Replication Materials, **MCBS 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

February 2024

**Overview:** This file describes how we obtained and processed the MCBS data.

**Data Access**: The MCBS data we used are not publicly available and thus not provided here. We used the “*cost\_and\_use*” data, which were hosted on the NBER UNIX servers.

**Programs:** Below we describe the programs that use MCBS data to impute Medicaid spending for the HRS. The programs reside in

/Programs and Data/wealthcouples/impute/MCBS\_imputation\_code.

1. **Directory: .../research/clean**
   1. **clean/YYYY.do:**

YYYY is every year from 1996 to 2012

*Inputs: Raw MCBS data files from /disk/agedisk2/mcbs/cost\_and\_use/data/.*

*Ouputs: YYYY.dta*

1. **Directory: .../research/merge**
   1. **Merge.do:**

Main file (merged clean files), with no price deflation

Appends clean/1996/1996.dta-clean/2012/2012.dta

*Output:* /research/merge/merge.dta

List of variables in merge.dta:

* baseid
* weights: population weights
* year
* pce: 2005 base
* male
* age
* dead
* heal: health status, 1 if bad, 0 if good
* married
* nurshome: = 1 if in a nursing home for 60 days or more
* nurshomedays: days spent in a nursing home
* medicaidind: MCBS administrative variable = 1 if individual participates in Medicaid. Note that this variable excludes a number of individuals for whom Medicaid payments is positive.
* workdind: = 1 if individual works. this variable is missing for 1996-98
* income
* totalexp: total expenditures
* oop: costs paid out of pocket
* medicare: reimbursements from Medicare
* medicaid: reimbursements from Medicaid
* privins: reimbursements from private insurance (Medigap, etc…)
* otherexp: reimbursements from other sources
* secondaryinsprem: premia paid to non-Medicare sources
* partapremium: Medicare Part A premium
* partbpremium: Medicare Part B premium
* totpremium: sum of secondaryinsprem, partapremium, partbpremium
* ooppluspremium: sum of totpremium and oop
  1. **gen\_mergecouples.do:**

Uses merge.dta. Deflates all nominal quantities to 2014 dollars. Also computes the Medicaid premium variable and an indicator for Medicaid receipt.

*Inputs: merge/merge.dta*

*Output: mergecouples.dta. Note: in “merge” prices are nominal, and in “mergecouples” they are deflated to 2014 dollars.*

* 1. **Impute\_mcbs.do:**

Creates HRS consistent variables. Estimates regression coefficients for conditional mean matching imputation. Output files contain the coefficient vector from this regression and a separate file with the residual and value of the x\*beta prediction. Imputation coefficients and residuals are estimated for Medicaid payments as well as for the total medical expenses of Medicaid recipients. Aggregates data to a two-year frequency and estimates the same regressions.

*Input: mergecouples.dta*

*Output: impute\_coeffs.dta, impute \_donor.dta impute \_coeffs2yr.dta, impute \_donor2yr.dta, impute\_totalexp\_coeffs.dta, impute\_totalexp\_donor.dta impute\_totalexp\_coeffs2yr.dta, impute\_totalexp\_donor2yr.dta.*

*Note: files ending with “coeffs” contain a vector of coefficients. Files ending with “donor” contain the value of the residuals and the xbeta prediction. Files with “totalexp” contain total expenditures=Medicaid+medicare+OOP and files with “2yr” have biennial data.*

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