

Estimation of the Variance Using Bootstrap Weights User's Guide for the BOOTVARE_V21.SPS Program (VERSION 2.1)

1. Introduction

This guide is for users of the SPSS program BOOTVARE_V21.SPS 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 institution 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_V21.SPS 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:

It is in fact the first official SPSS version of BOOTVARE. A Beta version was available before on request only. 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 have been removed from the program. These commands are now in the program MACROE_V21.SPS, 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 analyze 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 SPSS version 9.0 and 10.0.

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_V21.SPS 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_V21.SPS Program

The BOOTVARE_V21.SPS 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 SPSS 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_V21.SPS (and MACRO_V21.SPS) to estimate the variances.

Step 1: Creation of the Analysis File

The user needs to create a SPSS 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_V21.SPS 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.SPS). IT IS NOT NECESSARY TO USE A SPSS PROGRAM TO CREATE THE ANALYSIS DATA FILE. DICHOTOMOUS VARIABLES CAN BE COMPUTED DIRECTLY IN THE DATA EDITOR WINDOW AND THE ANALYSIS DATA FILE CAN BE CREATED FROM THERE. (Don't forget to keep only the necessary variables.)

Step 2: Variance Calculation Using the BOOTVARE_V21.SPS Program

Once the new SPSS data file is created in step 1, the next step consists of running the BOOTVARE_V21.SPS program. Before running it, the desired parameters and analysis must be specified. This program calls the MACROE_V21.SPS program.

MACROE_V21.SPS contains the different macros that enable the calculation of the estimates. *For standard use of the variance estimation program, no modification of the MACROE_V21.SPS program by the user is necessary.* Changes may be required in certain cases explained later.

The BOOTVARE_V21.SPS 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 and directory 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_V21.SPS

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 *Results* in the MACROE_V21.SPS program, to suit their needs (see the notes included in the BOOTVARE_V21.SPS 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 by the user is a modification of the number of weights to use in the first part of the BOOTVARE_V21.SPS program.

Results Obtained with BOOTVARE_V21.SPS

The following results are obtained after running BOOTVARE_V21.SPS for totals, ratios and differences between ratios. See Appendix B for interpretation of the results:

TYPE:	Type of estimate (total, ratio,diff ratio)
VARIABLE:	Variable(s) used to calculate the estimates.
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_V21.SPS for linear and logistic regressions. See Appendix B for interpretation of the results:

TYPE:	Type of estimate (regression, logistic regression, ...)
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_V21.SPS program, preceded by an example of a program to use to prepare the analysis file (STEP1.SPS). 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.SPS Program

(Used as an example; the users can use their own program or use the Data Editor window)

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

```
*****
*                               STEP1.SPS                               *
*                               *                                       *
* This program creates the SPSS datafile containing the necessary variables for the *
* BOOTVARE_V21.SPS program.                                           *
*                               *                                       *
* IT IS NOT NECESSARY TO USE THIS PROGRAM. Dichotomous variables can *
* be computed directly in the Data Editor window and the analysis data file can be *
* created from there.                                                 *
*****

SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO.

* Creation of the SPSS "sav" 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.

DEFINE !Go()
!LET !Layout = 'name_and_location_of_layout'
!LET !DataIn = 'name_and_location_of_source_file'
!LET !DataOut = 'name_of_resulting_file_in_SPSS_format' /* <- (analysis file to be used in BOOTVARE_V21.SPS)

* Do not modify the following block of instructions.
FILE HANDLE INFILE/NAME = !QUOTE(!DataIn).
INCLUDE FILE=!QUOTE(!Layout).
SET PRI ON.

*****
* Creation of Dichotomous Variables ( -> REPLACE all following instructions (up to "SAVE" ) *
* (examples are presented below using NPHS cycle 3 variables) *
*****

NUMERIC diab males females mdiab fdiab.

IF (ccc8_1j=1) diab=1.
IF (ccc8_1j~=1) diab=0.

IF (sex=1) males=1.
IF (sex~=1) males=0.
IF (sex=2) females=1.
IF (sex~=2) females=0.

compute mdiab = males*diab.
compute fdiab = females*diab.

execute.
```

Appendix A

SAVE OUTFILE !QUOTE(!DataOut)
/KEEP = **list of variables to keep** .

* It is recommended that only the necessary variables be kept *
* in order to reduce the runtime of BOOTVARE_V21.SPS. *
* IMPORTANT: the identification variables and, if necessary, *
* the breakdown variable(s) (ex: province, sex) must be kept. The *
* weight variable also must be kept if point estimates are *
* calculated at this step *

* Replace "SAVE" by "XSAVE" in previous command if point estimates are desired at this time.

*** DO NOT FORGET TO RUN THE LAST LINES OF THE PROGRAM (AFTER THE EXAMPLE) ***.

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

*** Examples: ***.

WEIGHT BY **weight_variable**.

FREQ VAR = **variable list**.
LOGISTIC REG VAR = **var** WITH **variable list** /PRI SUM
/CRI BCO(0.0001) LCO(0) ITE(25) EPS(10E-12).

WEIGHT OFF.

* DO NOT DELETE the following instructions

SET PRI NO.
!ENDDDEFINE /* !Go.

SET MPR ON.
!Go.
SET MPR NO.

BOOTVARE V21.SPS 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.) ***
***                                     ***
*****
```

SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO.

DEFINE !Bootvar()

!Let !Tot=!Null
!Let !Reg=!Null

```
*****
* SPECIFY THE ANALYSIS DATA FILE TO BE USED (created in step 1) (Folder and file) : *
*****
```

!LET !Mfile=**'Folder_and_file_name_of_the_analysis_file(step 1)'** /* <- (ex: c:\data\analysis.sav).

```
*****
* SPECIFY THE BOOTSTRAP WEIGHTS FOLDER AND FILE NAMES : *
* 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 SPSS FORMAT (remove the " * ") *
*****
```

*!LET !Bwsav=**'Folder_and_file_names_of_the_bootstrap_weights_in_SPSS_format(with_extension)'**

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

*!LET !BWTXT=**'Folder_and_file_names_of_the_bootstrap_weights_in_text_format(with_extension)'**

*!LET !Bwlay=**'Folder_and_file_names_of_the_bootstrap_layout_file_in_SPSS_syntax(with_extension)'**

*!LET !Bwsav=

'Folder_and_file_names_of_the_bootstrap_weights_in_SPSS_format(created_in_this_section)with_extension'

*FILE HANDLE BOOTLAY/NAME = !QUOTE(!BWTXT).

*INCLUDE FILE=!QUOTE(!Bwlay).

*SAVE OUT=!QUOTE(!Bwsav)/UNC.

```
*****
* 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 two single quotes (%LET !Classes = ' ') *
* - If more than one variable, leave a space between each variable ( %LET !Classes = ' var1 var2 ' ) *
* - DO NOT ERASE OR COMMENT OUT THIS COMMAND *
*****
```

!LET !Classes = **'breakdown_variable(s)_or_empty_(keep_the_2_single_quotes)'**

Appendix A

```
*****
* 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_V21.SPS PROGRAM MUST THEN BE RUN.
*
*   - FOR TESTING, B MUST BE >= 2
*   - REFER TO APPENDIX C IN THE DOCUMENTATION TO FIND THE NUMBER OF WEIGHTS
*   THAT THE WEIGHTS FILE CONTAINS
*****
```

!LET !B = *number_of_bootstrap_weights*

```
*****
* SPECIFY THE DIRECTORY AND THE NAME OF THE FILE THAT CONTAINS THE MACROS *
* (THE PROGRAM MACROE_V21.SAS IF NO MODIFICATIONS HAVE BEEN MADE BY THE USER) *
*****
```

INCLUDE FILE=*directory_name_of_macroe_v21.sps*MACROE_V21.SPS'.

```
* DO NOT MODIFY THE NEXT LINE. (GO TO SECTION 2)
*****
!Prepare B=!B /Classes=!Classes /mfile=!Mfile /bwsav=!bwsav.
```

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

...
...
...

```
* TO OBTAIN VARIANCE ESTIMATES OF A TOTAL, RUN:
*****
```

```
*!Total !B !Classes/ Var = name_of_the_variable .
*!Let !Tot=1.
```

```
* TO OBTAIN VARIANCE ESTIMATES OF A RATIO, RUN:
*****
```

```
*!Ratio !B !Classes/ Varlist = name_of_numerator name_of_denominator .
*!Let !Tot=1.
```

```
* TO OBTAIN VARIANCE ESTIMATES OF A DIFFERENCE OF RATIO, RUN:
*****
****   NOTE: see the comment at the beginning of section 2 ...   ****
```

```
*!Diffrat !B !Classes/ Varlist = variable1 variable2 variable3 variable4 .
*!Let !Tot=1.
```

```

* where: variable1 : the numerator variable of the first ratio
*         variable2 : the denominator variable of the first ratio
*         variable3 : the numerator variable of the second ratio
*         variable4 : the denominator variable of the second ratio
*
```

```
* TO OBTAIN VARIANCE ESTIMATES OF REGRESSION PARAMETERS, RUN:
*****
```

```
*!Regress !B !Classes/ Dep = dependant_variable Indep = independent_variable(s)
*!Let !Reg=1.
```

Appendix A

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

*!Log_Reg !B !Classes/ Cri=SAS Dep = dependant_variable Indep = Independent_variable(s)
*!Let !Reg=1.

    * NOTE: Default iteration criteria differs between SAS and SPSS.
    *       Use Cri=SAS to use SAS default values.
    *       Indep= must be specified LAST because of the unknown number of independent variables.

*****
*       DO NOT MODIFY THE REMAINING LINES       *
*****

SET PRI NO MPR NO.

**** Printing of results ****.
!!F (!Tot=1) !THEN
+ !Print_T !Classes. /* Printing of Totals, Ratios and Differences of Ratios results
!!FEND
!!F (!Reg=1) !THEN
+ !Print_R !Classes. /* Printing of Linear and Logistic Regressions results
!!FEND
*****

!Stop !NBlocks.
!ENDDEFINE /* !Bootvar.

SET MPR ON.
!Bootvar.
SET MPR NO.

*** END OF BOOTVARE_V21.SPS PROGRAM ***.
```

APPENDIX B

This is a complete example showing how to use the program BOOTVARE_V21.SPS. First, the analysis data file is created (step 1). Then, BOOTVARE_V21.SPS 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 population and the men, 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.SPS                     *
*                               *
* This program creates the SPSS datafile containing the necessary variables for the *
* BOOTVARE_V21.SPS program.      *
*                               *
* IT IS NOT NECESSARY TO USE THIS PROGRAM. Dichotomous variables can *
* be computed directly in the Data Editor window and the analysis data file can be *
* created from there.           *
*                               *
*****

SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO.

* Creation of the SPSS "sav" 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.

DEFINE !Go()
!LET !Layout = 'D:\LAYOUT\h35_i.sps'
!LET !DataIn = 'D:\DATA\h35.txt'
!LET !DataOut = 'C:\BOOTVAR\diabetes.sav' /* <- (analysis file to be used in BOOTVARE_V21.SPS)

* Do not modify the following block of instructions.
FILE HANDLE INFILE/NAME = !QUOTE(!DataIn).
INCLUDE FILE=!QUOTE(!Layout).
SET PRI ON.
```

```
*****
* Creation of Dichotomous Variables ( -> REPLACE all following instructions (up to "SAVE" ) *
* (examples are presented below using NPHS cycle 3 variables) *
*****
```

Appendix B

NUMERIC diab nonproxy total males females mdiab fdiab.

IF (ccc8_1j=1) diab=1.
IF (ccc8_1j~=1) diab=0.

compute nonproxy = 0.
IF (am58_pxy>2) nonproxy=\$SYSMIS.
IF (am58_pxy=2) nonproxy=1.

compute total = 1.

IF (dhc8_sex=1) males=1.
IF (dhc8_sex~=1) males=0.
IF (dhc8_sex=2) females=1.
IF (dhc8_sex~=2) females=0.

compute mdiab = males*diab.
compute fdiab = females*diab.

SELECT IF (prc8_cur = 10 or prc8_cur = 24 or prc8_cur = 35 or prc8_cur = 59).

execute.

*****.

SAVE OUTFILE !QUOTE(!DataOut)
/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_V21.SPS. *
* IMPORTANT: the identification variables and, if necessary, *
* the breakdown variable(s) (ex: province, sex) must be kept. The *
* weight variable also must be kept if point estimates are *
* calculated at this step *

* Replace "SAVE" by "XSAVE" in previous command if point estimates are desired at this time.

*** DO NOT FORGET TO RUN THE LAST LINES OF THE PROGRAM (AFTER THE EXAMPLE) ***.

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

*** Examples: ***.

SORT CASES BY prc8_cur.
SPLIT FILE BY prc8_cur.

WEIGHT BY wt58.

FREQ VAR=diab mdiab fdiab.
LOGISTIC REG VAR=diab WITH nonproxy females /PRI SUM
/CRI BCO(0.0001) LCO(0) ITE(25) EPS(10E-12).

WEIGHT OFF.
SPLIT FILE OFF.

* DO NOT DELETE the following instructions
*****.

SET PRI NO.
!ENDDDEFINE /* !Go.

SET MPR ON.
!Go.
SET MPR NO.

Step 2 - BOOTVARE V21.SPS program:

```

. . .
. . .
. . .

```

```

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

```

```
SET COM NO HEA ON MES ON ERR ON RES ON JOU ON LEN NONE WID 132 PRI NO MPR NO.
```

```
DEFINE !Bootvar()
```

```
!Let !Tot=!Null
!Let !Reg=!Null
```

```

*****
* SPECIFY THE ANALYSIS DATA FILE TO BE USED (created in step 1) (Folder and file): *
*****

```

```
!LET !Mfile='C:\BOOTVAR\diabetes.sav' /* <- (ex: c:\data\analysis.sav).
```

```

*****
* SPECIFY THE BOOTSTRAP WEIGHTS FOLDER AND FILE NAMES : *
* 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 SPSS FORMAT (remove the " * ") *
*****

```

```
!LET !BWsav='D:\Bootstrp\data\b5h35.sav'
```

```

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

```

```

*!LET !BWTXT=' Folder_and_file_names_of_the_bootstrap_weights_in_text_format_(with_extension) '
*!LET !Bwlay=' Folder_and_file_names_of_the_bootstrap_layout_file_in_SPSS_syntax_(with_extension) '
*!LET !Bwsav=' Folder_and_file_names_of_the_bootstrap_weights_in_SPSS_format_(created_in_this_section)_with_extension '

```

```

*FILE HANDLE BOOTLAY/NAME = !QUOTE(!BWTXT).
*INCLUDE FILE=!QUOTE(!Bwlay).
*SAVE OUT=!QUOTE(!BWsav) /UNC.

```

```

*****
* 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 two single quotes (%LET !Classes = ' ' ) *
* - If more than one variable, leave a space between each variable ( %LET !Classes = ' var1 var2 ' ) *
* - DO NOT ERASE OR COMMENT OUT THIS COMMAND *
*****

```

```
!LET !Classes = 'prc8_cur'
```

Appendix B

```
*****
* 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_V21.SPS PROGRAM MUST THEN BE RUN.        *
*                                                                                     *
*   - FOR TESTING, B MUST BE >= 2                                                  *
*   - REFER TO APPENDIX C IN THE DOCUMENTATION TO FIND THE NUMBER OF WEIGHTS       *
*   THAT THE WEIGHTS FILE CONTAINS                                                 *
*****

!LET !B = 500

*****
* SPECIFY THE DIRECTORY AND THE NAME OF THE FILE THAT CONTAINS THE MACROS          *
* (THE PROGRAM MACROE_V21.SAS IF NO MODIFICATIONS HAVE BEEN MADE BY THE USER)    *
*****

INCLUDE FILE='C:\BOOTVAR\MACROE_V21.SPS'.

* DO NOT MODIFY THE NEXT LINE. (GO TO SECTION 2)
*****
!Prepare B=!B /Classes=!Classes /mfile=!Mfile /bwsav=!bwsav.

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

...
...
...

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

*!Total !B !Classes/ Var = name_of_the_variable .
*!Let !Tot=1.

!Total !B !Classes/ Var = diab.
!Total !B !Classes/ Var = mdiab.
!Let !Tot=1.

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

*!Ratio !B !Classes/ Varlist = name_of_numerator  name_of_denominator .
*!Let !Tot=1.

!Ratio !B !Classes/ Varlist = diab total.
!Ratio !B !Classes/ Varlist = mdiab males.
!Let !Tot=1.

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

*!Diffrrat !B !Classes/ Varlist = variable1 variable2 variable3 variable4 .
*!Let !Tot=1.

* where: variable1 : the numerator variable of the first ratio
*         variable2 : the denominator variable of the first ratio
*         variable3 : the numerator variable of the second ratio
*         variable4 : the denominator variable of the second ratio *
```

Appendix B

* TO OBTAIN VARIANCE ESTIMATES OF REGRESSION PARAMETERS, RUN:

```
*!Regress !B !Classes/ Dep = dependant_variable Indep = Independent_variable(s)
*!Let !Reg=1.
```

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

```
*!Log_Reg !B !Classes/ Cri=SAS Dep = dependant_variable Indep = Independent_variable(s)
*!Let !Reg=1.
```

```
!Log_Reg !B !Classes/ Cri=SAS Dep = diab Indep = nonproxy females
!Let !Reg=1.
```

* NOTE: Default iteration criteria differs between SAS and SPSS.

* Use Cri=SAS to use SAS default values.

* Indep= must be specified LAST because of the unknown number of independent variables.

```
* DO NOT MODIFY THE REMAINING LINES *
```

SET PRI NO MPR NO.

**** Printing of results ****.

```
!!F (!Tot=1) !THEN
```

```
+ !Print_T !Classes. /* Printing of Totals, Ratios and Differences of Ratios results
```

```
!!FEND
```

```
!!F (!Reg=1) !THEN
```

```
+ !Print_R !Classes. /* Printing of Linear and Logistic Regressions results
```

```
!!FEND
```

```
!Stop !NBlocks.
```

```
!ENDDFINE /* !Bootvar.
```

SET MPR ON.

```
!Bootvar.
```

SET MPR NO.

*** END OF BOOTVARE_V21.SPS 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 the twelfth line. 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* as the numerator of the ratio and *males* as the denominator (VARIABLES). 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 female in Ontario (ninth line) 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 7.9964 (WALD) and $p=0.004687$ (PVALUE) respectively. The estimates of the standard deviation for the parameter estimate is 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

21 Mar 02 Bootstrap [500 weights] Variance Estimate

PRC8_CUR	TYPE.	VARIABLE	YHAT	BS_SD	BS_CV	CIL95	CIU95
10	Total	DIAB	20741.32	1778.61	8.58	17255.24	24227.40
10	Total	MDIAB	7029.13	1356.31	19.30	4370.75	9687.51
10	Ratio	DIAB TOTAL	3.85	.33	8.58	3.21	4.50
10	Ratio	MDIAB MALES	2.63	.51	19.30	1.64	3.63
24	Total	DIAB	205292.27	16330.49	7.95	173284.50	237300.04
24	Total	MDIAB	110452.80	10818.85	9.79	89247.85	131657.75
24	Ratio	DIAB TOTAL	2.87	.23	7.95	2.42	3.32
24	Ratio	MDIAB MALES	3.12	.31	9.80	2.52	3.72
35	Total	DIAB	362439.60	20692.51	5.71	321882.28	402996.92
35	Total	MDIAB	198237.67	16369.18	8.26	166154.08	230321.26
35	Ratio	DIAB TOTAL	3.22	.18	5.71	2.86	3.58
35	Ratio	MDIAB MALES	3.57	.29	8.26	2.99	4.14
59	Total	DIAB	110375.39	10661.17	9.66	89479.50	131271.28
59	Total	MDIAB	62808.64	8301.60	13.22	46537.50	79079.78
59	Ratio	DIAB TOTAL	2.83	.27	9.66	2.29	3.37
59	Ratio	MDIAB MALES	3.24	.43	13.22	2.40	4.08

21 Mar 02 Bootstrap [500 weights] Variance Estimate

PRC8_CUR	TYPE	BETA	BHAT	ODDS	WALD	PVALUE	BS_SD	BS_CV	CIL95	CIU95
10	Logistic Regression	Const_	-4.00379	.01825	255.3510	.00000000	.25055	6.26	.01117	.02982
10	Logistic Regression	NONPROXY	.85788	2.35815	11.8177	.00058671	.24955	29.09	1.44594	3.84587
10	Logistic Regression	FEMALES	.46627	1.59403	2.8588	.09087612	.27577	59.14	.92845	2.73674
24	Logistic Regression	Const_	-3.89894	.02026	852.2376	.00000000	.13356	3.43	.01560	.02633
24	Logistic Regression	NONPROXY	.89123	2.43812	34.0970	.00000001	.15263	17.13	1.80774	3.28831
24	Logistic Regression	FEMALES	-.36991	.69080	5.7831	.01618064	.15382	41.58	.51100	.93387
35	Logistic Regression	Const_	-3.57575	.02799	961.1974	.00000000	.11533	3.23	.02233	.03510
35	Logistic Regression	NONPROXY	.60946	1.83943	23.5198	.00000124	.12567	20.62	1.43785	2.35318
35	Logistic Regression	FEMALES	-.34507	.70817	7.9964	.00468716	.12203	35.36	.55753	.89952
59	Logistic Regression	Const_	-3.99966	.01832	301.8469	.00000000	.23021	5.76	.01167	.02877
59	Logistic Regression	NONPROXY	1.05766	2.87963	18.8915	.00001384	.24334	23.01	1.78732	4.63951
59	Logistic Regression	FEMALES	-.46337	.62916	4.5912	.03213705	.21625	46.67	.41180	.96126

```

-----
Actual Time:                21-MAR-2002 14:35:20.700
Starting Time:              21-MAR-2002 13:45:22.048
Total Elapsed Time to run Program:  0:49:58.652
-----

```

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	LN6F.txt	B5LONGF	REALUKEY PERSONID	WT66LF (M) WT66SLF (S)	500
Cycle 2 - Partial (Master only)	LN6P.txt	B5LONGP	REALUKEY PERSONID	WT66LP (M)	500

(M) - Master Files and Public Use Micro-Data 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	B5LNGF	REALUKEY PERSONID	WT60LF (M) WT60SLF (S)	500
Cycle 4 - Square (Master only)	LONG.txt	B5LONG	REALUKEY PERSONID	WT60SLF (M)	500
Cycle 4 - Full Cycle1 and 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 and Public Use Micro-Data Files

(S) - Share Files