

Estimation of the Variance Using Bootstrap Weights User's Guide for the BOOTVARE_V20.SAS Program (VERSION 2.0)

1. Introduction

This guide is for users of the SAS program BOOTVARE_V20.SAS which was created to estimate the variance using the bootstrap method. This program enables the estimation of variances for the National Population Health Survey (NPHS) (the household and institutional components) and for the Canadian Community Health Survey (CCHS).

Section 2 of this guide briefly explains the resampling method used to estimate the variance. Section 3 gives a detailed explanation of the rules to follow when using the BOOTVARE_V20.SAS program, as well as the preliminary steps that must be taken. The programs to use are in Appendix A. Appendix B contains a complete example (programs and results). Finally, the particularities of each survey (file names, identification variables to use, etc.) are in Appendix C.

Changes from the Previous Version:

The biggest change made to the program is that it is easier for the users to use. All of the commands that need to be changed have been grouped together. The required changes are stated more clearly and the commands that do not need to be modified (the macros, for example) have been removed from the program. These commands are now in the program MACROE_V20.SAS, which the user does not need to modify.

It is also easier to test the program. The only change that needs to be made is to specify a smaller number of bootstrap weights to use (at least 2) at the beginning of the program.

The macro for general linear models is not available anymore. This macro needed a lot of modifications by the user in order to adapt it for each specific analysis. For the analysis of differences between means, it is suggested to use the macro “Difference Between Ratios”, which is easier to use.

Finally, the program can be used with the CCHS data as well as with the NPHS data. No changes are necessary because the program automatically detects which survey the user is working with.

Please note that the program was tested and works with SAS version 6.12 and 8.2.

2. Bootstrap Method

The sampling designs for health surveys are complex. Since the variance for such designs cannot be calculated with simple formulas, a resampling method is necessary to calculate the variance.

The bootstrap method consists of subsampling the initial sample. Within each stratum, a simple random sample (SRS) is selected, with replacement, from $n-1$ clusters within the n clusters of the stratum. This creates B new samples (or repetitions). The same estimate is then calculated for each of the B samples, which gives B different estimates. To obtain each of the B estimates, a specific weight for each sample is necessary. In each SRS sample, the weight is then recalculated for each record in the stratum. These B weights, the bootstrap weights, have been produced and are available with the data.

With the BOOTVARE_V20.SAS program, the bootstrap weights are used to obtain precise estimates of the variance for simple statistics such as totals and ratios, as well as for more complex analysis like regressions. The same rules for confidentiality and release guidelines apply to the variance estimates obtained through the bootstrap method.

Here are the major steps to follow to obtain a valid estimate for the variance of a particular estimate.

- A) Calculate an estimate (total, ratio, etc.) using the final weight included in the data file. This estimate is the point estimate.
- B) Calculate the same estimate, this time using each of the B bootstrap weights contained in the bootstrap file. B estimates (total, ratio, etc) are then obtained.
- C) Finally, calculate the variance of the B estimates. This variance is the estimate of the variance of the point estimate calculated in A.

3. Variance Estimation with the BOOTVARE_V20.SAS Program

The BOOTVARE_V20.SAS program enables the calculation of estimates of the variance for totals, ratios, differences between ratios and linear or logistic regression parameters.

Variance estimation is done in *two steps* and involves the use of three SAS programs. The *first step* consists of creating a data file containing the variables to be used to do the analysis (first program). The *second step* is the use of BOOTVARE_V20.SAS (and MACROE_V20.SAS) to estimate the variances.

Step 1: Creation of the Analysis File

The user needs to create a SAS data file which will be used as the input file for the program estimating the variance in step 2. The following tasks must be done in this step:

1. Reading of the input file
2. Creation of the variables required for the analysis

1 - Reading of the input file: The analysis file is created from the file containing the survey data. This file must be read, and the layout, which is also provided, must be used to specify the variables contained in the file. See Appendix C for the file and variable names to use.

2 - Creation of the variables required for the analysis: Dichotomous variables (1 or 0) must be created by using the qualitative variables of interest contained in the initial data file. This is to identify which records have the different characteristics that are being studied. The dichotomous variable will take a value of 1 for records that have the characteristic of interest and a value of 0 otherwise. For example, for the estimation of totals, ratios and differences between ratios, these dichotomous variables will identify the records with the characteristic of interest in order to sum their weights to obtain the total or ratio in step 2. See the example in Appendix B for more details.

The analysis file must contain:

- The necessary variables for the analysis (dichotomous variables and other variables that do not need to be modified). To reduce the runtime of the program, DO NOT keep unnecessary variables.
- The identification variable(s) of the respondents.
- If needed, the breakdown variable(s), identifying the groups for which a separate analysis is desired (ex: province, sex, etc...).
- If the analysis is only of interest for a certain subgroup (for example, a province or an age group), keep only the records that are part of this subgroup.

REMARKS:

- It is recommended that point estimates be calculated at this step to be sure that the desired estimate is being calculated correctly, and that the program `BOOTVARE_V20.SAS` correctly calculates the same estimate. In this case, it is necessary to keep the weight variable when creating the analysis file.
- Percentage estimates are ratios. Two dichotomous variables must then be created; one for the numerator identifying the individuals having the characteristic of interest, and one for the denominator identifying the individuals that are part of the group of interest.
- Mean estimates can be considered as ratios. The numerator is the variable identifying the characteristic of interest within the group of interest, and the denominator is the variable identifying the individuals that are part of the group of interest.

The users must create their own program to prepare the analysis file containing the necessary variables for the analysis. An example of a program that creates this file is included in Appendix A (the program `STEP1.SAS`)

Step 2: Variance Calculation Using the `BOOTVARE_V20.SAS` Program

Once the new SAS data file is created in step 1, the next step consists of running the `BOOTVARE_V20.SAS` program. Before running it, the desired parameters and analysis must be specified. This program calls the `MACROE_V20.SAS` program.

MACROE_V20.SAS contains the different macros that enable the calculation of the estimates. *For standard use of the variance estimation program, no modification of the MACROE_V20.SAS program by the user is necessary.* Changes may be required in certain cases explained later.

The BOOTVARE_V20.SAS program is included in Appendix A. The parts to be changed are in ***bold characters***. The rest of the program does not need to be changed. The program is divided into ***two sections***. The ***first section*** enables the users to define the different parameters, and the ***second section*** enables the specification of the desired analysis.

Section 1:

In this section, the user must specify:

- The name of the directories where to find the analysis file created in step 1 and where the results will be saved.
- The name of the data file to use (analysis file created in step 1)
- The name and directory of the bootstrap weights file
- The breakdown variable(s) (ie: if the analysis is done separately for specific sub-groups (ex: provinces, sex))
- The number of bootstrap weights to use
- The name of the directory where to find the program MACROE_V20.SAS

N.B. Appendix C contains information about the files the user will need (file names, certain variable names, number of bootstrap weights)

Section 2:

This section enables the specification of the desired analysis to obtain the estimates of the variance for:

- Totals
- Ratios
- Differences between ratios
- Regression models (linear or logistic)

For differences between ratios: if the user wants to calculate the difference between ratios, they must modify, *if necessary*, the macro *diff_rat* in the MACROE_V20.SAS program, to suit their needs (see the notes included in the BOOTVARE_V20.SAS program for more details.)

For regressions: Categorical variables will be treated as continuous variables. Dichotomous variables must be created in step 1 for each possible value (except one) of the categorical variable in order to treat this variable properly.

Modification to the program for testing purposes: Running the program could take a while (especially for complex model analysis). It is possible to reduce the number of bootstrap weights

used in order to test the program. **(However, to obtain the final estimates of the variance, it is important to use all of the bootstrap weights provided.)** To test the program, all that must be done is a modification of the number of weights to use in the first part of the BOOTVARE_V20.SAS program.

Results Obtained with BOOTVARE_V20.SAS

The following results are obtained after running BOOTVARE_V20.SAS for totals, ratios and differences between ratios. See Appendix B for interpretation of the results:

TYPE:	Type of estimate (total, ratio,diff_ratio)
VAR1 & VAR2:	Variables used to calculate the estimates. For a total, VAR2 = None.
VAR3 & VAR4:	Variables used to calculate the difference between ratios.
YHAT:	Parameter estimate (in % for ratios)
BS_SD:	Standard deviation
BS_CV:	Coefficient of variation
CIL95:	Lower limit of the 95% confidence interval
CIU95:	Upper limit of the 95% confidence interval

Note: For the means (considered as ratios), the parameter estimate and the confidence interval limits must be divided by 100 since the results are in percentage.

The following results are obtained after running BOOTVARE_V20.SAS for linear and logistic regressions. See Appendix B for interpretation of the results:

BETA:	Parameter to estimate
BHAT:	Parameter estimate
ODDS:	Odds ratio (logistic regression only)
WALD:	Wald statistic (logistic regression only)
PVALUE:	P-value of the Wald statistic (logistic regression only)
BSVAR:	Variance of the parameter estimate
BS_SD:	Standard deviation of the parameter
BS_CV:	Coefficient of variation for the parameter estimate
CIL95:	Lower limit of the 95% confidence interval (for the odds ratio if logistic regression)
CIU95:	Upper limit of the 95% confidence interval (for the odds ratio if logistic regression)

Appendix A contains the BOOTVARE_V20.SAS program, preceded by an example of a program to use to prepare the analysis file (STEP1.SAS). Appendix B contains a complete example (programs and results). Finally, Appendix C contains information about the files that the user needs (file names, names of certain variables, number of bootstrap weights).

APPENDIX A - Programs to Run

STEP 1.SAS Program

(Used as an example; the users can use their own program)

The parts in **bold** need to be changed.

```
*****
*                               *
*           STEP1.SAS           *
*                               *
*   This program creates the SAS datafile
*   containing the necessary variables
*   for the BOOTVARE_V20.SAS program
*****;

LIBNAME in1 'directory_to_save_file_in';

*** Creation of the SAS data file containing the variables and cases required for the analysis. Note
*** that this file should be as small as possible (containing only necessary variables and cases) in
*** order to reduce time and memory requirements especially if regression type analysis are to be done. ;

data in1.Name_of_analysis_file;          /* file to be used with BOOTVARE_V20.SAS */
  %let datafid= " name_and_location_of_source_file ";
  %include " name_and_location_of_layout ";

*** Creation of Dichotomous Variables ***
*** (examples are presented below using NPHS cycle 3 variables) ***;

  /* diabetes */
  if ccc8_1j=1 then diab=1;
  else diab=0;

  /* sex */
  if dhc8_sex=1 then males=1;
  else males=0;
  if dhc8_sex=2 then females=1;
  else females=0;

  /* diabetes*sex */
  mdiab = diab * males;    /* male diabetics */
  fdiab = diab * females;  /* female diabetics */

  keep list of variables to keep;

  * It is recommended that only the necessary variables be kept
  * in order to reduce the runtime of BOOTVARE_V20.SAS.
  * IMPORTANT: the identification variables and, if necessary,
  * the breakdown variable (ex: province, sex) must be kept. The
  * weight variable also must be kept if point estimates are
  * calculated at this step
  *
run;

*****
* Calculation of point estimates
* Suggested, but not required...
*****;

proc freq data=in1.Name_of_analysis_file;
  table variables of interest;
  weight weight_variable;
run;

proc logistic data=in1.Name_of_analysis_file ;
  model dependent_variable = independent variables ;
  weight weight_variable;
run;
```

BOOTVARE V20.SAS Program

The parts in ***bold*** need to be changed.

```

...
...
...
/*****
/****          SECTION 1          ****
/*****
/****
/**** This section lets the user specify the different parameters of ****
/**** interest (variable names, directory names, file names, etc.) ****
/****
/*****

*****
** SPECIFY THE NAME OF THE FOLLOWING 2 DIRECTORIES (directories only): **
*****
libname in1 "name_of_the_directory_containing_analysis_file_step_1"; /* (ex: c:\data) */
libname out "name_of_the_directory_to_save_results_in"; /* (ex: c:\output)*/

*****
** SPECIFY THE NAME OF THE ANALYSIS FILE (CREATED IN STEP 1) (without extension): **
*****
%let Mfile = in1.Name_of_analysis_file;

*****
** SPECIFY THE NAME OF THE FILE CONTAINING THE BOOTSTRAP WEIGHTS: **
** NB: Only run one of the two following series of commands **
** (comment the other one out, or erase it): **
*****
* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN SAS FORMAT (remove the " * ")
-----
*          libname in2 "directory_name_containing_bootstrap_weights_file"; /* (ex: c:\bootstrp)*/
*          %let bsamp=in2.SAS_file_name_containing_the_weights_without_extension ;

* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN ASCII (.TXT) FORMAT (remove the " * ")
-----
*          data bootwt;
*          %let datafid="(directory)location_and_bootstrap_weights_file(with_extension) ";
*          %include "(directory)location_and_file_name_of_layout(with_extension)";
*          run;
*          %let bsamp=bootwt;

*****
** SPECIFY, IF DESIRED, THE BREAKDOWN VARIABLE(S) (EG: PROVINCE, SEX, ETC...): **
** write the name of the breakdown variable(s) below. **
**
** - If the analysis includes all of the data in the file created in step 1, put **
** a dot. (%let classes=.) **
** - If more than one variable, leave a blank between each variable **
** (%let classes=var1 var2) **
** - DO NOT ERASE OR COMMENT OUT THIS COMMAND **
*****
%let classes = breakdown_variable(s)_or_a_dot;

*****
** SPECIFY THE NUMBER OF BOOTSTRAP WEIGHTS TO USE: **
** IMPORTANT: IT IS NECESSARY TO USE ALL THE BOOTSTRAP WEIGHTS WHEN PERFORMING THE **
** FINAL ANALYSIS. THE COMPLETE BOOTVARE_V20.SAS PROGRAM MUST THEN BE RUN. **
**
** - Refer to Appendix C in the documentation to find the number of **
** weights that the weights file contains. **
** - For testing, B must be >= 2. **
*****
%let B = number_of_weights_to_use ;

*****
** SPECIFY THE DIRECTORY AND THE NAME OF THE FILE THAT CONTAINS THE MACROS **
** (THE PROGRAM MACROE_V20.SAS IF NO MODIFICATIONS HAVE BEEN MADE BY THE USER) **
*****
%include "directory_name_of_macro_e_v20.sas\MACROE_V20.SAS";

```

```

/*****
***          SECTION 2          ***
****
**** This section lets the user specify the different analyses of interest. ****
****
****
****
...
...
...

* TO OBTAIN VARIANCE ESTIMATES OF A TOTAL, RUN:
-----;

    * %total(variable_name);

* TO OBTAIN VARIANCE ESTIMATES OF A RATIO, RUN:
-----;

    * %ratio (numerator_variable, denominator_variable);

* TO OBTAIN VARIANCE ESTIMATES OF A DIFFERENCE BETWEEN RATIOS, RUN:
-----;
    *NOTE: see the comment at the beginning of section 2 ... ;

    * %diff_rat(VAR1, VAR2, VAR3, VAR4);

                * where: var1 : the numerator variable of the first ratio      *
                *          var2 : the denominator variable of the first ratio   *
                *          var3 : the numerator variable of the second ratio     *
                *          var4 : the denominator variable of the second ratio  *;

* TO OBTAIN VARIANCE ESTIMATES OF REGRESSION PARAMETERS, RUN:
-----;

    * %regress(dependent_variable, independent_variables(no comma)) ;

* TO OBTAIN VARIANCE ESTIMATES OF LOGISTIC REGRESSION PARAMETERS, RUN:
-----;

    * %logreg(dependent_variable, independent_variables(no comma));

%output; /*Displays the results on the screen. Do not modify. */

* TO SAVE THE RESULTS IN A FILE, RUN: (remove the "")
-----;

    * data out.Results_filename ;
    * set &result ;
    * run;

/* end of BOOTVARE_V20.SAS program */

```


APPENDIX B

This is a complete example showing how to use the program BOOTVARE_V20.SAS. First, the analysis data file is created (step 1). Then, BOOTVARE_V20.SAS is adapted to obtain the desired analysis. The results that are produced follow the programs.

Example:

This example uses the cycle 3 (1998) cross-sectional file of the NPHS general component. This example:

- 1- Computes the total number and the proportion of diabetics in the general population and among men only, for each province (only four provinces will be kept).
- 2- Studies the relationship between diabetes, sex and type of interview (proxy or not), for each province.

Step 1:

```
*****
*                               *
*               STEP1.SAS       *
*                               *
*   This program creates the SAS datafile      *
*   containing the necessary variables          *
*   for the BOOTVARE_V20.SAS program          *
*                               *
*****;
```

```
LIBNAME in1 'C:\BOOTVAR\';

*** Creation of the SAS data file containing the variables and cases required for the analysis. Note
*** that this file should be as small as possible (containing only necessary variables and cases) in
*** order to reduce time and memory requirements especially if regression type analysis are to be done. ;

data in1.diabetes;
    %let datafid='D:\Data\h35.txt'; /* file to be used with BOOTVARE_V20.SAS */
    %include 'D:\Layout\h35_i.sas';

*** keep only 4 provinces;
    if prc8_cur in (10 24 35 59);
```

Appendix B

```

*** Creation of Dichotomous Variables ***
*** (examples are presented below using NPHS cycle 3 variables) ***;

/* diabetes */
if ccc8_1j=1 then diab=1;
else diab=0;

/* Dichotomous variable (0/1) for type of interview */

nonproxy=0;
if am58_pxy>2 then nonproxy=.;
if am58_pxy=2 then nonproxy=1;

/* sex */
total=1;
if dhc8_sex=1 then males=1;
else males=0;
if dhc8_sex=2 then females=1;
else females=0;

/* diabetes*sex */
mdiab = diab * males; /* male diabetics */
fdiab = diab * females; /* female diabetics */

keep diab total males females mdiab fdiab nonproxy wt58 realukey personid prc8_cur;

* It is recommended that only the necessary variables be kept *
* in order to reduce the runtime of BOOTVARE_V20.SAS. *
* IMPORTANT: the identification variables and, if necessary, *
* the breakdown variable (ex: province, sex) must be kept. The *
* weight variable also must be kept if point estimates are *
* calculated at this step *;

run;

*****
* Calculation of point estimates *
* Suggested, but not required... *
*****;

PROC SORT DATA=in1.diabetes; BY prc8_cur; RUN;

proc freq data=in1.diabetes;
table diab mdiab fdiab;
by prc8_cur;
weight wt58;
run;

proc logistic data=in1.diabetes;
model diab = nonproxy females;
by prc8_cur;
weight wt58;
TITLE "Relationship between diabetes, sex and type of interview";
run;

```

Step 2 - BOOTVARE V20.SAS program:

```

...
...
...
/*****
/****              SECTION 1              ****
/*****
/**** This section lets the user specify the different parameters of ****
/**** interest (variable names, directory names, file names, etc.) ****
/****              ****
/*****

*****
** SPECIFY THE NAME OF THE FOLLOWING 2 DIRECTORIES (directories only): **
*****
libname in1 "C:\BOOTVAR";          /* (ex: c:\data) */
libname out "C:\BOOTVAR";          /* (ex: c:\output)*/

*****
** SPECIFY THE NAME OF THE ANALYSIS FILE (CREATED IN STEP 1) (without extension): **
*****
%let Mfile = in1.diabetes;

*****
** SPECIFY THE NAME OF THE FILE CONTAINING THE BOOTSTRAP WEIGHTS: **
** NB: Only run one of the two following series of commands **
** (comment the other one out, or erase it): **
*****

* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN SAS FORMAT (remove the " * ");
-----
libname in2 "D:\bootstrp\DATA";
%let bsamp = in2.b5h35;

* EXECUTE THIS PART IF THE BOOTSTRAP WEIGHTS ARE IN ASCII (.TXT) FORMAT (remove the " * ");
-----

* data bootwt;
* %let datafid="(directory)location_and_bootstrap_weights_file_(with_extension) ";
* %include "(directory)location_and_file_name_of_layout_(with_extension)";
* run;
* %let bsamp=bootwt;

*****
** SPECIFY, IF DESIRED, THE BREAKDOWN VARIABLE(S) (EG: PROVINCE, SEX, ETC...): **
** Write the name of the breakdown variable(s) below. **
** **
** - If the analysis includes all of the data in the file created in step 1, put **
** a dot. (%let classes = .) **
** - If more than one variable, leave a blank between each variable **
** (%let classes=var1 var2) **
** - DO NOT ERASE OR COMMENT OUT THIS COMMAND **
*****
%let classes = prc8_cur;

*****
** SPECIFY THE NUMBER OF BOOTSTRAP WEIGHTS TO USE: **
** IMPORTANT: IT IS NECESSARY TO USE ALL THE BOOTSTRAP WEIGHTS WHEN PERFORMING THE **
** FINAL ANALYSIS. THE COMPLETE BOOTVARE_V20.SAS PROGRAM MUST THEN BE RUN. **
** **
** - Refer to Appendix C in the documentation to find the number of **
** weights that the weights file contains. **
** - For testing, B must be >= 2. **
*****
%let B = 500 ;

*****
** SPECIFY THE DIRECTORY AND THE NAME OF THE FILE THAT CONTAINS THE MACROS **
** (THE PROGRAM MACROE_V20.SAS IF NO MODIFICATIONS HAVE BEEN MADE BY THE USER) **
*****
%include "C:\BOOTVAR\MACROE_V20.SAS";

```

```

/*****
/****                               SECTION 2                               ****
/*****                               ****
/**** This section lets the user specify the different analyses of interest. ****
/****                               ****
/*****

...
...
...

* TO OBTAIN VARIANCE ESTIMATES OF A TOTAL, RUN:
-----;

    * %total(variable_name);

    %total(diab);
    %total(mdiab);

* TO OBTAIN VARIANCE ESTIMATES OF A RATIO, RUN:
-----;

    * %ratio (numerator_variable, denominator_variable);

    %ratio(diab,total);
    %ratio(mdiab,males);

* TO OBTAIN VARIANCE ESTIMATES OF A DIFFERENCE BETWEEN RATIOS, RUN:
-----;
    *NOTE: see the comment at the beginning of section 2 ... ;

    * %diff_rat(VAR1,VAR2,VAR3,VAR4);

                * where:  var1   : the numerator variable of the first ratio   *
                *         var2   : the denominator variable of the first ratio  *
                *         var3   : the numerator variable of the second ratio   *
                *         var4   : the denominator variable of the second ratio  *

* TO OBTAIN VARIANCE ESTIMATES OF REGRESSION PARAMETERS, RUN:
-----;

    * %regress(dependent_variable, independent_variables_(no comma)) ;

* TO OBTAIN VARIANCE ESTIMATES OF LOGISTIC REGRESSION PARAMETERS, RUN:
-----;

    * %logreg(dependent_variable, independent_variables_(no comma));
    %logreg(diab,nonproxy females);

%output; /*Displays the results on the screen. Do not modify. */

* TO SAVE THE RESULTS IN A FILE, RUN: (remove the "")
-----;

    data out.results ;
    set &result ;
    run;

/* end of BOOTVARE_V20.SAS program */

```

Results and interpretation:

The tables on the next page present the results of the analyses done using the programs from the example. Results for the totals and ratios are presented in the first table. For example, if we want the ratio of the number of diabetic males to the total number of males, in Ontario, we look at observation 12. The region 35 corresponds to the province of Ontario (see the data dictionary document included on the CD-ROM for the codes associated with each province) and the variable Type indicates the type of analysis, in this case a ratio. We find the variables *mdiab* (VAR1) as the numerator of the ratio and *males* (VAR2) as the denominator. The estimate of the ratio is 3.57% (YHAT) with a standard deviation of 0.29 (BS_SD) and a coefficient of variation of 8.26 (BS_CV). The 95% confidence interval for this estimate is (2.99%, 4.14%) (CIL95, CIU95).

Results from the logistic regression are shown in the second table. For example, the estimate of the parameter for the variable *females* in Ontario (observation 9) is -0.34507 (BHAT) and the odds ratio is 0.70817 (ODDS). The Wald's statistic for this parameter and its associated p-value are 7.996 (WALD) and p=0.004687 (PVALUE) respectively. The estimates of the variance and the standard deviation for the parameter estimate are 0.014891 (BS_VAR) and 0.12203 (BS_SD) and the coefficient of variation is 35.36 (BS_CV). Finally, the confidence interval for the odds ratio is (0.55753, 0.89952) (CIL95, CIU95).

Following the results, is the execution time of this program.

Appendix B

Variance estimation using 500 bootstraps for Totals and Ratios

Obs	PRC8_CUR	type	var1	var2	yhat	bs_sd	bs_cv	cil95	ciu95
1	10	Total	diab	blank	20741.31	1778.61	8.58	17255.23	24227.39
2	10	Total	mdiab	blank	7029.11	1356.31	19.30	4370.73	9687.48
3	10	Ratio	diab	total	3.85	0.33	8.58	3.21	4.50
4	10	Ratio	mdiab	males	2.63	0.51	19.30	1.64	3.63
5	24	Total	diab	blank	205292.21	16330.49	7.95	173284.44	237299.98
6	24	Total	mdiab	blank	110452.77	10818.85	9.80	89247.83	131657.72
7	24	Ratio	diab	total	2.87	0.23	7.95	2.42	3.32
8	24	Ratio	mdiab	males	3.12	0.31	9.80	2.52	3.72
9	35	Total	diab	blank	362439.56	20692.51	5.71	321882.24	402996.88
10	35	Total	mdiab	blank	198237.67	16369.18	8.26	166154.07	230321.26
11	35	Ratio	diab	total	3.22	0.18	5.71	2.86	3.58
12	35	Ratio	mdiab	males	3.57	0.29	8.26	2.99	4.14
13	59	Total	diab	blank	110375.38	10661.17	9.66	89479.48	131271.27
14	59	Total	mdiab	blank	62808.64	8301.60	13.22	46537.50	79079.77
15	59	Ratio	diab	total	2.83	0.27	9.66	2.29	3.37
16	59	Ratio	mdiab	males	3.24	0.43	13.22	2.40	4.08

Variance estimation using 500 bootstraps for Logistic regressions Dependent variable: diab

Obs	PRC8_CUR	beta	bhat	Odds	wald	pvalue	bs_var	bs_sd	bs_cv	cil95	ciu95
1	10	Intercept	-4.00372	0.01825	255.389	0.000000	0.062766	0.25053	6.26	0.01117	0.02982
2	10	nonproxy	0.85783	2.35803	11.819	0.000586	0.062260	0.24952	29.09	1.44595	3.84543
3	10	females	0.46625	1.59401	2.859	0.090860	0.076036	0.27575	59.14	0.92848	2.73660
4	24	Intercept	-3.89892	0.02026	852.643	0.000000	0.017829	0.13352	3.42	0.01560	0.02633
5	24	nonproxy	0.89121	2.43807	34.116	0.000000	0.023281	0.15258	17.12	1.80787	3.28795
6	24	females	-0.36991	0.69080	5.783	0.016180	0.023660	0.15382	41.58	0.51100	0.93387
7	35	Intercept	3.57575	0.02799	961.191	0.000000	0.013302	0.11534	3.23	0.02233	0.03510
8	35	nonproxy	0.60946	1.83943	23.520	0.000001	0.015793	0.12567	20.62	1.43785	2.35318
9	35	females	-0.34507	0.70817	7.996	0.004687	0.014891	0.12203	35.36	0.55753	0.89952
10	59	Intercept	-3.99955	0.01832	301.873	0.000000	0.052991	0.23020	5.76	0.01167	0.02877
11	59	nonproxy	1.05756	2.87934	18.892	0.000014	0.059203	0.24332	23.01	1.78722	4.63883
12	59	females	-0.46337	0.62916	4.591	0.032133	0.046763	0.21625	46.67	0.41180	0.96125

Length of time required to run the program

Obs	start	stop
1	20MAR02:10:16:17	20MAR02:10:23:51

APPENDIX C

Variable and File Names to Use With the Program

Note: For a complete list of the available variables, consult the data dictionary provided with the documentation.

NPHS - Household Component					
	Name of data file	Name of bootstrap weights file (ASCII format: .txt SAS format: .sd2 or .sas7bdat)	Identification variables	Weight variables	# of weights
General File :					
Cycle 1	H3H5.txt	B5H35	REALUKEY PERSONID	WT54 (M) SHRWT5 (S)	500
Cycle 2	H35.txt	B5H35	REALUKEY PERSONID	WT56 (M) WT56_S (S)	100
Cycle 3	H35.txt	B5H35	REALUKEY PERSONID	WT58 (M) WT58_S (S)	500
Health File :					
Cycle 1	H3H5H6.txt	B5H356	REALUKEY PERSONID	WT64 (M) SHRWT6 (S)	500
Cycle 2	H356.txt	B5H356	REALUKEY PERSONID	WT66 (M) WT66_S (S)	500
Cycle 2 <i>Health Promotion Survey Questions</i>	H356.txt	B5H356A	REALUKEY PERSONID	WT66_N (M) WT66_SN (S)	500
Cycle 2 <i>Child Health Services Questions Man. and Alb.</i>	H356.txt	B5H356C	REALUKEY PERSONID	WT66_N (M) WT66_SN (S)	500
Cycle 3	H356.txt	B5H356	REALUKEY PERSONID	WT68 (M) WT68_S (S)	500
Longitudinal File :					
Cycle 2 - Full	LNGF.txt	B5LONGF	REALUKEY PERSONID	WT66LF (M) WT66SLF (S)	500
Cycle 2 - Partial (Master only)	LNGP.txt	B5LONGP	REALUKEY PERSONID	WT66LP (M)	500

(M) - Master Files
(S) - Share Files

Appendix C

Cycle 2 - Square (Master only)	LONG.txt	B5LONGS	REALUKEY PERSONID	WT66LF (M)	500
Cycle 3 - Full	LNGF.txt	B5LNGF	REALUKEY PERSONID	WT68LF (M) WT68_SLF(S)	500
Cycle 4 - Full	LONG.txt (M) LNGF.txt (S)	B5LNGF	REALUKEY PERSONID	WT60LF (M) WT60SLF (S)	500
Cycle 4 - Square (Master only)	LONG.txt	B5LONG	REALUKEY PERSONID	WT60SLF (M)	500
Cycle 4 - Full Cycle 1 & Cycle 4 (Master only)	LONG.txt	B5LNGFE	REALUKEY PERSONID	WT60LFE (M)	500

NPHS - Institution Component					
	Name of data file	Name of bootstrap weights file <i>(ASCII format: .txt SAS format: .sd2 or .sas7bdat)</i>	Identification variables	Weight variables	# of weights
Longitudinal File:					
Cycle 3	LNGF.txt	B2LNGF	UNIQUEID	WTI8LF (M) WTI8_S (S)	2000
Cycle 4	LNGF.txt	B2LNGF	UNIQUEID	WTI0LF (M) WTI0SLF (S)	2000

CCHS					
	Name of data file	Name of bootstrap weights file <i>(ASCII format: .txt SAS format: .sd2 or .sas7bdat)</i>	Identification variables	Weight variables	# of weights
Cycle 1.1	HS.txt	B5	SAMPLEID PERSONID	WTSA_M (M) WTSA_S (S)	500
Cycle 1.1 P.E.I buy-in	HS.txt	B5_PEI	SAMPLEID PERSONID	WTSA_PEM (M) WTSA_PES (S)	500
Cycle 1.1 Fourth Quarter	HS.txt	B5_Q4	SAMPLEID PERSONID	WTSA_Q4M (M) WTSA_Q4S (S)	500

(M) - Master Files
(S) - Share Files