

## Using the RAND HRS Longitudinal File and RAND HRS Fat Files

### Sample Programs for HRS Summer Institute Workshop

This document is intended to provide users with some examples of how to both set up and perform some simple descriptive analyses in SAS and Stata, using the RAND HRS data products. The sample programs assume that the data being used are set up in a folder called **C:\RandHRS**, and that any files you create are stored in a separate folder called **C:\MyPaper**. There are four programs, each of which should be run in the specified order:

**Part #1:** Run descriptive analyses using variables from the RAND HRS Longitudinal File (randhrs1992\_2016v1).

- Run tables on variables that describe the interview response status during each wave.
  - INW1 – INW13: These are flags that indicate whether an individual responded at a given wave, where 0 = “Non-response” and 1 = “Response, Alive”.
  - RwiWSTAT: These variables provide (where “w” is the wave, e.g., R10IWSTAT for Wave 10) another way to examine interview status. In addition to telling you whether the individual responded or not, these variables also indicate mortality status.

**Part #2:** Create a file that contains a subset of variables from the RAND HRS Longitudinal File (randhrs1992\_2016v1).

- HACOHORT: This variable identifies the entry cohort subsample (e.g., HRS, Ahead, CODA, WB, EBB, MBB).
- RAGENDER: This variable identifies the gender of the respondent, where 1 = “Male” and 2 = “Female”.
- RARACEM: This variable identifies the race of the respondent, where 1 = “White/Caucasian”, 2 = “Black/African American”, and 3 = “Other”. Note that respondents who are Hispanic can be identified using the RAHISPAN variable.
- R8PENINC – R10PENINC: These variables indicate, for Waves 8 (2006), 9 (2008) and 10 (2010), whether the respondent is currently receiving any pension income. The corresponding variables for the spouse are S8PENINC – S10PENINC.
- R8LBRF – R10LBRF: These variables summarize, for Waves 8 (2006), 9 (2008) and 10 (2010), the labor force status for the respondent at each wave as working full-time (=1), working part-time (=2), unemployed (=3), partly retired (=4), retired (=5), disabled (=6), or not in the labor force (=7). The corresponding variables for the spouse are S8LBRF – S10LBRF.
- R8AGEY\_E – R10AGEY\_E: These variables, for Waves 8 (2006), 9 (2008) and 10 (2010), indicate the age of the respondent (in years) at the final interview date. The corresponding variables for the spouse are S8AGEY\_E – S10AGEY\_E.

**Part #3:** Run descriptive analyses using the dataset created in **Part #2**.

- Generate summary statistics for respondents’ pension income receipt in Waves 8 - 10.
- Produce tables that examine/compare the following:
  - Respondents’ pension income receipt by cohort for Wave 10.
  - Respondents’ pension income receipt to that of the spouse for Wave 9.
  - Respondents’ Wave 9 pension income receipt to that in Wave 10.
  - Respondents’ labor force status by pension income receipt for Waves 9 and 10.
  - Respondents’ pension income receipt and labor force status in Wave 10 by gender.

**Part #4:** Merge the dataset created in **Part #2** with the RAND HRS Fat Files for Waves 9 and 10. The raw variables that are selected from the fat files (described below) include questions that ask respondents whether their employers offer pension plans.

- Merge the dataset created in Part #2 with the fat files for Wave 9 (h08f3a) and Wave 10 (hd10f5e).
  - Important things to note:**
    - For sorting and merging you can use HHIDPN, or HHID & PN together. The results should be the same regardless of which format of the respondent identifier you use.
    - The randhrs1992\_2016v1 dataset and the fat files are all sorted by HHIDPN (HHID & PN). However, Stata requires you to resort the files before merging, and only allows two files to be merged at once.
- Keep HHIDPN, HHID and PN from the fat files, as well as the LJ325 (Wave 9) and MJ325 (Wave 10) variables described below.
- In the final merged dataset, keep only the individuals who responded to Waves 9 and 10.
- Produce some tables that examine/compare the following:
  - Response status flags for Waves 9 and 10.
  - Questions in Waves 9 and 10 about whether employers offer pension plans.
  - Questions in Waves 9 and 10 about whether employers offer pension plans by gender.

```
=====
LJ325      DOES EMPLOYER OFFER ANY PLANS

Section: J      Level: Respondent      Type: Numeric      Width: 1      Decimals: 0
Ref: SecJ.CURRENTPENSIONNEW.J325_

Does your employer offer any (such) retirement plans?
.....

          487          1. YES
         1162          5. NO
           64          8. DK (Don't Know)
            3          9. RF (Refused)
        15501      Blank. INAP (Inapplicable); Partial Interview
```

```
=====
MJ325      DOES EMPLOYER OFFER ANY PLANS

Section: J      Level: Respondent      Type: Numeric      Width: 1      Decimals: 0
Ref: SecJ.CURRENTPENSIONNEW.J325_

Does your employer offer any (such) retirement plans?
.....

          322          1. YES
         1846          5. NO
           80          8. DK (Don't Know)
            1          9. RF (Refused)
        19785      Blank. INAP (Inapplicable); Partial Interview
```

## SAS Code

### Creating a formats catalog

```
/* The sasfmts.sas7bat and randhrs1992_2016v1.sas7bdat files are stored in c:\RandHRS */  
  
libname library "c:\RandHRS";  
  
proc format library=library CNTLIN=library.sasfmts;  
run;
```

### Part #1:

```
proc freq data=library.randhrs1992_2016v1;  
  table inw8-inw10 r8iwstat r9iwstat r10iwstat / missing list;  
run;
```

### Part #2:

```
libname mylib "c:\MyPaper"; /* this is where the output file will be stored */  
  
data mylib.wkshop;  
  set library.randhrs1992_2016v1 (keep=inw8-inw10 hhidpn hacohort ragender raracem  
                                r8iwstat r9iwstat r10iwstat  
                                r8peninc r9peninc r10peninc  
                                r8lbrf r9lbrf r10lbrf  
                                r8agey_e r9agey_e r10agey_e  
                                s8peninc s9peninc s10peninc  
                                s8lbrf s9lbrf s10lbrf  
                                s8agey_e s9agey_e s10agey_e);  
  where inw8=1 or inw9=1 or inw10=1;  
run;  
  
proc contents data=mylib.wkshop;  
run;
```

### Part #3:

```
proc means data=mylib.wkshop;  
  var r8peninc r9peninc r10peninc;  
run;  
  
proc freq data=mylib.wkshop;  
  table r10peninc*hacohort  
        r9peninc*s9peninc r9peninc*r10peninc  
        r9lbrf*r9peninc r10lbrf*r10peninc  
        (r10peninc r10lbrf)*ragender  
        ragender*r10lbrf*r10peninc  
        /missprint;  
run;
```

#### **Part #4:**

```
data mylib.wkplus;
  merge mylib.wkshop
        library.h08f3a (keep=hhidpn hhid pn lj325)
        library.hd10f5e (keep=hhidpn hhid pn mj325);
  by hhidpn;
  if inw9=1 and inw10=1;

run;

proc freq data=mylib.wkplus;
  table lj325 mj325
        (lj325 mj325)*ragender
        /missprint;
run;
```

## Stata Code

### Part #1:

```
set memory 200m
set maxvar 10000
#delimit ;

use inw* hhidpn hacohort ragender raracem
r8iwstat r9iwstat r10iwstat
r8peninc r9peninc r10peninc
r8lbrf r9lbrf r10lbrf
r8agey_e r9agey_e r10agey_e
s8peninc s9peninc s10peninc
s8lbrf s9lbrf s10lbrf
s8agey_e s9agey_e s10agey_e
using "c:\RandHRS\randhrs1992_2016v1" ;

#delimit cr
tab inw8, m
tab inw9, m
tab inw10, m
tab r8iwstat, m
tab r9iwstat, m
tab r10iwstat, m
```

### Part #2:

```
keep if inw8==1 | inw9==1 | inw10==1

save "c:\MyPaper\wkshop", replace
```

### Part #3:

```
use "c:\MyPaper\wkshop"

sum r8peninc r9peninc r10peninc
tab r10peninc hacohort, m
tab r9peninc s9peninc, m
tab r9peninc r10peninc, m
tab r9lbrf r9peninc, col m
tab r10lbrf r10peninc, col m
tab r10peninc ragender, col m
tab r10lbrf ragender, col m
tab r10lbrf r10peninc if ragender==1, col m
tab r10lbrf r10peninc if ragender==2, col m
```

#### **Part #4:**

```
use hhidpn hhid pn lj325 using "c:\RandHRS\h08f3a"
sort hhidpn
save "c:\MyPaper\h08x", replace

use hhidpn hhid pn mj325 using "c:\RandHRS\hd10f5e"
sort hhidpn
save "c:\MyPaper\h10x", replace

clear

use "c:\MyPaper\wkshop"
sort hhidpn

merge 1:1 hhidpn using "c:\MyPaper\h08x", generate(mrg08)
tab mrg08, m
sort hhidpn

merge 1:1 hhidpn using "c:\MyPaper\h10x", generate(mrg10)
tab mrg10, m
sort hhidpn

save "c:\MyPaper\wkplus", replace

keep if inw9==1 & inw10==1

tab mrg08 inw9, m
tab mrg10 inw10, m
tab lj325, m
tab mj325, m
tab lj325 ragender, m
tab mj325 ragender, m
```

## SPSS Code

### Part #1:

```
get file "c:\RandHRS\randhrs1992_2016v1.sav"  
  /keep = inw8 inw9 inw10  
        hhidpn hacohort ragender raracem  
        r8iwstat r9iwstat r10iwstat  
        r8peninc r9peninc r10peninc  
        r8lbrf r9lbrf r10lbrf  
        r8agey_e r9agey_e r10agey_e  
        s8peninc s9peninc s10peninc  
        s8lbrf s9lbrf s10lbrf  
        s8agey_e s9agey_e s10agey_e.  
  
freq  
  /variables = inw8 inw9 inw10 r8iwstat r9iwstat r10iwstat  
  /missing = include.
```

### Part #2:

```
select if inw8 = 1 or inw9 = 1 or inw10 = 1.  
  
save outfile = "c:\MyPaper\wkshop.sav"  
  /map.
```

### Part #3:

```
get file "c:\MyPaper\wkshop.sav".  
  
descriptives  
  /variables = r8peninc r9peninc r10peninc.  
  
crosstabs  
  /tables r10peninc by hacohort  
  /tables r9peninc by s9peninc  
  /tables r9peninc by r10peninc  
  /tables r9lbrf by r9peninc  
  /tables r10lbrf by r10peninc  
  /tables r10peninc by ragender  
  /tables r10lbrf by ragender  
  /missing = include.  
  
select if ragender = 1.  
  
crosstabs  
  /tables r10lbrf by r10peninc  
  /missing = include.  
  
select if ragender = 2.  
  
crosstabs  
  /tables r10lbrf by r10peninc  
  /missing = include.
```

## **Part #4:**

```
get file "c:\MyPaper\wkshop.sav".

sort cases by hhidpn.

save outfile = "c:\MyPaper\wkshops.sav"
  /map.

get file "c:\RandHRS\h08f3a.sav"
  /keep = hhidpn hhid pn lj325.

sort cases by hhidpn.

save outfile = "c:\MyPaper\h08f3as.sav"
  /map.

get file "c:\RandHRS\hd10f5e.sav"
  /keep = hhidpn hhid pn mj325.

sort cases by hhidpn.

save outfile = "c:\MyPaper\hd10f5es.sav"
  /map.

match files
  /file = "c:\MyPaper\wkshops.sav" / in = inrnd
  /file = "c:\MyPaper\h08f3as.sav" / in = in08
  /file = "c:\MyPaper\hd10f5es.sav" / in = in10
  /by hhidpn
  /map.

save outfile = "c:\MyPaper\wkplus.sav".

select if inw9 = 1 and inw10 = 1.

crosstabs
  /tables in08 by inw9
  /tables in10 by inw10
  /missing = include.

freq
  /variables = lj325 mj325
  /missing = include.

crosstabs
  /tables lj325 by ragender
  /tables mj325 by ragender
  /missing = include.
```