Replication Materials, **Raw 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 produce our analysis datasets from the raw HRS data products. All code is in Stata format. We describe the data analysis itself in the file **readme\_data.pdf**. That data analysis uses the output created by the files described in this readme file.

1. **Data Requirements**

The project used data from the core Health and Retirement Study (HRS) and its supplements available at https://hrs.isr.umich.edu/data-products. All of the datafiles are publicly available upon registration on the HRS website.

To run the programs, one needs to download the following HRS and RAND data products:

* HRS Cross-Wave Tracker File
* HRS Core files (1992-2020) + AHEAD Core files (1993 & 1995)
* HRS Core Imputations files (1992-2004) + AHEAD Core Imputations file (1995)
* HRS Exit files (1996-2020) + AHEAD Exit file (1995)
* HRS Post Exit files (1998-2020)
* HRS Exit Imputations files (1998 & 2000) + AHEAD Exit Imputations file (1995)
* RAND HRS Longitudinal File 2018v2

These data products provide the raw files used as inputs in the analysis.

1. **.do files used to produce the analysis dataset**

**Important: There are path dependencies** –run the files in this order.

1. Primary Files (no dependencies; can be done in any order)

* *asset.do*
* *demog.do*
* *employ.do*
* *exitassets.do*
* *exithealutil.do*
* *exitmedcost.do*
* *healstat.do*
* *healutil.do*
* *income.do*
* *insure.do*
* *medcost.do*
* *mortality.do*
* *pension.do*
* *retplans.do*
* *tenure.do*
* *unsettle.do*

1. Initial Merge (only after all primary files are complete; has to be done in this order)

* *merge1.do*
* *merge2.do*
* *merge3.do*

1. RAND files (after Initial Merge; in this order)

* *randreduce.do*
* *huswife.do*
* *randcode.do*

1. Second Merge (after RAND files; in this order)

* *merge4.do*
* *merge5.do*
* *insureanceimpute.do*

1. Final Merge (after Insuranceimpute.do; in this order)

* *merge6.do*
* *merge7.do*

1. Final Preparation of Analysis Data (after merge7)

* *Dataprep\_couples0.do (calls impute\_Medicaid.do)*

1. **Description of Individual .do files**: .do files reside in \merge with folder location controlled by $folder
2. **all\_prep\_JPE.do**:

* this master file runs all of the files below, in proper order
* global $folder for setting paths

1. **asset.do**

* reads in
  + $folder/hrs\wave1\househld
  + $folder/hrs \wave2\w2d
  + $folder/hrs \wave2\h94i\_d
  + $folder/hrs \wave2\h94i\_k
  + $folder/hrs \wave3\h96i\_er
  + $folder/hrs \wave3\h96i\_fh
  + $folder/hrs \wave3\h96i\_jh
  + $folder/hrs \wave4\h98i\_fh
  + $folder/hrs \wave4\h98i\_jh
  + $folder/hrs \wave5\H00CS\_R
  + $folder/hrs \wave5\H00A\_R
  + $folder/hrs \wave5\H00PR\_R
  + $folder/hrs \wave5\h00i\_fh
  + $folder/hrs \wave5\h00i\_jh
  + $folder/hrs \wave6\H02A\_R
  + $folder/hrs \wave6\H02i\_hh
  + $folder/hrs \wave6\h02i\_qh
  + $folder/hrs \wave7\H04i\_hh
  + $folder/hrs \wave7\h04i\_qh
  + $folder/hrs \wave7\H04A\_R
  + $folder/hrs \wave8\H06A\_R
  + $folder/hrs \wave9\H08A\_R
  + $folder/hrs \wave10\H10A\_R
  + $folder/hrs \wave11\H12A\_R
  + $folder/hrs \wave12\H14A\_R
  + $folder/hrs \aheadw1\bhh21
  + $folder/hrs\aheadw1\br21
  + $folder/hrs\aheadw2\A95E\_R
  + $folder/hrs\aheadw2\a95i\_fh
  + $folder/hrs\aheadw2\a95i\_jh
  + $folder/hrs\aheadw2\A95N\_H
* generates $folder/hrs\merge\asset
* Output dataset has asset data
* includes variables: house remort re liquid wheels trust nohouse assets

1. **demog.do**

* reads in
  + $folder/hrs \tracker\trk2010
  + $folder/hrs \wave1\health
  + $folder/hrs\wave2\w2a
  + $folder/hrs\wave3\H96a\_r
  + $folder/hrs\wave3\H96cs\_r
  + $folder/ hrs\wave4\h98cs\_r
  + c:\research \hrs\wave4\h98a\_r
  + c:\research \hrs\wave5\H00CS\_R
  + c:\research \hrs\wave5\H00A\_R
  + c:\research \hrs\wave6\H02A\_R
  + c:\research \hrs\wave6\H02B\_R
  + (Note: From Wave 7 onwards all demog data taken from tracker files instead)
  + c:\research \hrs\aheadw1\br21
  + c:\research \hrs\aheadw2\A95A\_R
  + c:\research \hrs\aheadw2\A95CS\_R
* generates c:\research \hrs\merge\demog
* Output dataset has demographic variables such as age, education, race, and marital status
* includes variables: birthm birthyr black hrscoh age school intm intyr male

1. **employ.do**

* reads in
  + $folder/hrs\aheadw1\BR21
  + $folder/hrs\aheadw2\A95G\_R
  + $folder/hrs\wave12\H14J\_R
  + $folder/hrs\wave11\H12J\_R
  + $folder/hrs\wave10\H10J\_R
  + $folder/hrs\wave9\H08J\_R
  + $folder/hrs\wave8\H06J\_R
  + $folder/hrs\wave7\H04J\_R
  + $folder/hrs\wave6\H02J\_R
  + $folder/hrs\wave5\H00G\_R
  + $folder/hrs\wave4\H98G\_R
  + $folder/hrs\wave3\H96G\_R
  + $folder/hrs\wave2\ w2fa
  + $folder/hrs\wave2\ w2fb
  + $folder/hrs\wave1\employer
* generates $folder/hrs\merge\employ.dta
* This dataset codes up hours, wages, occupation, industry
* includes variables: work same occ ind hours weeks wkrat lfpr pay sal wsal whourly piece other other se prof earn

1. **exitassets.do**

* reads in
  + $folder/hrs\aheadw2\x95i\_nr
  + $folder/hrs\aheadw2\x95N\_R
  + $folder/hrs\wave3\x96i\_nr
  + $folder/hrs\wave3\x96N\_R
  + $folder/hrs\wave4\x98i\_nr
  + $folder/hrs\wave4\X98N\_R
  + $folder/hrs\wave5\h2000xi
  + $folder/hrs\wave5\X00N\_R
  + $folder/hrs\wave6\x02T\_R
  + $folder/hrs\wave7\x04T\_R
  + $folder/hrs\wave7\px04T\_R
  + $folder/hrs\wave8\x06T\_R
  + $folder/hrs\wave8\px06T\_R
  + $folder/hrs\wave9\x08T\_R
  + $folder/hrs\wave9\px08T\_R
  + $folder/hrs\wave10\x10T\_R
  + $folder/hrs\wave10\px10T\_R
  + $folder/hrs\wave11\x12T\_R
  + $folder/hrs\wave12\x14T\_R
* generates $folder/hrs\merge\exitassets.dta
* Output dataset contains asset data for recently dead
* includes variables: estate inspay deathex deathins amspouse amchildren perchildren amcharities amsiblings amrelatives amfriends amothers estatesum

1. **exithealutil.do**

* reads in
  + $folder/hrs\aheadw2\x95CS\_R
  + $folder/hrs\aheadw2\x95PR\_R
  + $folder/hrs\aheadw2\x95E\_R
  + $folder/hrs\wave3\x96CS\_R
  + $folder/hrs\wave3\x96PR\_R
  + $folder/hrs\wave3\x96E\_R
  + $folder/hrs\wave4\x98CS\_R
  + $folder/hrs\wave4\x98PR\_R
  + $folder/hrs\wave4\x98E\_R
  + $folder/hrs\wave5\x00CS\_R
  + $folder/hrs\wave5\x00PR\_R
  + $folder/hrs\wave5\x00E\_R
  + $folder/hrs\wave6\x02A\_R
  + $folder/hrs\wave6\x02PR\_R
  + $folder/hrs\wave6\x02N\_R
  + $folder/hrs\wave7\x04A\_R
  + $folder/hrs\wave7\x04PR\_R
  + $folder/hrs\wave7\x04N\_R
  + $folder/hrs\wave8\x06A\_R
  + $folder/hrs\wave8\x06PR\_R
  + $folder/hrs\wave8\x06N\_R
  + $folder/hrs\wave9\x08A\_R
  + $folder/hrs\wave9\x08PR\_R
  + $folder/hrs\wave9\x08N\_R
  + $folder/hrs\wave10\x10A\_R
  + $folder/hrs\wave10\x10PR\_R
  + $folder/hrs\wave10\x10N\_R
  + $folder/hrs\wave11\x12A\_R
  + $folder/hrs\wave11\x12PR\_R
  + $folder/hrs\wave11\x12N\_R
  + $folder/hrs\wave12\x14A\_R
  + $folder/hrs\wave12\x14PR\_R
  + $folder/hrs\wave12\x14N\_R
* generates $folder/hrs\merge\exithealutil.dta
* Output dataset contains number of nights spent in nursing home in last wave of life
* includes variables: nurhmdate deathdate previntvw numdayscont mmiss night\_sincepreviw numdays

1. **exitmedcost.do**

* reads in
  + $folder/hrs\aheadw2\x95i\_er
  + $folder/hrs\aheadw2\X95R\_R
  + $folder/hrs\aheadw2\x95E\_R
  + $folder/hrs\wave3\x96R\_R
  + $folder/hrs\wave3\x96E\_R
  + $folder/hrs\wave3\x96i\_er
  + $folder/hrs\wave4\x98i\_er
  + $folder/hrs\wave4\x98R\_R
  + $folder/hrs\wave4\x98E\_R
  + $folder/hrs\wave4\x98PR\_R
  + $folder/hrs\wave4\x98CS\_R
  + $folder/hrs\wave5\h2000xi
  + $folder/hrs\wave5\X00R\_R
  + $folder/hrs\wave5\X00E\_R
  + $folder/hrs\wave5\x00PR\_R
  + $folder/hrs\wave5\x00CS\_R
  + $folder/hrs\wave6\x02N\_R
  + $folder/hrs\wave7\x04N\_R
  + $folder/hrs\wave8\x06N\_R
  + $folder/hrs\wave9\x08N\_R
  + $folder/hrs\wave10\x10N\_R
  + $folder/hrs\wave11\x12N\_R
  + $folder/hrs\wave12\x14N\_R
* generates $folder/hrs\merge\exitmedcost.dta
* Output dataset contains medical expense (out of pocket and total) information for recently dead
* includes variables: doop dmedicaid diprem dtc dmedc

1. **healstat.do**

* reads in
  + $folder/hrs\wave12\H14C\_R
  + $folder/hrs\wave11\H12C\_R
  + $folder/hrs\wave10\H10C\_R
  + $folder/hrs\wave9\H08C\_R
  + $folder/hrs\wave8\H06C\_R
  + $folder/hrs\wave7\H04C\_R
  + $folder/hrs\wave6\H02C\_R
  + $folder/hrs\wave5\H00B\_R
  + $folder/hrs\wave4\H98B\_R
  + $folder/hrs\wave3\H96B\_R
  + $folder/hrs\wave2\w2b
  + $folder/hrs\wave1\health
  + $folder/hrs\aheadw1\br21
  + $folder/hrs\aheadw2\A95B\_R
* generates $folder/hrs\merge\healstat.dta
* This dataset codes up some health variables
* includes variables: ghealth healthch pain

1. **healutil.do**

* reads in
  + $folder/hrs\wave1\Health
  + $folder/hrs\wave2\w2b
  + $folder/hrs\wave3\H96e\_r
  + $folder/hrs\wave4\H98E\_R
  + $folder/hrs\wave4\H98CS\_R
  + $folder/hrs\wave5\H00E\_R
  + $folder/hrs\ahead1\bhh21
  + $folder/hrs\ahead1\br21
  + $folder/hrs\ahead2\A95E\_R
* \*no wave 6-10 data
* generates $folder/hrs\merge\healutil
* health care utilization variables
* includes variables: hosp nursing drtimes outsurg dentist nodrug

1. **income.do**

* reads in
  + $folder/hrs\wave1\househld
  + $folder/hrs\wave1\h92i\_n
  + $folder/hrs\wave2\W2a
  + $folder/hrs\wave2\W2n
  + $folder/hrs\wave2\H94i\_n
  + $folder/hrs\wave3\H96a\_r
  + $folder/hrs\wave3\H96i\_jh
  + $folder/hrs\wave4\H98i\_jh
  + $folder/hrs\wave5\H00i\_jh
  + $folder/hrs\wave6\H02A\_R
  + $folder/hrs\wave6\H02i\_qh
  + $folder/hrs\wave7\H04A\_R
  + $folder/hrs\wave7\H04i\_qh
  + $folder/hrs\wave8\H06A\_R
  + $folder/hrs\wave9\H08A\_R
  + $folder/hrs\wave10\H10A\_R
  + $folder/hrs\wave11\H12A\_R
  + $folder/hrs\wave12\H14A\_R
  + c:\research \hrs\aheadw1\br21
  + c:\research \hrs\aheadw1\bhh21
  + c:\research \hrs\aheadw2\A95CS\_R
  + c:\research \hrs\aheadw2\A95J\_H
  + c:\research \hrs\aheadw2\a95i\_jh
* generates c:\research \hrs\merge\income
* traces sources of income
* includes variables: faminc SELFY LABY CAPY Businc Interestinc Capinc Peninc Vainc Socialinc Anninc Unempinc Welfareinc Otherinc SOCYinc Penan pend

1. **insure.do**

* reads in
  + $folder/hrs\wave1\non\_recoded\_househld
  + $folder/hrs\wave2\w2r
  + $folder/hrs\wave3\h96r\_r
  + $folder/hrs\wave4\H98R\_R
  + $folder/hrs\wave5\H00R\_R
  + $folder/hrs\wave6\H02J\_R
  + $folder/hrs\wave6\H02N\_R
  + $folder/hrs\wave7\H04J\_R
  + $folder/hrs\wave7\H04N\_R
  + $folder/hrs\wave8\H06J\_R
  + $folder/hrs\wave8\H06N\_R
  + $folder/hrs\wave9\H08J\_R
  + $folder/hrs\wave9\H08N\_R
  + $folder/hrs\wave10\H10J\_R
  + $folder/hrs\wave10\H010N\_R
  + $folder/hrs\wave11\H12J\_R
  + $folder/hrs\wave11\H012N\_R
  + $folder/hrs\wave12\H14J\_R
  + $folder/hrs\wave12\H014N\_R
  + $folder/hrs\ aheadw1\br21
  + $folder/hrs\ aheadw2\A95R\_R
* generates $folder/hrs\merge\insure
* Output dataset contains insurance information
* includes variables: medcare insgov epins insspa insspb epfra epfrb eprhia eprhib eprfra eprfrb eprspa eprsrb inspriv

1. **medcost.do**

* reads in
  + $folder/hrs\wave1\househld
  + $folder/hrs\wave2\w2r
  + $folder/hrs\wave2\medexp2
  + $folder/hrs\wave2\w2b
  + $folder/hrs\wave3\h96cs\_r
  + $folder/hrs\wave3\h96i\_er
  + $folder/hrs\wave3\H96R\_R
  + $folder/hrs\wave3\H96E\_R
  + $folder/hrs\wave4\h98pr\_r
  + $folder/hrs\wave4\h98i\_er
  + $folder/hrs\wave4\h98r\_r
  + $folder/hrs\wave5\h00pr\_r
  + $folder/hrs\wave5\h00i\_er
  + $folder/hrs\wave5\H00R\_R
  + $folder/hrs\wave6\h02i\_nr
  + $folder/hrs\wave6\H02N\_R
  + $folder/hrs\wave7\h04i\_nr
  + $folder/hrs\wave7\H04N\_R
  + $folder/hrs\wave8\H06N\_R
  + $folder/hrs\wave9\H08N\_R
  + $folder/hrs\wave10\H10N\_R
  + $folder/hrs\wave11\H12N\_R
  + $folder/hrs\wave12\H14N\_R
  + $folder/hrs\ aheadw1\bhh21
  + $folder/hrs\ aheadw1\ br21
  + $folder/hrs\ aheadw2\ A95cs\_r
  + $folder/hrs\ aheadw2\ a95i\_er
  + $folder/hrs\ aheadw2\ A95pr\_r
  + $folder/hrs\ aheadw2\ A95r\_r
* generates $folder/hrs\merge\medcost
* Output dataset contains medical expense (out of pocket and total) information
* includes variables: oop drugc tc medcost iprem ins

1. **mortality.do**

* reads in
  + $folder/hrs \tracker\trk2010
* generates $folder/hrs \merge\mortality.dta
* Dataset contains vital status variable and sub-household IDs
* includes variables: dead adead subxwx

1. **pension.do**

* reads in
  + $folder/hrs\wave1\employer
  + $folder/hrs\wave2\ w2fa
  + $folder/hrs\wave2\ w2fb
  + $folder/hrs\wave2\ w2fc
  + $folder/hrs\wave2\ w2g
  + $folder/hrs\wave2\ w2h
  + $folder/hrs\wave3\ H96G\_R
  + $folder/hrs\wave4\ H98G\_R
  + $folder/hrs\wave5\ H00G\_R
  + $folder/hrs\wave6\ H02J\_R
  + $folder/hrs\wave6\ H02K\_R
  + $folder/hrs\wave6\ H02L\_R
  + $folder/hrs\wave7\ H04J\_R
  + $folder/hrs\wave7\ H04K\_R
  + $folder/hrs\wave7\ H04L\_R
  + $folder/hrs\wave8\ H06J\_R
  + $folder/hrs\wave8\ H06K\_R
  + $folder/hrs\wave8\ H06L\_R
  + $folder/hrs\wave9\ H08J\_R
  + $folder/hrs\wave9\ H08K\_R
  + $folder/hrs\wave9\ H08L\_R
  + $folder/hrs\wave10\ H010J\_R
  + $folder/hrs\wave10\ H010K\_R
  + $folder/hrs\wave10\ H010L\_R
  + $folder/hrs\wave11\ H012J\_R
  + $folder/hrs\wave11\ H012K\_R
  + $folder/hrs\wave11\ H012L\_R
  + $folder/hrs\wave12\ H014J\_R
  + $folder/hrs\wave12\ H014K\_R
  + $folder/hrs\wave12\ H014L\_R
  + $folder/hrs\aheadw2\A95PR\_R
  + $folder/hrs\aheadw2\A95CS\_R
  + $folder/hrs\aheadw2\A95G\_R
* generates d:\hrs\merge\pension.dta
* contains information on pension type, pension wealth, working for the government and/or jobs that don’t pay into Social Security
* includes variables: pentypeDBxcurr pentypeDBxlast pentypeDBxprev pentypeDCxcurr pentypeDCxlast pentypeDCxprev penwDC1curr penwDC1last penwDC1prev penwDB1curr penwDB1last penwDB1prev

1. **retplans.do**

* reads in
  + $folder/hrs\wave1\employer
  + $folder/hrs\wave1\modulee
* generates $folder/hrs\merge\retplans.dta
* Contains retirement information on individuals preferences for work and retirement planning
* includes variables: V5009 V5011 V5012 V5013 V5014 V5015 V5016 V5017 V5018 V5019 V5031 V3319

1. **tenure.do**

* reads in
  + $folder/hrs\wave1\employer
  + $folder/hrs\wave2\w2fa
  + $folder/hrs\ wave3\H96g\_r
  + $folder/hrs\ wave4\H98g\_r
  + $folder/hrs\ wave5\H00g\_r
  + $folder/hrs\ wave6\H02J\_R
  + $folder/hrs\ wave7\H04J\_R
  + $folder/hrs\ wave8\H06J\_R
  + $folder/hrs\ wave9\H08J\_R
  + $folder/hrs\ wave10\H10J\_R
  + $folder/hrs\ wave11\H12J\_R
  + $folder/hrs\ wave11\H14J\_R
  + $folder/hrs\aheadw1\Br21
  + $folder/hrs\aheadw2\A95G\_R
* generates $folder/hrs\merge\tenure.dta
* employment tenure information
* includes variables: tenure job reason

1. **unsettle.do**

* reads in
  + $folder/hrs\wave1\househld
  + $folder/hrs\wave2\w2b
  + $folder/hrs\ wave3\H96cs\_r
  + $folder/hrs\ wave3\H96e\_r
  + $folder/hrs\ wave4\H98cs\_r
  + $folder/hrs\ wave4\H98e\_r
  + $folder/hrs\ wave5\H00CS\_R
  + $folder/hrs\ wave5\H00E\_R
  + $folder/hrs\ wave6\H02A\_R
  + $folder/hrs\ wave6\H02N\_R
  + $folder/hrs\ wave7\H04A\_R
  + $folder/hrs\ wave7\H04N\_R
  + $folder/hrs\ wave8\H06A\_R
  + $folder/hrs\ wave8\H06N\_R
  + $folder/hrs\ wave9\H08A\_R
  + $folder/hrs\ wave9\H08N\_R
  + $folder/hrs\ wave10\H10A\_R
  + $folder/hrs\ wave10\H10N\_R
  + $folder/hrs\ wave11\H12A\_R
  + $folder/hrs\ wave11\H12N\_R
  + $folder/hrs\ wave12\H14A\_R
  + $folder/hrs\ wave12\H14N\_R
  + $folder/hrs\aheadw1\bhh21
  + $folder/hrs\aheadw1\br21
  + $folder/hrs\aheadw2\A95CS\_R
  + $folder/hrs\aheadw2\A95E\_R
* generates $folder/hrs\merge\settle.dta
* includes variables: unset

1. **merge1.do**

* reads in
  + $folder/hrs\merge\employ
  + $folder/hrs\merge\demog
  + $folder/hrs\merge\asset
  + $folder/hrs\merge\income
  + $folder/hrs\merge\insure
  + $folder/hrs\merge\healstat
  + $folder/hrs\merge\mortality
  + $folder/hrs\merge\medcost
  + $folder/hrs\merge\exitmedcost
  + $folder/hrs\merge\exitassets
  + $folder/hrs\merge\exithealutil
  + $folder/hrs\merge\settle
  + $folder/hrs\merge\healutil
  + $folder/hrs\merge\pension
  + $folder/hrs\merge\tenure
  + $folder/hrs\merge\retplans
  + generates $folder/hrs\merge\merge1.dta
* This dataset contains almost all the variables of interest

1. **merge2.do**

* reads in
  + $folder/hrs\merge\merge1
* generates $folder/hrs\merge\merge2.dta

1. **merge3.do**

* reads in:
  + $folder/hrs\merge\merge2
  + $folder/hrs\wave2\w2fc
  + $folder/hrs\wave2\w2fa
  + $folder/hrs\wave2\w2fb
  + $folder/hrs\wave3\H96G\_R
  + $folder/hrs\wave4\H98G\_R
  + $folder/hrs\wave5\H00G\_R
  + $folder/hrs\wave6\H02J\_R
  + $folder/hrs\wave7\H04J\_R
  + $folder/hrs\wave8\H06J\_R
  + $folder/hrs\wave9\H08J\_R
  + $folder/hrs\wave10\H10J\_R
  + $folder/hrs\wave11\H12J\_R
  + $folder/hrs\wave12\H14J\_R
  + $folder/hrs\ aheadw2\A95g\_r
* generates d:\hrs\merge\merge3.dta
* This generates labor force participation variables for the times in between waves (e.g., we get their responses in 1992 and 1994 – what was their lfpr in 1993?)

1. **huswife.do**

* reads in
  + $folder/hrs\merge\merge3.dta
* generates $folder/hrs\merge\huswife1
* merges husbands and wives together

1. **randreduce.do**

* reads in
  + $folder/hrs\rand\rndhrs\_p
* generates $folder/hrs\rand\rndhrs\_c
* cleans RAND data

1. **randcoec.do**

* reads in
  + $folder/hrs\merge\huswife1
  + $folder/hrs\rand\rndhrs\_c
* generates $folder/hrs\merge\randdata1.dta
* produces RAND imputations for relevant variables.
* **A lot of the previously-produced variables are overwritten with their RAND-produced equivalents.**

1. **merge4.do**

* reads in
  + $folder/hrs\merge\randdata1
* generates
  + $folder/hrs\merge\hrs1 to $folder/hrs\merge\hrs10
  + $folder/hrs\merge\hrs102 to $folder/hrs\merge\hrs110
  + $folder/hrs\merge\hrs202 to $folder/hrs\merge\hrs210
* This program generates datasets for each wave.

1. **merge5.do**

* appends $folder/hrs\merge\hrs1 to $folder/hrs\merge\hrs10 (also 102 to 110 and 202-210)
* generates $folder/hrs\merge\merge5.dta
* This also generates some hazard rates for job exit and entry and does some deflating.

1. **insureanceimpute.do**

* reads in
  + $folder/hrs\merge\merge5
* generates $folder/hrs\merge\insuranceimpute.dta
* recodes some insurance variables

1. **merge6.do**

* reads in
  + $folder/hrs\merge\merge5
* generates $folder/hrs\merge\merge6.dta
* recodes some pension and health insurance variables

1. **merge7.do**

* reads in
  + $folder/hrs\merge\merge6
* generates $folder/hrs\merge\merge7.dta
* produces lagged health insurance variables, age dummies

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