Replication Materials, **Data Component**, for

WHY DO COUPLES AND SINGLES SAVE DURING RETIREMENT? HOUSEHOLD HETEROGENEITY AND ITS AGGREGATE IMPLICATIONS

by Mariacristina De Nardi, Eric French, John Bailey Jones and Rory McGee

JPE MS 2021-11-29

March 2024

Overview: This file describes how to run the analyses that produce the data inputs for our model and MSM estimation procedure. All code is in Stata format, except for the medical spending persistence code, which is in GAUSS format.

Details on how we convert raw HRS data downloads into the datasets used in these analyses can be found in the file **readme_rawHRS.pdf**.

1. **The data inputs are contained in two directories: 1) /Programs and Data/wealthcouples, and 2) /Programs and Data/wealthcouples/impute**
   a. The main HRS data file is "dataprep0_impute.dta" in /wealthcouples. This file contains imputed values from the MCBS. This file is pre-processed (i.e., there are many files used to construct this file), and the computer code used to create this dataset are described in this readme file and in readme_rawHRS.pdf and readme_rawMCBS.pdf. If you wish to modify what we have done, we strongly recommend that you begin at this step.
   b. The HRS exit interview data on transfers are contained in "exitassets_amothers.dta" in /wealthcouples. This dataset is pre-processed (i.e., there are many files used to construct this file), and the computer code used to create this dataset are described in this readme file and in readme_rawHRS.pdf. These data are not used for the main results, and are only used as inputs for the file JPErevision_amothers.do.
   c. The MCBS data used to impute Medicaid spending resides in /wealthcouples/impute. There are two files:
      - "impute_coeffs2yr.dta" contains regression coefficients.
      - "impute_donors2yr.dta" contains donor residuals.

      These two files are used by impute_medicaid.do (see point 2.b below) to create our main HRS data file dataprep0_impute.dta. We are unable to provide the underlying MCBS data. We have, however, provided the programs used to generate both the imputation files which are in the directory: /wealthcouples/impute/MCBS_imputation_code. See point 3 below for more on this.

2. **Descriptions of the .do files for analysis**

   a. **dataprep_couples0.do**:
      - Uses /merge/insuranceimpute.dta

- Drops households not in the AHEAD cohort
- Defines manmarstat, womanmarstat, mandead, womandead
- Defines dumpthis variables for dropping people with marital transitions and inconsistencies that aren't permitted (i.e. hh who split up, people who are "separated" or "partnered", people who died and then came back to life, hh where one spouse is dead and the other claims to be divorced, hh where spouses claim different marital status in different years, etc.)
- Defines firsthh
- Saves file as dataprep0.dta

b. **impute_medicaid.do**
- Uses "dataprep0.dta"
- Reads in "impute\impute_donors2yr.dta" and "\impute_coeffs2yr.dta"
- Imputes Medicaid payments using MCBS data (see point 3 below)
- Saves file as "dataprep0_impute.dta"
  - We provide this file as we are unable to provide raw MCBS data
- This program is called by "dataprep_couples1.do"

c. **dataprep_couples1.do**
- Uses "dataprep0.dta"
- Executes drops defined in dataprep0.do, and drops workers
- Fixes some inconsistencies in marriage, death variables
- Fixes possible inconsistencies between spouses in household level variables
- Defines imputed spousal age (and hhage) and drops couples with too large an age gap
- Moves some exit information to correct wave
- Generates (wo)mandied, (wo)manheal, (wo)mannursing
- Saves "dataprep1x.dta"
- Drops wives to have household-level observations
- Defines hhstatus (and single_m, single_w, couple)
- Defines permanent income
- Creates shifters for incprof
- Saves "dataprep0x.dta"
- Saves "dataprep1.dta"

d. **dataprep_couples2.do**
- Uses "dataprep1.dta"
- Defines (wo)mannursndead and changes (wo)manheald accordingly
- Lags age, age_wife, hhage, couple, manheal, and womanheal for healtrans.do
- Imputes lagged (wo)manheal if missing died==1
- Imputes lagged (wo)manheal if missing before entering NH via nursndead
- Bottom codes hhmedcost at 250

- Saves “Dataprep2_4states.dta”
- We provide a copy of the output .dta file in subfolder “May2022”

e. **healtrans_specB.do**
- Uses “dataprep2_4states.dta”
- Defines dummies and other variables for multivariate logit estimation*:* *mlogit heal age age2 age3 PI PI2 PIage lcouple lcoupleage lheal_bad lhealage_bad lheal_good lhealage_good male maleage malePI male_lcouple lcouplePI, baseoutcome(0*)
- Outputs regression coefficients for Gauss life expectancy code in “hstransh_041713_sing_1mtx_specB”

f. **medexFE_res**
- Uses “dataprep2_4states.dta”
- Drops wave 2 and makes medcost household-level
- Defines dummies/necessary variables for fixed effects regression: *xtreg medcost age1 age2 age3 age4 single_m single_w manbadheal womanbadheal mannursheal womannursheal single_mage single_wage single_mage2 single_wage2 mandied womandied mandiedPI womandiedPI PIage, fe i(HHID)*
- Defines shifters and generates medical expense profiles for couples and singles in each health state, as well as for couples in which one spouse dies at age 80.
- Has switches for medcost in logs or levels (with appropriate variance correction in log case)
- Saves normalized, var(1), medical spending residuals in "\normalizedresid.dta" which is used in "\Medex_AR\momenthealX.do"

g. **incprof.do**
- Uses “dataprep2_4states.dta”
- Uses shifters defined in “dataprep_couples1.do” to generate income profiles for couples and singles separately (at the 20% and 80% PI levels), as well as couples that lose a spouse at age 80
- Outputs regression coefficients for Gauss in “\incprof”

h. **isim96_c.do**:
- Uses Dataprep2_4states.dta.dta
- Recodes firsthh to be compatible with Gauss life expectancy code
- Produces an initial distribution of (age PI firsthh manheal womanheal) either for singles or for everyone depending on whether couples are dropped

i. **wlthmat_adj.do**
- Uses “dataprep2_4states.dta.dta”

- Replace hhmedcost with Rand version and bottomcode it
- Recode mannurisng and womannursing to zero in wave2 (since everyone should have started non-institutionalized
- Save as “wlth.dta”
- Keep a selection of variables and reshape to Stata’s “wide” data format
- Set variables to zero if missing, but flag them
- Drop those with missing hhstatus
- “outsheet” using wlthmat and resave “wlthmat.dta”

j. **JPErevision_amothers.do**
- Uses “dataprep2_4states.dta”
- Uses “exit_amothers.dta”
- Computes the intensive and extensive margins of bequests made when the first spouse dies, as shown in Table 2

k. **assetprof_growth_v6.do**
- uses “dataprep2_4states.dta”
- Computes the asset growth regressions shown in Table 1.
- For this analysis we used a more recent extract of the HRS. Please contact the authors if interested in this exact extract.

The next two files are in the folder \Medex_AR

l. **momentheal.do**
- uses "\normalizedresid.dta"
- sets up data to export to gauss
- exports "\Medex_AR\moments.dta", "\Medex_AR\momentm.out"’ "\Medex_AR\mydat.out" and "\Medex_AR\momentm.dta"

m. **heavcvXl.GAU**
- Uses " Medex_AR\moments.dta", "\Medex_AR\momentm.out"’ "\Medex_AR\mydat.out" and "\Medex_AR\momentm.dta"
- Estimates error component models for medical expenditure residuals
- Exports "\Medex_AR\cmodel.fmt", "\Medex_AR\covest.fmt", "\Medex_AR\c.fmt" and "\Medex_AR\cdata.fmt"

3. We provide exported coefficients and residual values from the MCBS. We document our MCBS data in the companion file **readme_MCBS.pdf**. Please note the MCBS data itself is not provided.

4. Please contact us if something is unclear, so that we can improve the documentation and make it clearer for everyone.