

HRS Complex Sample Specification for SAS, Stata, and WesVar

by Robert B. Nielsen¹ and Martin C. Seay²

The following information is presented for public use HRS data. If utilizing RAND HRS data, simply replace STRATUM and SECU with RAESTRAT and RAEHSAMP, respectively.

WesVar:

To generate replicate weights

1. Open a SAS dataset with WesVar.
2. Specify the sample weight in "Full Sample" frame.
3. Bring all other variables into "variables" frame. No ID variable needed.
4. Save the dataset to convert to WesVar formatted dataset.
5. After dataset created, click on "Create Weights" button.
6. For "Method" click "BRR".
7. For "VarStrat" use STRATUM; for "VarUnit" use SECU.
8. No other frames need to be filled...click OK to create replicate weights that may be used in other programs or within WesVar.
9. Wesvar supports few procedures (bivariate estimates plus logistic and OLS regression) but you may export the replicate weights for use elsewhere (such as the svy procedures in Stata BRR or SAS repweights commands in SAS).

SAS:

```
proc survey___ data=__your_datafile___;      *(Taylor series);
  *your_secondary_command and your variable(s) or model statement;
  strata      STRATUM ;                      * this is the variance strata variable;
  cluster     SECU ;                        * this is the half sample variable;
  weight      _your_weight_;               * person/family/household weight variable;
run;
```

Of course, change the **proc survey___** for other procedures. For example:

```
proc surveylogistic data=__your_datafile___; *(Taylor series);
  model      DEPENDENT (event='1')= INDEPENDENT;
  strata     STRATUM ;                      * this is the variance strata variable;
  cluster    SECU ;                        * this is the half sample variable;
  weight     _your_weight_;               * person/family/household weight variable;
run;
```

```
proc surveylogistic data=your_datafile varmethod=BRR(fay);*(Replicate weights);
  model      DEPENDENT (event='1')= INDEPENDENT;
  weight     _your_weight_;               * person/family/household weight variable;
  repweights RepWt_1-RepWt_n;           * replicate weights 1 to n ;
run;
```

¹ Associate Professor, Department of Housing and Consumer Economics, University of Georgia, 205 Consumer Research Center, Athens, GA 30602. Email: rnielsen@uga.edu

² Assistant Professor, School of Family Studies and Human Services, 318 Justin Hall, Kansas State University, Manhattan, KS. 66506 Email: mseay@ksu.edu

Stata: (using the estimation of a mean value as our example)

Stata allows users to specify the complex sample design parameters prior to any analysis or with inline commands. Excellent examples are found in Leacock (2006). The Taylor series and replicate weights methods are straightforward with the use of the information below:

Taylor series linearization method using menus

Sampling unit is *SECU*; Strata is *STRATUM*.
 Sampling weight variable is your regular full sample *pweight*.
 Standard error can be specified as linearized (for Taylor) or other.
 Within the procedure, specify "survey data estimation" in SE.

Taylor series linearization method using code

```
First svyset SECU [pweight=_yourwgt_], strata(STRATUM) vce(linearized)
Then svy: mean _yourvariable_
```

Replicate weight method using menus

Sampling weight variable is your full sample *pweight*.
 Balanced repeated replicate(BRR) weight variables are *RepWt_1-RepWt_n*
 Fay's adjustment is .5
 Within the procedure, specify "survey data estimation" in SE.

Replicate weight method using code

```
First svyset [pweight=_yourwgt_], brrweight(RepWt_1-RepWt_n) fay(.5)
      vce(brr)
Then svy: mean _yourvariable_
```

As noted in HRS documentation, "...the sampling error codes that are provided should enable analysts to conduct either Taylor Series or Replicated estimation of sampling errors for survey statistics." (Heeringa and Conner, 1995; P. 48)

References

- Leacock, C. (Ed.) (2006). Getting started with the Health and Retirement Survey. Institute for Social Research: University of Michigan Survey Research Center. Available: <http://hrsonline.isr.umich.edu/sitedocs/dmgt/IntroUserGuide.pdf>
- Heeringa, S. G., & Conner, J. H. (1995). Technical Description of the Health and Retirement Survey Sample Design. Institute for Social Research: University of Michigan Survey Research Center. Available: <http://hrsonline.isr.umich.edu/sitedocs/userg/HRSSAMP.pdf>

Suggested citation:

Nielsen, R. B., & Seay, M. C. (2012). *HRS complex sample specification for SAS, Stata, and WesVar*. Technical note, Department of Housing and Consumer Economics, University of Georgia, Athens, GA.