio
  - Table 68 – Wage distribution
- Earnings inequality
  - Data for Figure 28
- Employment
  - Table 55 – employment at age 25

"Stats from CPS Races.do"

- Imports data extracted from http://cps.ipums.org/ and performs basic data cleaning. Saves file CPS_minorities.dta for future use to obtain statistics.
- Table 118 – citizenship status
- Figures 29 and 30 – fraction married
- Data for Figure 31 – Employment by age
- Data for Figure 32 – Government transfers by age
- Table 123 – Wage distribution

HRS

Contains the following sub-folders:

1. Do Files/
   - Do files 1-5
   - DO ALL.do
2. HRS Data Files/

i. Raw HRS Data files/ - Downloaded “randhrs1992_2016v1.dta” and file with help variables to be saved here.
ii. Intermediate Data Files/ – intermediate files saved here.

“DO ALL.do” – executes all of the following 5 do files for all statistics from HRS:

1. 0. Preparing HRS data.do
   - Uses 2 raw data files: “randhrs1992_2016v1.dta” and “Help_variables_HRS.dta”
   - basic data cleaning and saves file for analysis “HRS_Analysis.dta”
2. “1. HRS Survival.do”
   - Tables 83 and 119 – survival logit regression
   - Exports survival probabilities to “parameters_survival.txt" for all 3 racial groups.
3. “2. HRS Marital Trans.do”
   - Figure 27 – married to widower transitions
   - Exports parameters on marital transitions for model: “Marriage3.txt” for all 3 racial groups.
4. “4. HRS SS and OOP.do”
   - Tables 76, 77, and 79 – SS benefits for retirees, and work experience at retirement.
   - Table 125 – retirement income for minorities.
   - Exports “SS_inc.txt" and “Income_spouse_ret.txt" with retirement income (own and spouse) – does this for all racial groups.
   - Table 92 bottom panel – spouse’s OOP when retired
5. “5. HRS Assets.do”
   - Table 48 – correlation between health and wealth
   - Table 53 – Assets statistics

Simulated data

The file structure is the following:

Simulated data/

- Calibration/
  - Benchmark Data/
    - Data files/
    - Dictionaries/
  - Do Files/
    - “Benchmark Calibration.do”

        - Do files numbered 1-8
- Calibration Minorities/
    - Benchmark Data Minorities/
        - Data files/
        - Dictionaries/
    - Do Files/
        - “Benchmark Calibration.do”
        - Do files numbered 1-3
- Experiments/
    - Experiments Data/
        - Counterfactuals Stata Files/
        - Dictionaries/
        - Experiments Stata Files/
        - Temp/
    - Do Files/
        - “Run All.do”
        - Do files numbered 1-9
- Experiments Minorities/
    - Experiments Data/
        - Counterfactuals Stata Files/
        - Dictionaries/
        - Experiments Stata Files/
        - Temp/
    - Do Files/
        - “Import Experiments Data.do”
        - “Results Minorities.do”
-

Calibration

“Calibration/Do Files/Benchmark Calibration.do” runs all of the following do files that produce tables and figures for the Appendix and paper. We specify when tables and figures are found in the manuscript, otherwise they are in the Appendix, mostly in the calibration sections which compare data and model.

1. “1. Import Benchmark Data.do”
2. “2. Plots for Calibration.do”
    - Figures 4, 5, 6, 7, 9, 11, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 28
3. “3. Calibrate Emp.do”
    - Tables 46, 52, 57 and 58 right panels, 78
4. “4. Gov Transfers.do”

- Tables 72, 74, 75

5. "5. Calibrate H R Shocks.do"
   - ***Table 1 in paper***
   - Tables 12, 17, 20, 22, 23, 25, 47
6. "6. Assets.do"
   - Tables 48 and 54
7. "7. Calibrate Pay Treat.do"
   - Tables 37, 38, 39, 40, 42, 43, 44
8. "8. Wages.do"
   - Tables 63, 64, 65, 67, 68

Important notes: The Benchmark data in Stata format is in the folder "Calibration\Benchmark Data\Data files". Unless we re-run the model and import the new data into Stata using the Dictionary files, we do not need to execute this. Thus, it is commented out in the do file "Benchmark Calibration.do." Uncomment it and set the correct path in Dictionary files to import newly simulated data if re-running the model.

Calibration Minorities

"Calibration Minorities/Do Files/Benchmark Calibration.do" runs all of the following do files:

1. "1. Import Benchmark Data.do"
2. "2. Plots for Calibration.do"
   - Figures 31 – 38 (all figures for calibration section)
3. "3. Tables for Calibration.do"
   - Tables 123, 124, 127, 128, 129 (all tables for calibration in Appendix L)

Important notes: Again, the Benchmark data in Stata format is in the folder "Calibration Minorities\Benchmark Data Minorities\Data files". Unless we re-run the model and import the new data into Stata using the Dictionary files, we do not need to execute this. Thus, it is commented out in the do file "Benchmark Calibration.do." Uncomment it and set the correct path in Dictionary files to import newly simulated data if re-running the model. (The Benchmark for minorities corresponds to the setting jBenchmark_exp=0 in the simulation code found in Model/Fortran code/Experiments/Experiments_B/H.)

Experiments

"Run All.do" runs all of the following do files:

1. "1. Import Benchmark Data.do"
2. "2. Severe Health Shocks Table and Figures.do"

    - ***Figures 1 and 2 in the paper*** and **Table 2 in paper**
3. "3. Effects of Shocks Experiments.do"
    - ***Table 3 in paper***
4. "4. Effects of Shocks Decomposed.do"
    - ***Tables 4, 5 and 6 in paper***
5. "5. Health spending.do"
    - ***Table 7 in paper***
    - Table 130 in Appendix
6. "6. Insurance Experiment.do"
    - ***Table 8 in paper***
    - Table 132 in Appendix
7. "7. Mechanisms.do"
    - ***Table 9 in paper***
8. "8. Extra Tables.do"
    - All tables in Appendix J (Extra Results) for whites: Tables 106-110
    - Table 14 Appendix
    - Figure 12 Appendix
9. "9. Sensitivity Analysis.do"
    - All in Appendix G (Sensitivity Analysis): Tables 93-95

Important notes: The experiments data sets in Stata format are in the folder "Experiments\Experiments Stata Files\". Unless we re-run the model and import the new data into Stata using the Dictionary files, we do not need to execute the first do file: "Import Experiments Data.do." Thus, it is commented out in the do file "Run All.do." Uncomment it and set the correct paths in the dictionary files to import newly simulated data if re-running the model.

Experiments Minorities

"Results Minorities.do" produces all the results for racial minorities:

- Imports data using "Import Experiments Data.do"
- ***Table 10 in paper***
- Tables 131, 133 Appendix

Important notes: As with the previous cases, unless we re-run the model and re-import the new data into Stata using the Dictionary files, we do not need to execute the first do file: "Import Experiments Data.do."