

USER'S DIRECTIONS FOR DOWNLOADING AND ANALYZING HINTS 1 (2003) DATA-SAS

DOWNLOADING AND FORMATTING THE DATA FILE ON YOUR PERSONAL PC

Note: This will create a formatted temporary data file called HINTS2003.

- 1) Unzip the data file and place it in a folder on your computer;
- 2) Download the format file into the same folder as in step #1;
- 3) Copy and paste the following syntax into the editor window in SAS and run:

```
filename FORMATS "full pathway where format file is foundformats.hints2003.d2006_06_02.public.sas";  
filename FILE1 "full pathway where data file is found\hints2003.d2006_06_02.public.v8x";
```

```
proc format;  
%include FORMATS;
```

```
proc cimport file=FILE1 data=HINTS2003;  
run;
```

*/*You now have a temporary dataset named HINTS2003 that can be used in data steps, procedures, etc. later in the same program or in another program. For example:*/*

```
proc freq data=HINTS2003;  
  tables rslt;  
run;
```

ANALYZING HINTS DATA

If you plan on doing inferential statistical testing using the HINTS data, it is important that you utilize a statistical program that can incorporate the replicate weights that are included in the HINTS database. The issue is that the standard errors in your analyses will probably be underestimated if you don't incorporate the jackknife replicate weights; therefore, your p-values will be smaller than they "should" be, your tests will be more liberal, and you are more likely to make a type I error. Statistical programs like SUDAAN, STATA and Wesvar can handle the replicate weights. Below you will find some sample SUDAAN syntax to help guide your analysis.

Documentation and explanation of the weighting procedure for the survey data can be found in the 'Final Report' included in the data and documents download.

Note that analyses of large HINTS domains usually produce reliable estimates, but analyses of small domains may yield unreliable estimates, as indicated by their large variances. The analyst should pay particular attention to the standard error and coefficient of variation (relative standard error) for estimates of means, proportions, and totals, and the analyst should report these when writing up results. It is important that the analyst realizes that small sample sizes for particular analyses will tend to result in unstable estimates.

Denominator Degrees of Freedom (DDF)

The HINTS 1 (2003) database contains a set of 50 replicate weights to compute accurate standard errors for statistical testing procedures. These replicate weights were created using a jackknife minus one replication method; when analyzing one iteration of HINTS data, the proper denominator degrees of freedom (ddf) is 49. Thus, analysts who are only using the HINTS 1 (2003) data should use 49 ddf in their statistical models. HINTS statistical analyses that involve more than one iteration of data will typically utilize a set of $50 \times k$ replicate weights, where they can be viewed as being created using a stratified jackknife method with k as the number of strata, and $49 \times k$ as the appropriate ddf. Analysts who were merging two iterations of data and making comparisons should adjust the ddf to be 98 (49×2) etc.

Note:

Final sample weight variable name: fwgt

Replicate weight names (50): fwgt1 through fwgt50

SAMPLE SUDAAN CODE

Use the following code in SUDAAN procedures to specify the sample design:

```
proc procedurename data=datasetname design=jackknife;  
  weight fwgt;  
  jackwgt fwgt1-fwgt50 / adjjack=.98;
```

SAMPLE STATA SYNTAX

```
#delimit,  
svyset [pw=fwgt], jkrw(fwgt1-fwgt50, multiplier(.98));  
INFORMATION FOR USING WESVAR
```

WesVar is an interactive program centered on sessions called "workbooks" that are linked to a specific WesVar data file. It is software developed by Westat that uses replication methods to compute estimates and their

standard errors. When creating a WesVar data file, the exact method used to create the replicate weights should be specified properly. The sample design information is incorporated in the replicate weights, so no additional statements for sampling strata and primary sampling units are needed.

For the HINTS, a user first needs to transform the SAS data file(s) containing the final weight, replicate weights, and data into a WesVar file with the following specification:

Field(s)/Specification	Variable(s)/Value	Note
ID	CaseNo	Case Identifier
Full Sample	FWGT	Full sample weight
Replicates	FWGT1,...,FWGT50	Replicate sample weights
Method	JK1	Replication method
Df	49	Degree of freedom
Variables	As many needed for the analyses	Analysis variables

In a workbook, a user can request descriptive statistics, regression analyses, and tabulations of analysis variables. Regression requests support both linear and logistic models. Output listings include statistics such as sum of weights, means, and percentages, along with their corresponding standard errors, design effects, coefficients of variation (CV), and confidence intervals.

For more information, a user may check out the following Westat's website:

http://www.westat.com/westat/expertise/information_systems/wesvar/