

# WTADJUST Example #2

## *SUDAAN Statements and Results Illustrated*

- REFLEVEL
- WTMAX
- SETENV
- UPPERBD
- LOWERBD

## *Input Data Set(s):*

### *Example*

*Using data from the 2006 National Health Interview Survey (NHIS) Sample Adult Public Use File, assume records assigned to Weeks 9-13 within each quarter of data collection are nonrespondents and compute an appropriate nonresponse adjustment for those records associated with Weeks 1-8.*

*Evaluate the potential relative bias implications of omitting data collected in Weeks 9-13 of each quarter on the following three outcome measures:*

- o *Told by a doctor that you have cancer,*
- o *Current heavier drinker and*
- o *Chances of Getting HIV/AIDs = High, Moderate; Already Have AIDS*

### *Solution*

Many ongoing surveys face the issue of whether to cut off data collection after a certain point in time or after a certain number of call-backs. If one were to allow for a longer period of data collection or a greater number of call-backs, variance and bias might be reduced due to the larger resulting respondent sample, though the data collection costs will also increase. Cost and variance implications of a smaller sample can generally be estimated fairly easily by looking at historical data. To measure bias implications, one method is to examine the potential relative bias, (i.e., use historical data and derive estimates computed using only those respondents that participated early in the data collection effort, or after a fewer number of call-backs) and compare these estimates against those estimates generated from the full sample. This example illustrates the process of estimating the relative bias.

Households are assigned to the NHIS interviewers on a weekly basis, so the weekly samples are independent of one another. Therefore, there is no reason to assume that there would be any bias implications by stopping data collection within each quarter at Week 8. This analysis is being conducted with the NHIS data for illustration purposes only. All output in this example was generated using the SAS-callable version of SUDAAN.

*Exhibit 1* presents an initial DATA statement that was used to read in the NHIS public use file. In summary, the following variables are created/kept:

<u>Variable</u>	<u>Definition</u>
RECID	Observation number
STRAT_P	Pseudo-Strata indicator for variance estimation

PSU_P	Pseudo-Primary Sampling Unit (PSU) indicator for variance estimation
WTFA_SA	Final Weight
ASSIGNWK	Assignment week (1-13)
RESP	0/1 respondent indicator. In this example, those cases worked in ASSIGNWK=1,...,8 are considered respondents and those records worked in ASSIGNWK=9,...,13 are considered nonrespondents.
INTV_QRT	Interview quarter (1-4)
SEX	Gender
REGION	Census Region
AGE_P	Age category
HISP	Hispanic indicator
RACE	Race indicator
CANCER	0/1 indicator for whether the respondent was ever told by a doctor that they had cancer.
DRINKER	0/1 indicator for whether the person is considered a current, heavier drinker
HIV	0/1 indicator for whether a person's responses indicated their chances of getting HIV/AIDs was high, moderate or whether they already have AIDs.

## Exhibit 1. SAS Code to Read In NHIS Data

```
proc format;
value hisp 1="1 Hispanic"
           2="2 Nonhispanic";

value race 1="1 White"
           2="2 Black/African American"
           3="3 Other";

run;

data one(keep=recid strat_p psu_p wtfa_sa
          assignwk resp
          intv_qrt sex region age_p hisp race cancer drinker hiv);
set indsn.samadult;

recid=_n_;

format hisp hisp.;
if 00 le hispan_i le 11 then hisp=1;
else hisp=2;

format race race.;
if mracrpi2 in (1,2) then race=mracrpi2;
else race=3;

if 1 le assignwk le 8 then resp=1;
else resp=0;

/*Ever told by doctor you had cancer = Yes*/
if canev=1 then cancer=1;
else if canev=2 then cancer=0;

/*Current, heavier drinker*/
if alcstat in (8) then drinker=1;
else if alcstat in (1,2,3,4,5,6,7,8) then drinker=0;

/*Chances of Getting HIV/AIDs = High, Moderate, Already Have AIDS*/
if chnsadsp in (1,2) then hiv=1;
else if chnsadsp in (1,2,3,4) then hiv=0;

run;
```

We assume that this is a with-replacement design, and therefore, the appropriate SUDAAN design statements are:

```
PROC _____ DESIGN=WR;
WEIGHT wtfa_sa;
NEST strat_p psu_p;
```

Given the above assumptions, it is more efficient for SUDAAN if the file was sorted by the NEST statement variables before running any SUDAAN procedures. In this example, we will use the NOTSORTED option on the PROC statement with the first SUDAAN procedure to indicate that the dataset is not sorted. We will also use the OUTDATA option on the PROC statement to save a copy of the sorted file for subsequent SUDAAN runs.

In this example, we will use the main effects of quarter, region, sex, race, Hispanic indicator and age as explanatory variables in the weight adjustment model. All variables will be considered categorical except for the age variable.

Following the recommended steps outlined in *Exhibit 1*, we first compute the weighted sums to which the final nonresponse adjusted, weighted sample should sum. Since we are using both categorical and continuous variables in the model, we use both SUDAAN's CROSSTAB and DESCRIPT to get these weighted sums. *Exhibit 2* shows the code used to derive these initial totals.

## Exhibit 2. SAS Code Used to Compute Initial Control Totals

```
PROC CROSSTAB DATA=one DESIGN=WR USEVMEM=1 NOTSORTED OUTDATA=onesort;
WEIGHT wtfa_sa;
NEST strat_p psu_p;
CLASS intv_qrt region sex race hisp resp;
TABLES (intv_qrt region sex race hisp)*resp;
RLABEL race="Race of Respondent"
        hisp="Hispanic Indicator"
        resp="0/1 Respondent Indicator";
SETENV COLWIDTH=15;
PRINT NSUM WSUM ROWPER /
      FILETYPE=RTF FILENAME="Example_12_2_OUT1.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE "Example 15-2";
run;

PROC DESCRIPT DATA=onesort DESIGN=WR USEVMEM=1;
WEIGHT wtfa_sa;
NEST strat_p psu_p;
CLASS resp;
TABLES resp;
VAR age_p;
RLABEL resp="0/1 Respondent Indicator";
SETENV COLWIDTH=15;
PRINT NSUM TOTAL /
      FILETYPE=RTF FILENAME="Example_12_2_OUT2.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE "Example 15-2";
run;
```

In this example, we also the USEVMEM=1 option on the PROC statement to tell SUDAAN to always used the external drive to store information during processing. This will minimize the amount of internal memory (i.e., RAM) needed by the SUDAAN job.

The output from these procedures is presented in *Exhibit 3* and *Exhibit 4*.

**Exhibit 3. Example\_12\_2\_OUT1.rtf**

Date: 05-29-2008

SUDAAN

Page: 1

Time: 09:50:45

Table:1

Variance Estimation Method: Taylor Series (WR)

Example 15-2

by: Interview Quarter, 0/1 Respondent Indicator.

Interview Quarter		0/1 Respondent Indicator		
		Total	0	1
Total	Sample Size	24275	10021	14254
	Weighted Size	220266693.00	94730690.00	125536003.00
	Row Percent	100.00	43.01	56.99
1 Quarter 1	Sample Size	6155	2977	3178
	Weighted Size	54831586.00	27009735.00	27821851.00
	Row Percent	100.00	49.26	50.74
2 Quarter 2	Sample Size	7180	2885	4295
	Weighted Size	54980247.00	23352243.00	31628004.00
	Row Percent	100.00	42.47	57.53
3 Quarter 3	Sample Size	3705	1354	2351
	Weighted Size	55148712.00	21426670.00	33722042.00
	Row Percent	100.00	38.85	61.15
4 Quarter 4	Sample Size	7235	2805	4430
	Weighted Size	55306148.00	22942042.00	32364106.00
	Row Percent	100.00	41.48	58.52

**Exhibit 3. Example\_12\_2\_OUT1.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:45

SUDAAN

Page: 2  
Table:2

Variance Estimation Method: Taylor Series (WR)  
Example 15-2  
by: Region, 0/1 Respondent Indicator.

Region		0/1 Respondent Indicator		
		Total	0	1
Total	Sample Size	24275	10021	14254
	Weighted Size	220266693.00	94730690.00	125536003.00
	Row Percent	100.00	43.01	56.99
1 Northeast	Sample Size	4151	1588	2563
	Weighted Size	39033183.00	16664752.00	22368431.00
	Row Percent	100.00	42.69	57.31
2 Midwest	Sample Size	5303	2251	3052
	Weighted Size	51565416.00	22664919.00	28900497.00
	Row Percent	100.00	43.95	56.05
3 South	Sample Size	9254	3912	5342
	Weighted Size	83510919.00	35799776.00	47711143.00
	Row Percent	100.00	42.87	57.13
4 West	Sample Size	5567	2270	3297
	Weighted Size	46157175.00	19601243.00	26555932.00
	Row Percent	100.00	42.47	57.53

**Exhibit 3. Example\_12\_2\_OUT1.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 3	
Time: 09:50:45				Table:3	
Variance Estimation Method: Taylor Series (WR)					
Example 15-2					
by: Sex, 0/1 Respondent Indicator.					
Sex		0/1 Respondent Indicator			
		Total	0	1	
Total	Sample Size	24275	10021	14254	
	Weighted Size	220266693.00	94730690.00	125536003.00	
	Row Percent	100.00	43.01	56.99	
1 Male	Sample Size	10715	4453	6262	
	Weighted Size	106252219.00	45794627.00	60457592.00	
	Row Percent	100.00	43.10	56.90	
2 Female	Sample Size	13560	5568	7992	
	Weighted Size	114014474.00	48936063.00	65078411.00	
	Row Percent	100.00	42.92	57.08	

**Exhibit 3. Example\_12\_2\_OUT1.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 4	
Time: 09:50:45				Table:4	
Variance Estimation Method: Taylor Series (WR)					
Example 15-2					
by: Race of Respondent, 0/1 Respondent Indicator.					
Race of Respondent		0/1 Respondent Indicator			
		Total	0	1	
Total	Sample Size	24275	10021	14254	
	Weighted Size	220266693.00	94730690.00	125536003.00	
	Row Percent	100.00	43.01	56.99	
1 White	Sample Size	18422	7868	10554	
	Weighted Size	180681666.00	79125115.00	101556551.00	
	Row Percent	100.00	43.79	56.21	
2 Black/African American	Sample Size	4182	1430	2752	
	Weighted Size	26696785.00	9791517.00	16905268.00	
	Row Percent	100.00	36.68	63.32	
3 Other	Sample Size	1671	723	948	
	Weighted Size	12888242.00	5814058.00	7074184.00	
	Row Percent	100.00	45.11	54.89	

**Exhibit 3. Example\_12\_2\_OUT1.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 5
Time: 09:50:45				Table:5
Variance Estimation Method: Taylor Series (WR)				
Example 15-2				
by: Hispanic Indicator, 0/1 Respondent Indicator.				
Hispanic Indicator		0/1 Respondent Indicator		
		Total	0	1
Total	Sample Size	24275	10021	14254
	Weighted Size	220266693.00	94730690.00	125536003.00
	Row Percent	100.00	43.01	56.99
1 Hispanic	Sample Size	4227	1497	2730
	Weighted Size	28663833.00	10613779.00	18050054.00
	Row Percent	100.00	37.03	62.97
2 Nonhispanic	Sample Size	20048	8524	11524
	Weighted Size	191602860.00	84116911.00	107485949.00
	Row Percent	100.00	43.90	56.10

The row percents, under the Respondent Indicator = 1 column in *Exhibit 3*, display the weighted response rates. Looking at the second page of *Exhibit 3* (for example), the table indicates that we have 2,563 respondents located in the Northeast Census region, and their weights sum to 22,368,431. Additionally, there are 4,151 records in the entire sample in the northeast, and their weights sum to 39,033,183. We seek to create a nonresponse adjustment that will force the weighted sum of the 2,563 respondent records to equal 39,033,183.

**Exhibit 4. Example\_12\_2\_OUT2.rtf**

Date: 05-29-2008		SUDAAN		Page: 1
Time: 09:50:47				Table:1
Variance Estimation Method: Taylor Series (WR)				
Example 15-2				
by: Variable, 0/1 Respondent Indicator.				
Variable		0/1 Respondent Indicator		
		Total	0	1
Age	Sample Size	24275	10021	14254
	Total	10061218601.00	4320097971.00	5741120630.00

The “Total” row in *Exhibit 4* provides the desired weighted sum from the PROC DESCRIPT. This row indicates that the weighted sum of the age variable over all respondent records is 5,741,120,630. We seek to create a nonresponse adjustment that will force this weighted sum to equal 10,061,218,601.



Again, following the steps presented in *Exhibit 1*, we first run a PROC WTADJUST with no parameters set on the WTMIN, WTMAX, LOWERBD, CENTER or UPPERBD statements (Step 2 in *Exhibit 1*). This code is presented in *Exhibit 5*, and the resulting output is presented in *Exhibit 6*.

### Exhibit 5. Initial WTADJUST Procedure

```

PROC WTADJUST DATA=onesort DESIGN=WR ADJUST=NONRESPONSE USEVMEM=1;
WEIGHT wtfa_sa;
NEST strat_p psu_p;
REFLEVEL intv_qrt=1 region=1;
CLASS intv_qrt region sex race hisp;
MODEL resp=intv_qrt region sex race hisp age_p;
RLABEL race="Race of Respondent"
        hisp="Hispanic Indicator"
        resp="0/1 Respondent Indicator";
SETENV COLWIDTH=15;
PRINT BETA SEBETA P_BETA / BETAfmt=F10.3 SEBETAfmt=F10.3
      FILETYPE=RTF FILENAME="Example_12_2_OUT3.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT UNWTD RR WTD RR /
      FILETYPE=RTF FILENAME="Example_12_2_OUT3.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT INITWTM N INITWTM NTRIMMED /
      FILETYPE=RTF FILENAME="Example_12_2_OUT3.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT MARGADJ ADJMIN ADJMAX /
      FILETYPE=RTF FILENAME="Example_12_2_OUT3.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT UWEORIG UWETRIM UWEFINAL /
      FILETYPE=RTF FILENAME="Example_12_2_OUT3.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE "Example 15-2";
run;

```

Similar to the example presented in *Section 24.13.1*, notice that several PRINT statements are included in this run of WTADJUST. These print statements specify output to the same EXAMPLE\_12\_2\_OUT3.rtf file. When multiple PRINT statements like this appear in the same SUDAAN procedure, SUDAAN appends these tables to one another. This is an effective way to control which columns appear together on the same page of output.

In the code above, we also specify an alternate label for the variables RACE, HISP and RESP. This is done with the RLABEL statement.

Finally, for illustration purposes, we use the REFLEVEL statement to change the reference level associated with the categorical variables INTV\_QRT and REGION. Ordinarily, the last level is assumed to be the reference level (in this case, “4” for both variables). However, the REFLEVEL statement tells SUDAAN to make the first level the reference level.

**Exhibit 6. Example\_12\_2\_OUT3.rtf**

Date: 05-29-2008

SUDAAN

Page: 1

Time: 09:50:47

Table: 1

Variance Estimation Method: Taylor Series (WR)  
 Response variable RESP: 0/1 Respondent Indicator  
 Nonresponse Adjustment  
 Example 15-2  
 by: Independent Variables and Effects.

Independent Variables and Effects		Beta Coeff.	SE Beta	P-value T-Test B=0
Intercept		0.319	0.122	0.0094
Interview Quarter	1 Quarter 1	0.000	0.000	.
	2 Quarter 2	-0.219	0.083	0.0087
	3 Quarter 3	-0.326	0.097	0.0009
	4 Quarter 4	-0.248	0.090	0.0060
Region	1 Northeast	0.000	0.000	.
	2 Midwest	0.017	0.103	0.8656
	3 South	0.020	0.095	0.8308
	4 West	0.009	0.097	0.9293
Sex	1 Male	0.006	0.023	0.8046
	2 Female	0.000	0.000	.
Race of Respondent	1 White	-0.020	0.071	0.7755
	2 Black/African American	-0.265	0.080	0.0010
	3 Other	0.000	0.000	.
Hispanic Indicator	1 Hispanic	-0.229	0.053	0.0000
	2 Nonhispanic	0.000	0.000	.
Age		-0.001	0.001	0.2054

**Exhibit 6. Example\_12\_2\_OUT3.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:47

SUDAAN

Page: 2  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

<b>Independent Variables and Effects</b>		<b>Unweighted Response Rate (Percent)</b>	<b>Weighted Response Rate (Percent)</b>
Intercept		58.72	56.99
Interview Quarter	1 Quarter 1	51.63	50.74
	2 Quarter 2	59.82	57.53
	3 Quarter 3	63.45	61.15
	4 Quarter 4	61.23	58.52
Region	1 Northeast	61.74	57.31
	2 Midwest	57.55	56.05
	3 South	57.73	57.13
	4 West	59.22	57.53
Sex	1 Male	58.44	56.90
	2 Female	58.94	57.08
Race of Respondent	1 White	57.29	56.21
	2 Black/African American	65.81	63.32
	3 Other	56.73	54.89
Hispanic Indicator	1 Hispanic	64.58	62.97
	2 Nonhispanic	57.48	56.10
Age		.	.

**Exhibit 6. Example\_12\_2\_OUT3.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:47

SUDAAN

Page: 3  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

Independent Variables and Effects		Minimum Initial Weight Among Respondents	Maximum Initial Weight Among Respondents	Number of Trimmed Weights
Intercept		1294.00	98166.00	0
Interview Quarter	1 Quarter 1	1697.00	64708.00	0
	2 Quarter 2	1318.00	36179.00	0
	3 Quarter 3	2435.00	98166.00	0
	4 Quarter 4	1294.00	60560.00	0
Region	1 Northeast	1318.00	61357.00	0
	2 Midwest	1381.00	98166.00	0
	3 South	1374.00	69634.00	0
	4 West	1294.00	73837.00	0
Sex	1 Male	1425.00	98166.00	0
	2 Female	1294.00	69634.00	0
Race of Respondent	1 White	1400.00	98166.00	0
	2 Black/African American	1318.00	60388.00	0
	3 Other	1294.00	50798.00	0
Hispanic Indicator	1 Hispanic	1400.00	56616.00	0
	2 Nonhispanic	1294.00	98166.00	0
Age		.	.	.

**Exhibit 6. Example\_12\_2\_OUT3.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:47

SUDAAN

Page: 4  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

Independent Variables and Effects		Marginal Weight Adjustment	Minimum Adjustment Factor Among Respondents	Maximum Adjustment Factor Among Respondents
Intercept		1.7546	1.2672	2.1072
Interview Quarter	1 Quarter 1	1.9708	1.5109	2.1072
	2 Quarter 2	1.7383	1.3386	1.8588
	3 Quarter 3	1.6354	1.2672	1.7519
	4 Quarter 4	1.7089	1.3083	1.8329
Region	1 Northeast	1.7450	1.2672	2.0842
	2 Midwest	1.7842	1.3027	2.1064
	3 South	1.7503	1.2958	2.1072
	4 West	1.7381	1.2986	2.0932
Sex	1 Male	1.7575	1.2704	2.1072
	2 Female	1.7520	1.2672	2.1005
Race of Respondent	1 White	1.7791	1.4370	2.0871
	2 Black/African American	1.5792	1.2672	1.8152
	3 Other	1.8219	1.4681	2.1072
Hispanic Indicator	1 Hispanic	1.5880	1.2672	1.8452
	2 Nonhispanic	1.7826	1.4236	2.1072
Age		1.7525	.	.

**Exhibit 6. Example\_12\_2\_OUT3.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 5
Time: 09:50:47				Table: 1
Variance Estimation Method: Taylor Series (WR)				
Response variable RESP: 0/1 Respondent Indicator				
Nonresponse Adjustment				
Example 15-2				
by: Independent Variables and Effects.				
Independent Variables and Effects		Original Unequal Weighting Effect	Trimmed Unequal Weighting Effect	Final Unequal Weighting Effect
Intercept		1.5093	1.5093	1.5251
Interview Quarter	1 Quarter 1	1.3984	1.3984	1.4308
	2 Quarter 2	1.3590	1.3590	1.3895
	3 Quarter 3	1.4120	1.4120	1.4513
	4 Quarter 4	1.3897	1.3897	1.4222
Region	1 Northeast	1.4417	1.4417	1.4622
	2 Midwest	1.4500	1.4500	1.4466
	3 South	1.5724	1.5724	1.6028
	4 West	1.4978	1.4978	1.4997
Sex	1 Male	1.4849	1.4849	1.4892
	2 Female	1.5134	1.5134	1.5408
Race of Respondent	1 White	1.4580	1.4580	1.4614
	2 Black/African American	1.5384	1.5384	1.5236
	3 Other	1.5928	1.5928	1.5749
Hispanic Indicator	1 Hispanic	1.5319	1.5319	1.5264
	2 Nonhispanic	1.4821	1.4821	1.4853
Age		.	.	.

Page 3 of *Exhibit 6* indicates that the original weights range from 1,294 to 98,166, and page 4 indicates that the initial nonresponse adjustment ranges from 1.2672 to 2.1072.

In some applications, it may be desirable at this stage to conduct a backward elimination process and remove nonsignificant explanatory variables from the model. One may also want to check the significance of interaction terms. For this example, we will retain this set of main effects variables.

Moving next to Step 4 of *Exhibit 1*, we determine an appropriate weight truncation factor. In other words, we select appropriate values for WTMIN and WTMAX.

There are several methods one can use to establish appropriate weight truncation bounds. *Section 15.3.1* of the **SUDAAN 10 Language Manual** presents a discussion of how one might want to examine either the standard deviation or the interquartile range (IQR) of the sample weights. An initial value of WTMIN and WTMAX could set as follows:

$$WTMIN = WT_{Median} - 3 \cdot WT_{IQR}$$

$$WTMAX = WT_{Median} + 3 \cdot WT_{IQR}$$

If desired, the medians and IQR's can be computed within groups.

For illustration purposes in this example, we will set WTMAX at 50,000 for all records and not specify a value for WTMIN. The associated code for the next run of WTADJUST is presented in *Exhibit 7*, and the output is presented in *Exhibit 8*.

### Exhibit 7. WTADJUST Procedure

```
PROC WTADJUST DATA=onesort DESIGN=WR ADJUST=NONRESPONSE USEVMEM=1;
WEIGHT  wtfa_sa;
NEST    strat_p psu_p;
REFLEVEL intv_qrt=1 region=1;
CLASS   intv_qrt region sex race hisp;
WTMAX   50000;
MODEL   resp=intv_qrt region sex race hisp age_p;
RLABEL  race="Race of Respondent"
         hisp="Hispanic Indicator"
         resp="0/1 Respondent Indicator";
SETENV  COLWIDTH=15;
PRINT   BETA SEBETA P_BETA / BETAfmt=F10.3 SEBETAfmt=F10.3
         FILETYPE=RTF FILENAME="Example_12_2_OUT4.rtf" REPLACE
         FONTNAME="Arial" FONTSIZE=10
         TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT   UNWTD RR WTD RR /
         FILETYPE=RTF FILENAME="Example_12_2_OUT4.rtf" REPLACE
         FONTNAME="Arial" FONTSIZE=10
         TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT   INITWTMN INITWTMX NTRIMMED /
         FILETYPE=RTF FILENAME="Example_12_2_OUT4.rtf" REPLACE
         FONTNAME="Arial" FONTSIZE=10
         TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT   MARGADJ ADJMIN ADJMAX /
         FILETYPE=RTF FILENAME="Example_12_2_OUT4.rtf" REPLACE
         FONTNAME="Arial" FONTSIZE=10
         TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT   UWEORIG UWETRIM UWEFINAL /
         FILETYPE=RTF FILENAME="Example_12_2_OUT4.rtf" REPLACE
         FONTNAME="Arial" FONTSIZE=10
         TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE  "Example 15-2";
run;
```

**Exhibit 8. Example\_12\_2\_OUT4.rtf**

Date: 05-29-2008

SUDAAN

Page: 1

Time: 09:50:49

Table: 1

Variance Estimation Method: Taylor Series (WR)  
 Response variable RESP: 0/1 Respondent Indicator  
 Nonresponse Adjustment  
 Example 15-2  
 by: Independent Variables and Effects.

Independent Variables and Effects		Beta Coeff.	SE Beta	P-value T-Test B=0
Intercept		0.320	0.122	0.0091
Interview Quarter	1 Quarter 1	0.000	0.000	.
	2 Quarter 2	-0.222	0.083	0.0078
	3 Quarter 3	-0.316	0.097	0.0012
	4 Quarter 4	-0.250	0.090	0.0055
Region	1 Northeast	0.000	0.000	.
	2 Midwest	0.019	0.103	0.8520
	3 South	0.024	0.095	0.8014
	4 West	0.009	0.097	0.9225
Sex	1 Male	0.008	0.023	0.7295
	2 Female	0.000	0.000	.
Race of Respondent	1 White	-0.015	0.071	0.8372
	2 Black/African American	-0.264	0.080	0.0011
	3 Other	0.000	0.000	.
Hispanic Indicator	1 Hispanic	-0.235	0.053	0.0000
	2 Nonhispanic	0.000	0.000	.
Age		-0.001	0.001	0.1315



**Exhibit 8. Example\_12\_2\_OUT4.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:49

SUDAAN

Page: 2  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

<b>Independent Variables and Effects</b>		<b>Unweighted Response Rate (Percent)</b>	<b>Weighted Response Rate (Percent)</b>
Intercept		58.72	56.99
Interview Quarter	1 Quarter 1	51.63	50.74
	2 Quarter 2	59.82	57.53
	3 Quarter 3	63.45	61.15
	4 Quarter 4	61.23	58.52
Region	1 Northeast	61.74	57.31
	2 Midwest	57.55	56.05
	3 South	57.73	57.13
	4 West	59.22	57.53
Sex	1 Male	58.44	56.90
	2 Female	58.94	57.08
Race of Respondent	1 White	57.29	56.21
	2 Black/African American	65.81	63.32
	3 Other	56.73	54.89
Hispanic Indicator	1 Hispanic	64.58	62.97
	2 Nonhispanic	57.48	56.10
Age		.	.

**Exhibit 8. Example\_12\_2\_OUT4.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:49

SUDAAN

Page: 3  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

Independent Variables and Effects		Minimum Initial Weight Among Respondents	Maximum Initial Weight Among Respondents	Number of Trimmed Weights
Intercept		1294.00	98166.00	26
Interview Quarter	1 Quarter 1	1697.00	64708.00	5
	2 Quarter 2	1318.00	36179.00	0
	3 Quarter 3	2435.00	98166.00	20
	4 Quarter 4	1294.00	60560.00	1
Region	1 Northeast	1318.00	61357.00	2
	2 Midwest	1381.00	98166.00	4
	3 South	1374.00	69634.00	18
	4 West	1294.00	73837.00	2
Sex	1 Male	1425.00	98166.00	15
	2 Female	1294.00	69634.00	11
Race of Respondent	1 White	1400.00	98166.00	23
	2 Black/African American	1318.00	60388.00	2
	3 Other	1294.00	50798.00	1
Hispanic Indicator	1 Hispanic	1400.00	56616.00	1
	2 Nonhispanic	1294.00	98166.00	25
Age		.	.	.

**Exhibit 8. Example\_12\_2\_OUT4.rtf-cont.**

Date: 05-29-2008  
 Time: 09:50:49

SUDAAN

Page: 4  
 Table: 1

Variance Estimation Method: Taylor Series (WR)  
 Response variable RESP: 0/1 Respondent Indicator  
 Nonresponse Adjustment  
 Example 15-2  
 by: Independent Variables and Effects.

Independent Variables and Effects		Marginal Weight Adjustment	Minimum Adjustment Factor Among Respondents	Maximum Adjustment Factor Among Respondents
Intercept		1.7587	1.2655	2.1149
Interview Quarter	1 Quarter 1	1.9735	1.4978	2.1149
	2 Quarter 2	1.7383	1.3252	1.8628
	3 Quarter 3	1.6472	1.2655	1.7695
	4 Quarter 4	1.7094	1.2942	1.8384
Region	1 Northeast	1.7465	1.2655	2.0878
	2 Midwest	1.7891	1.3075	2.1124
	3 South	1.7560	1.2993	2.1149
	4 West	1.7407	1.3023	2.0975
Sex	1 Male	1.7630	1.2702	2.1149
	2 Female	1.7547	1.2655	2.1054
Race of Respondent	1 White	1.7839	1.4348	2.1022
	2 Black/African American	1.5809	1.2655	1.8237
	3 Other	1.8221	1.4633	2.1149
Hispanic Indicator	1 Hispanic	1.5886	1.2655	1.8445
	2 Nonhispanic	1.7873	1.4230	2.1149
Age		1.7553	.	.

**Exhibit 8. Example\_12\_2\_OUT4.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 5
Time: 09:50:49				Table: 1
Variance Estimation Method: Taylor Series (WR)				
Response variable RESP: 0/1 Respondent Indicator				
Nonresponse Adjustment				
Example 15-2				
by: Independent Variables and Effects.				
Independent Variables and Effects		Original Unequal Weighting Effect	Trimmed Unequal Weighting Effect	Final Unequal Weighting Effect
Intercept		1.5093	1.4850	1.5051
Interview Quarter	1 Quarter 1	1.3984	1.3852	1.4180
	2 Quarter 2	1.3590	1.3590	1.3910
	3 Quarter 3	1.4120	1.3713	1.4089
	4 Quarter 4	1.3897	1.3857	1.4194
Region	1 Northeast	1.4417	1.4334	1.4575
	2 Midwest	1.4500	1.4190	1.4220
	3 South	1.5724	1.5422	1.5757
	4 West	1.4978	1.4802	1.4868
Sex	1 Male	1.4849	1.4544	1.4641
	2 Female	1.5134	1.4965	1.5269
Race of Respondent	1 White	1.4580	1.4325	1.4402
	2 Black/African American	1.5384	1.5225	1.5128
	3 Other	1.5928	1.5916	1.5778
Hispanic Indicator	1 Hispanic	1.5319	1.5271	1.5250
	2 Nonhispanic	1.4821	1.4561	1.4641
Age		.	.	.

Page 3 of *Exhibit 8*, above, indicates that 26 weights were trimmed, and page 5 indicates that the unequal weighting effect was reduced from 1.5093 to 1.4850. Note that weight truncation is done to minimize the effect of unequal weights on the variance of estimates; the reduction in variance can be approximated by examining the reduction in the unequal weighting effect. However, truncating the weights also introduced bias into the estimates. Therefore, we recommend not truncating a large number of weights during this process.

The next step in the weight adjustment process is to determine appropriate bounds on the resulting weight adjustment (Step 6 of *Exhibit 1*). In this example, we assign LOWERBD to 1, UPPERBD to 2 and let CENTER retain its default value. The resulting code is displayed in *Exhibit 9*, and the associated output is presented in *Exhibit 10*.

## Exhibit 9. WTADJUST Procedure

```
PROC WTADJUST DATA=onesort DESIGN=WR ADJUST=NONRESPONSE USEVMEM=1;
WEIGHT wtfa_sa;
NEST strat_p psu_p;
IDVAR recid resp intv_qrt region sex race hisp age_p cancer drinker hiv;
REFLEVEL intv_qrt=1 region=1;
CLASS intv_qrt region sex race hisp;
WTMAX 50000;
UPPERBD 2.00;
LOWERBD 1.00;
MODEL resp=intv_qrt region sex race hisp age_p;
RLABEL race="Race of Respondent"
hisp="Hispanic Indicator"
resp="0/1 Respondent Indicator";
SETENV COLWIDTH=15;
PRINT BETA SEBETA P_BETA / BETAFMT=F10.3 SEBETAFMT=F10.3
FILETYPE=RTF FILENAME="Example_12_2_OUT5.rtf" REPLACE
FONTNAME="Arial" FONTSIZE=10
TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT UNWTD RR WTD RR /
FILETYPE=RTF FILENAME="Example_12_2_OUT5.rtf" REPLACE
FONTNAME="Arial" FONTSIZE=10
TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT INITWTMN INITWTMX NTRIMMED /
FILETYPE=RTF FILENAME="Example_12_2_OUT5.rtf" REPLACE
FONTNAME="Arial" FONTSIZE=10
TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT MARGADJ ADJMIN ADJMAX /
FILETYPE=RTF FILENAME="Example_12_2_OUT5.rtf" REPLACE
FONTNAME="Arial" FONTSIZE=10
TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
PRINT UWEORIG UWETRIM UWEFINAL /
FILETYPE=RTF FILENAME="Example_12_2_OUT5.rtf" REPLACE
FONTNAME="Arial" FONTSIZE=10
TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
OUTPUT / PREDICTED=ALL FILENAME=outsud FILETYPE=SAS REPLACE;
RTITLE "Example 15-2";
run;
```

Notice in this example that we added an IDVAR statement and an OUTPUT statement. This was done in order to merge the final weights onto the analysis file and to check the final adjusted weights.

**Exhibit 10. Example\_12\_2\_OUT5.rtf**

Date: 05-29-2008

SUDAAN

Page: 1

Time: 09:50:51

Table: 1

Variance Estimation Method: Taylor Series (WR)  
 Response variable RESP: 0/1 Respondent Indicator  
 Nonresponse Adjustment  
 Example 15-2  
 by: Independent Variables and Effects.

Independent Variables and Effects		Beta Coeff.	SE Beta	P-value T-Test B=0
Intercept		0.627	0.571	0.2735
Interview Quarter	1 Quarter 1	0.000	0.000	.
	2 Quarter 2	-0.491	0.543	0.3660
	3 Quarter 3	-0.577	0.544	0.2902
	4 Quarter 4	-0.519	0.549	0.3453
Region	1 Northeast	0.000	0.000	.
	2 Midwest	0.020	0.125	0.8728
	3 South	0.024	0.111	0.8306
	4 West	0.009	0.112	0.9343
Sex	1 Male	0.009	0.027	0.7348
	2 Female	0.000	0.000	.
Race of Respondent	1 White	-0.021	0.099	0.8305
	2 Black/African American	-0.270	0.108	0.0131
	3 Other	0.000	0.000	.
Hispanic Indicator	1 Hispanic	-0.234	0.065	0.0004
	2 Nonhispanic	0.000	0.000	.
Age		-0.002	0.001	0.1450

**Exhibit 10. Example\_12\_2\_OUT5.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:51

SUDAAN

Page: 2  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

<b>Independent Variables and Effects</b>		<b>Unweighted Response Rate (Percent)</b>	<b>Weighted Response Rate (Percent)</b>
Intercept		58.72	56.99
Interview Quarter	1 Quarter 1	51.63	50.74
	2 Quarter 2	59.82	57.53
	3 Quarter 3	63.45	61.15
	4 Quarter 4	61.23	58.52
Region	1 Northeast	61.74	57.31
	2 Midwest	57.55	56.05
	3 South	57.73	57.13
	4 West	59.22	57.53
Sex	1 Male	58.44	56.90
	2 Female	58.94	57.08
Race of Respondent	1 White	57.29	56.21
	2 Black/African American	65.81	63.32
	3 Other	56.73	54.89
Hispanic Indicator	1 Hispanic	64.58	62.97
	2 Nonhispanic	57.48	56.10
Age		.	.

**Exhibit 10. Example\_12\_2\_OUT5.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:51

SUDAAN

Page: 3  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

Independent Variables and Effects		Minimum Initial Weight Among Respondents	Maximum Initial Weight Among Respondents	Number of Trimmed Weights
Intercept		1294.00	98166.00	26
Interview Quarter	1 Quarter 1	1697.00	64708.00	5
	2 Quarter 2	1318.00	36179.00	0
	3 Quarter 3	2435.00	98166.00	20
	4 Quarter 4	1294.00	60560.00	1
Region	1 Northeast	1318.00	61357.00	2
	2 Midwest	1381.00	98166.00	4
	3 South	1374.00	69634.00	18
	4 West	1294.00	73837.00	2
Sex	1 Male	1425.00	98166.00	15
	2 Female	1294.00	69634.00	11
Race of Respondent	1 White	1400.00	98166.00	23
	2 Black/African American	1318.00	60388.00	2
	3 Other	1294.00	50798.00	1
Hispanic Indicator	1 Hispanic	1400.00	56616.00	1
	2 Nonhispanic	1294.00	98166.00	25
Age		.	.	.



**Exhibit 10. Example\_12\_2\_OUT5.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:51

SUDAAN

Page: 4  
Table: 1

Variance Estimation Method: Taylor Series (WR)  
Response variable RESP: 0/1 Respondent Indicator  
Nonresponse Adjustment  
Example 15-2  
by: Independent Variables and Effects.

Independent Variables and Effects		Marginal Weight Adjustment	Minimum Adjustment Factor Among Respondents	Maximum Adjustment Factor Among Respondents
Intercept		1.7587	1.1312	1.9897
Interview Quarter	1 Quarter 1	1.9735	1.7549	1.9897
	2 Quarter 2	1.7383	1.1807	1.8665
	3 Quarter 3	1.6472	1.1312	1.8070
	4 Quarter 4	1.7094	1.1480	1.8526
Region	1 Northeast	1.7465	1.1312	1.9884
	2 Midwest	1.7891	1.1782	1.9896
	3 South	1.7560	1.1665	1.9897
	4 West	1.7407	1.1727	1.9888
Sex	1 Male	1.7630	1.1359	1.9897
	2 Female	1.7547	1.1312	1.9892
Race of Respondent	1 White	1.7839	1.3268	1.9887
	2 Black/African American	1.5809	1.1312	1.9574
	3 Other	1.8221	1.3808	1.9897
Hispanic Indicator	1 Hispanic	1.5886	1.1312	1.9627
	2 Nonhispanic	1.7873	1.3083	1.9897
Age		1.7553	.	.

**Exhibit 10. Example\_12\_2\_OUT5.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 5
Time: 09:50:51				Table: 1
Variance Estimation Method: Taylor Series (WR)				
Response variable RESP: 0/1 Respondent Indicator				
Nonresponse Adjustment				
Example 15-2				
by: Independent Variables and Effects.				
Independent Variables and Effects		Original Unequal Weighting Effect	Trimmed Unequal Weighting Effect	Final Unequal Weighting Effect
Intercept		1.5093	1.4850	1.5069
Interview Quarter	1 Quarter 1	1.3984	1.3852	1.3896
	2 Quarter 2	1.3590	1.3590	1.3972
	3 Quarter 3	1.4120	1.3713	1.4258
	4 Quarter 4	1.3897	1.3857	1.4290
Region	1 Northeast	1.4417	1.4334	1.4602
	2 Midwest	1.4500	1.4190	1.4243
	3 South	1.5724	1.5422	1.5755
	4 West	1.4978	1.4802	1.4910
Sex	1 Male	1.4849	1.4544	1.4668
	2 Female	1.5134	1.4965	1.5277
Race of Respondent	1 White	1.4580	1.4325	1.4413
	2 Black/African American	1.5384	1.5225	1.5215
	3 Other	1.5928	1.5916	1.5805
Hispanic Indicator	1 Hispanic	1.5319	1.5271	1.5354
	2 Nonhispanic	1.4821	1.4561	1.4650
Age		.	.	.

Page 4 of *Exhibit 10* indicates that the final nonresponse adjustment ranged from 1.1312 to 1.9897, and page 5 indicates that the unequal weighting effect changed from 1.4850 after weight trimming, to 1.5069 after the nonresponse adjustment was applied.

The final step in the weight adjustment process is to check the sum of the nonresponse adjusted sample weights. This is done by running the same CROSSTAB and DESCRIPT as was run in *Exhibit 2*. This code is presented below in *Exhibit 11*, and the associated output is presented in *Exhibit 12* and *Exhibit 13*.

## Exhibit 11. Code to Check the Sum of the Nonresponse Adjusted Weights

```
PROC CROSSTAB DATA=outsud DESIGN=WR USEVMEM=1;
WEIGHT wtfinal;
NEST strat_p psu_p;
CLASS intv_qrt region sex race hisp resp;
TABLES (intv_qrt region sex race hisp)*resp;
RLABEL race="Race of Respondent"
        hisp="Hispanic Indicator"
        resp="0/1 Respondent Indicator";
SETENV COLWIDTH=15;
PRINT NSUM WSUM ROWPER /
      FILETYPE=RTF FILENAME="Example_12_2_OUT6.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE "Example 15-2";
run;

PROC DESCRIPT DATA=outsud DESIGN=WR USEVMEM=1;
WEIGHT wtfinal;
NEST strat_p psu_p;
CLASS resp;
TABLES resp;
VAR age_p;
RLABEL resp="0/1 Respondent Indicator";
SETENV COLWIDTH=15;
PRINT NSUM TOTAL /
      FILETYPE=RTF FILENAME="Example_12_2_OUT7.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE "Example 15-2";
run;
```

**Exhibit 12. Example\_12\_2\_OUT6.rtf**

Date: 05-29-2008

SUDAAN

Page: 1

Time: 09:50:53

Table:1

Variance Estimation Method: Taylor Series (WR)

Example 15-2

by: Interview Quarter, 0/1 Respondent Indicator.

Interview Quarter		0/1 Respondent Indicator	
		Total	1
Total	Sample Size	14254	14254
	Weighted Size	220266693.00	220266693.00
	Row Percent	100.00	100.00
1 Quarter 1	Sample Size	3178	3178
	Weighted Size	54831586.00	54831586.00
	Row Percent	100.00	100.00
2 Quarter 2	Sample Size	4295	4295
	Weighted Size	54980247.00	54980247.00
	Row Percent	100.00	100.00
3 Quarter 3	Sample Size	2351	2351
	Weighted Size	55148712.00	55148712.00
	Row Percent	100.00	100.00
4 Quarter 4	Sample Size	4430	4430
	Weighted Size	55306148.00	55306148.00
	Row Percent	100.00	100.00

**Exhibit 12. Example\_12\_2\_OUT6.rtf-cont.**

Date: 05-29-2008  
 Time: 09:50:53

SUDAAN

Page: 2  
 Table:2

Variance Estimation Method: Taylor Series (WR)  
 Example 15-2  
 by: Region, 0/1 Respondent Indicator.

Region		0/1 Respondent Indicator	
		Total	1
Total	Sample Size	14254	14254
	Weighted Size	220266693.00	220266693.00
	Row Percent	100.00	100.00
1 Northeast	Sample Size	2563	2563
	Weighted Size	39033183.00	39033183.00
	Row Percent	100.00	100.00
2 Midwest	Sample Size	3052	3052
	Weighted Size	51565416.00	51565416.00
	Row Percent	100.00	100.00
3 South	Sample Size	5342	5342
	Weighted Size	83510919.00	83510919.00
	Row Percent	100.00	100.00
4 West	Sample Size	3297	3297
	Weighted Size	46157175.00	46157175.00
	Row Percent	100.00	100.00

**Exhibit 12. Example\_12\_2\_OUT6.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 3	
Time: 09:50:53				Table:3	
Variance Estimation Method: Taylor Series (WR)					
Example 15-2					
by: Sex, 0/1 Respondent Indicator.					
Sex		0/1 Respondent Indicator			
		Total	1		
Total	Sample Size	14254	14254		
	Weighted Size	220266693.00	220266693.00		
	Row Percent	100.00	100.00		
1 Male	Sample Size	6262	6262		
	Weighted Size	106252219.00	106252219.00		
	Row Percent	100.00	100.00		
2 Female	Sample Size	7992	7992		
	Weighted Size	114014474.00	114014474.00		
	Row Percent	100.00	100.00		

**Exhibit 12. Example\_12\_2\_OUT6.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 4	
Time: 09:50:53				Table:4	
Variance Estimation Method: Taylor Series (WR)					
Example 15-2					
by: Race of Respondent, 0/1 Respondent Indicator.					
Race of Respondent		0/1 Respondent Indicator			
		Total	1		
Total	Sample Size	14254	14254		
	Weighted Size	220266693.00	220266693.00		
	Row Percent	100.00	100.00		
1 White	Sample Size	10554	10554		
	Weighted Size	180681666.00	180681666.00		
	Row Percent	100.00	100.00		
2 Black/African American	Sample Size	2752	2752		
	Weighted Size	26696785.00	26696785.00		
	Row Percent	100.00	100.00		
3 Other	Sample Size	948	948		
	Weighted Size	12888242.00	12888242.00		
	Row Percent	100.00	100.00		

**Exhibit 12. Example\_12\_2\_OUT6.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 5	
Time: 09:50:53				Table:5	
Variance Estimation Method: Taylor Series (WR)					
Example 15-2					
by: Hispanic Indicator, 0/1 Respondent Indicator.					
Hispanic Indicator		0/1 Respondent Indicator			
		Total	1		
Total	Sample Size	14254	14254		
	Weighted Size	220266693.00	220266693.00		
	Row Percent	100.00	100.00		
1 Hispanic	Sample Size	2730	2730		
	Weighted Size	28663833.00	28663833.00		
	Row Percent	100.00	100.00		
2 Nonhispanic	Sample Size	11524	11524		
	Weighted Size	191602860.00	191602860.00		
	Row Percent	100.00	100.00		

**Exhibit 13. Example\_12\_2\_OUT7.rtf**

Date: 05-29-2008		SUDAAN		Page: 1	
Time: 09:50:53				Table:1	
Variance Estimation Method: Taylor Series (WR)					
Example 15-2					
by: Variable, 0/1 Respondent Indicator.					
Variable		0/1 Respondent Indicator			
		Total	1		
Age	Sample Size	14254	14254		
	Total	10061218601.00	10061218601.00		

As desired, the newly weighted totals displayed in *Exhibit 12* and *Exhibit 13* match the “Total” column in *Exhibit -3* and *Exhibit 4*, respectively.

The final run in this example is designed to test whether estimates generated from the Week 1-8 records, using the new nonresponse adjusted weight, are statistically different from the estimates generated from the full sample, using the original weight. This difference, if significant, is a measure of the potential bias that will result in estimates if records from Week 9-13 in each quarter are omitted from the analysis.

The code used to set up this statistical test is provided in *Exhibit 14*. Note that this code illustrates two “tricks”:

- If you are interested in estimating the statistical significance of the difference in two estimates that are generated from the same sample using different sets of variables, create a dummy analysis file where each record is output twice, as illustrated in the DATA statement in *Exhibit 14*. It is important to not change the NEST variables when doing this, so that SUDAAN properly accounts for the covariance in the significance test.
- If you are comparing two estimates and would like to see the two estimates, the difference in the estimates, and the statistical significance of the difference all in the same table, simply specify a set of CONTRAST statements as was done with the DESCRIPT procedure in *Exhibit 14*. Notice how the first CONTRAST will provide the estimates and the standard errors of the estimates computed using the original weight. The second CONTRAST will provide the same estimates computed using the reduced sample (Week 1-8) and the newly adjusted weights, and the third CONTRAST will provide an estimate of the difference and the p-value associated with the difference.

In this example, the relative bias is computed by groups defined by SEX for illustration purposes.

#### Exhibit 14. Code to Create Bias Estimates and Test Significance

```

Data compare;
set outsud;
weight=wtfa_sa; domain=1; output;
weight=wtfinal; domain=2; output;
run;

PROC DESCRIPT DATA=compare DESIGN=WR USEVMEM=1;
WEIGHT weight;
NEST strat_p psu_p;
CLASS domain sex;
VAR cancer drinker hiv;
TABLES sex;
CONTRAST domain=(1 0) / name="Estimate with Orig Weight";
CONTRAST domain=(0 1) / name="Estimate with Adjusted Weight";
CONTRAST domain=(1 -1) / name="Difference in Estimates";
SETENV COLWIDTH=15;
PRINT MEAN SEMEAN P_MEAN / MEANFMT=F10.6 SEMEANFMT=F10.6 P_MEANFMT=F10.6
      FILETYPE=RTF FILENAME="Example_12_2_OUT8.rtf" REPLACE
      FONTNAME="Arial" FONTSIZE=10
      TOPINCH=1 LEFTINCH=1 RIGHTINCH=2.5 BOTTOMINCH=2.5;
RTITLE "Example 15-2";
run;

```

The output from the above DESCRIPT is presented in *Exhibit 15*.



**Exhibit 15. Example\_12\_2\_OUT8.rtf**

Date: 05-29-2008

SUDAAN

Page: 1

Time: 09:50:54

Table:1

Variance Estimation Method: Taylor Series (WR)

Example 15-2

by: Variable, Sex, CONTRAST.

for: Variable = CANCER.

		CONTRAST		
		Estimate with Orig Weight	Estimate with Adjusted Weight	Difference in Estimates
<b>Sex</b> Total	Cntrst Mean	0.071909	0.072253	-0.000344
	SE Cntrst Mean	0.002126	0.002727	0.001911
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.857422
1 Male	Cntrst Mean	0.061473	0.061516	-0.000043
	SE Cntrst Mean	0.002756	0.003542	0.002331
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.985144
2 Female	Cntrst Mean	0.081641	0.082265	-0.000624
	SE Cntrst Mean	0.002950	0.003687	0.002750
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.820674

**Exhibit 15. Example\_12\_2\_OUT8.rtf-cont.**

Date: 05-29-2008  
Time: 09:50:54

SUDAAN

Page: 2  
Table:1

Variance Estimation Method: Taylor Series (WR)  
Example 15-2  
by: Variable, Sex, CONTRAST.

for: Variable = DRINKER.

		CONTRAST		
		Estimate with Orig Weight	Estimate with Adjusted Weight	Difference in Estimates
<b>Sex</b> Total	Cntrst Mean	0.050309	0.048524	0.001785
	SE Cntrst Mean	0.001842	0.002293	0.001767
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.313217
1 Male	Cntrst Mean	0.055424	0.054253	0.001171
	SE Cntrst Mean	0.002899	0.003732	0.002640
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.657637
2 Female	Cntrst Mean	0.045593	0.043244	0.002349
	SE Cntrst Mean	0.002282	0.002678	0.002123
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.269473

**Exhibit 15. Example\_12\_2\_OUT8.rtf-cont.**

Date: 05-29-2008		SUDAAN		Page: 3
Time: 09:50:54				Table:1
Variance Estimation Method: Taylor Series (WR)				
Example 15-2				
by: Variable, Sex, CONTRAST.				
for: Variable = HIV.				
Sex		CONTRAST		
		Estimate with Orig Weight	Estimate with Adjusted Weight	Difference in Estimates
Total	Cntrst Mean	0.022058	0.024380	-0.002322
	SE Cntrst Mean	0.001463	0.002040	0.001069
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.030540
1 Male	Cntrst Mean	0.024400	0.027277	-0.002877
	SE Cntrst Mean	0.002110	0.002729	0.001450
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.048141
2 Female	Cntrst Mean	0.019888	0.021686	-0.001798
	SE Cntrst Mean	0.001640	0.002352	0.001385
	P-value T-Test Cont. Mean=0	0.000000	0.000000	0.195364

Page 1 of *Exhibit 15*, above, indicates that the cancer estimate is 7.1909% using the original full sample with the original weight, and 7.2253% using the Week 1-8 sample and the new, nonresponse-adjusted weight. The corresponding standard errors are .2126% and .2727%, respectively, so the standard error is greater for estimates generated using the Week 1-8 subsample. This is not surprising since the sample size is smaller. The output also indicates that the p-value associated with the difference is 0.857422 (i.e., the difference is not significant).

The output indicates that the only differences that are significantly different at the .05 level are the HIV estimates for total sample (p=.030540) and for males (p=.048141). As mentioned earlier in this example, we really did not expect to see any differences. In the NHIS, the sample is randomly assigned to weeks, so there is no reason to believe the sample in Weeks 9-13 is really any different than the sample in Weeks 1-8. This dataset was chosen for this example for illustration purposes only.