Overview of Software that Will Produce Sample Weight Adjustments

Michael B. Witt

RTI International, P.O. Box 12194, Research Triangle Park, NC 27709

Abstract

Various algorithms have been developed for computing sample weight adjustments to account for features such as nonresponse and coverage errors (post-stratification). Some of these algorithms are now available in commercial software products such as SUDAAN® and WESVAR®. In addition, some methods for obtaining weight adjustments can be found by adapting existing procedures and functions in general purpose statistical software packages such as SAS®. This paper will present an overview of software that can be used to create weight adjustments. The focus is on options that already exist in software products and are relatively easy to use, particularly for small to mid-size studies.

Key Words: Weights, Weighting Software, Weight Adjustments, Post-stratification, Nonresponse

1. Introduction

Researchers might decompose a data collection effort, for example surveys, into a series of major tasks that include sample design, sample selection, data collection, weighting, imputation and data analysis. The "weighting" task is a common, generic term used to define the activities associated with deriving the final sample weights for a study. Sample weights are typically composed of several factors, including the unconditional inverse of the probability of selection, sometimes a weight trimming factor and various adjustments that account for nonresponse and/or coverage errors. These latter two adjustments are commonly referred to as **nonresponse weight adjustments** and **post-stratification weight adjustments**. For some studies, post-stratification adjustments might also be referred to as **benchmark adjustment factors**. In summary:

- **Nonresponse weight adjustments** are adjustments made to the design weights that account for the nonrespondents to a study. In some multistage studies, nonresponse can occur at more than one stage and therefore more than one nonresponse adjustment may be needed. For example, in an area probability, person-level study, nonresponse might occur at the screening interview (household-level) and at the questionnaire interview (person-level). In this case, separate nonresponse adjustments may be desired at both the household- and person-level.

- **Post-stratification weight adjustments** are adjustments made to calibrate weights so that they sum to known population totals obtained from some external source. These population totals may be gathered from a recent census of the target population or from another study. In general, post-stratification is done to reduce variance and bias in the final estimates from a study. Post-stratification adjustments are typically computed and applied after the nonresponse adjustments are finalized.

Large studies often have relatively large budgets and a longer amount of time allotted in the schedule to develop the final sample weights. In these larger studies, a considerable amount of time is spent identifying an optimal adjustment methodology and the appropriate variables to use in the sample weight adjustment process. Many of these

studies also have a rich history and a devoted audience who are interested in the results – an audience who often are very wary of any changes made to the final sample weight development process. The challenge, of course, is to derive sample weight adjustments that address the various multiple purposes and varied audiences of a large scale study – and this is why the large scale studies have the bigger budgets and longer amounts of time devoted to developing the sample weights. To efficiently derive weight adjustments for these studies, statisticians will often use home-grown macros and/or somewhat complicated procedures for developing the weight adjustments. See for example, (Chen et al., 2005).

Smaller and some medium-size studies do not always have the big budgets and a relatively long amount of time in the schedule devoted to doing sample weight adjustments. This paper summarizes and compares methodologies and software that can be used to derive sample weight adjustments, particularly for these smaller to medium size studies. This investigation focused on weight adjustment solutions that already exist in software packages and are relatively easy (and quick) to program. This investigation did not consider the use of any macros or home-grown solutions to doing weight adjustments as noted above.

Three software packages were considered and compared in this investigation: SAS® (SAS, 2003) WESVAR® (WESVAR, 2008) and SUDAAN® (SUDAAN, 2008). SAS is the most general statistical software package of the three, and offers a larger variety of approaches that can be adapted for computing sample weight adjustments. Only existing procedures and functions in SAS were considered in this investigation. Most of the SAS procedures and functions investigated are also available in other general purpose, commercial statistical software packages such as Stata® or SPSS®.

The weight adjustment approaches available in SAS that were considered in this investigation were:

- Weighting Class Adjustment (**Section 3.1**)
- Computing Weight Adjustments by Determining the Maximum Likelihood Estimates within a Multi-Dimensional Contingency Table (**Section 3.2**)
- Logistic Regression to Predict Response Propensity (**Section 3.3**)
- Computing Weight Adjustments by Solving a Nonlinear Optimization Problem (**Section 3.4**)

Weight adjustments approaches available in WESVAR are discussed in **Section 4** and approaches available in SUDAAN are discussed in **Section 5**.

All of the examples and results in this paper assume one is interested in doing a nonresponse adjustment. With the exception of the logistic regression approach (discussed in **Section 3.3**), all of these approaches can be easily adapted for post-stratification adjustment applications.

2. Data Used

Most of the examples presented in this paper used data from the public use file (PUF) associated with the 2007 National Immunization Survey (NIS). A comparison of the weight adjustment procedures considered in this paper is presented in **Section 6**. In **Section 6** results from doing a weight adjustment with the NIS data are also compared against a weight adjustment that was created using data from the 2007 National Survey on Drug Use and Health (NSDUH) and from the 2006 Fire Fighter Fatality Investigation

and Prevention Program Study (FFFIPP). These three federal studies were selected because they represented a variety of sample designs:

- NIS was a multistage, random digit dialing study with data collected primarily by telephone,
- NSDUH was a multistage, area probability sample with data collected via in-person interviews,
- And FFFIPP was a single stage, stratified simple random sample of establishments (fire departments) selected from a list frame.

A brief description of these studies is provided in **Exhibit 1**.

Exhibit 1.   Summary of Public Use Files Used

2007 National Immunization Survey (NIS). Most of the examples presented in this paper use data from the public use file (PUF) associated with the 2007 NIS. The NIS was sponsored by the National Center for Immunizations and Respiratory Diseases (NCIRD) and conducted jointly by NCIRD and the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention. The NIS was a national, random digit dialing (RDD) study that collected immunization data on children age 19-35 months.

One important component of this study was the collection of data from the child's health care provider(s). After establishing consent from the parent or guardian, a child's health care provider(s) was contacted and asked to provide additional immunization data for the children whose parents responded to the study. Some parents and guardians did not provide consent and for various other reasons, some health care provider(s) did not participate in the study. The examples considered in this paper address the health care provider nonresponse. Specifically, the final adjusted sample weight associated with each responding child record was the initial weight in the examples and a subsequent nonresponse adjustment was sought to account for the nonresponding provider data. The new adjusted weight could then be used to do analysis with the responding provider data. Variables from the NIS PUF that were used to create the provider nonresponse weight adjustment included:

Variable	Definition
RDDWT	Adjusted sample weight for responding child records
PDAT	0=No/1=Yes response indicator for health care providers
AGEGRP	Age category of child
FRSTBRN	First born status indicator for child
INCPOV1	Poverty status (4 levels)
M_AGEGRP	Age category of mother
EDUC1	Education category of mother
CWIC_02	Indicator for current receipt of WIC benefits

A complete description of the 2007 NIS can be found in (Centers for Disease Control and Prevention, 2007).

2007 National Survey on Drug Use and Health (NSDUH). NSDUH was sponsored by the Office of Applied Studies (OAS) within the Substance Abuse and Mental Health Services Administration (SAMHSA). The primary purpose of NSDUH was to measure the prevalence and correlates of drug use in the United States among people 12 years of age or older. This was a household, multistage study with data collected quarterly via in-person interviews.

Data from the NSDUH PUF were used only for the comparative analysis presented in **Section 6**. Data on household or person-level nonrespondents are not available on the PUF, so it was assumed that those respondents that indicated they used marijuana in the past year were "nonrespondents" to the study and the remaining non-past year marijuana users were the "respondents". A nonresponse adjustment was computed that corrected the weights of respondents for the nonrespondents. Variables considered in the adjustment process included gender, age group, education group, health status indicator, race and population density indicator.

A complete description of the 2007 NSDUH PUF can be found in (Substance Abuse and Mental Health Services Administration, 2007).

Exhibit 1 Summary of Public Use Files Used (Continued)

2006 Fire Fighter Fatality Investigation and Prevention Program Study (FFFIPP). The FFFIPP Fire Department Survey was sponsored by the National Institute of Occupational Safety and Health (NIOSH) and was based on a cross-sectional design with stratified random sampling. A stratified, probability sample of 3,000 fire departments representing 10 percent of the approximately 30,000 fire departments in the United States was selected for the survey. The primary method of collecting the data for the FFFIPP Fire Department Survey was via mail. Data was collected on safety practices employed in fire departments, and on the attitudes and knowledge of NIOSH's FFFIPP program.

  As with the NSDUH PUF data, data from the FFFIPP were only used for the comparative analysis presented in **Section 6**. A nonresponse adjustment was created at the fire department level that corrected for nonresponding fire departments. Variables that were used in the adjustment included an indicator for whether the fire department had a fatality or not, Census Region, urban/rural indicator, size of fire department and fire department type (career, volunteer or both).

  A complete description of the 2006 FFFIPP Survey can be found in (Centers for Disease Control and Prevention, 2006).

3. Weighting Class Adjustment Approaches Using Existing SAS Procedures and Functions

The more commonly used general purpose statistical software packages such as SAS, SPSS or Stata do not have procedures or functions that are explicitly designed to produce sample weight adjustments such as a nonresponse or post-stratification adjustment factor. However, these packages have tools that can adapted to provide weight adjustments, which is particularly appealing for the statistician who is interested in doing weight adjustment for a quick-turnaround or smaller study. A few of the procedures and functions that exist in SAS will be examined in this section. Similar procedures exist in other general purpose statistical software packages such as SPSS or Stata.

3.1 Weighting Class Adjustments

A simple, basic and very common method of computing a weight adjustment is the method known as the **weighting class approach** (see for example, Lessler and Kalsbeek, 1992, p. 182-187). With this approach, a multidimensional contingency table is formed using a set of categorical variables. Within each cell of the contingency table, a weight adjustment is computed and applied to each respondent in the cell. If one is interested in doing a nonresponse adjustment, this approach involves computing a nonresponse adjustment factor that gets applied to the respondents in the cell and accounts for the nonrespondents in the cell. If there are no nonrespondents in a cell, the nonresponse adjustment factor would be 1.00. Otherwise the factor is some value greater than 1.00.

Exhibit 2. Weighting Class Adjustment Using PROC MEANS

```
PROC MEANS DATA=indata;
VAR pdat;                /*** 1=Yes/0=No Response Indicator ***/
WEIGHT rddwt;
BY agegrp frstbrn;       /*** Weight Classing Variables ***/
OUTPUT OUT=outprop mean=resp_propensity;

DATA Final;
MERGE indata outprop;
BY agegrp frstbrn;
nonresp_adj=1/resp_propensity;
IF pdat=1 then finalwt=nonresp_adj * rddwt;
```

Performing a weighting class adjustment is relatively easy even for a beginner SAS programmer. In the example presented in **Exhibit 2**, a nonresponse adjustment is computed by classes formed using the categorical variables AGEGRP and FRSTBRN. The weighted mean of the 1=yes/0=no response indicator, PDAT, is computed. This is

the weighted response propensity. And the inverse of the weighted response propensity is the nonresponse weight adjustment that would get applied to the respondents in the corresponding AGEGRP by FRSTBRN cell.

One major advantage of this approach is that it is easy to implement and would work particularly well if one were interested in doing a weight adjustment with just a few categorical variables. However, if the number of categorical variables gets larger, the weight adjustment approach can lead to several unappealing results. For instance:

- A cell in the contingency table may have no respondents and one or more nonrespondents. This is a problem because there are no respondents in the cell to account for the nonrespondents, and simply discarding the nonrespondents from the adjustment process would mean the nonresponse-adjusted, weighted population estimate derived from the respondent data would not sum to the correct population totals. The nonresponse-adjusted, weighted population totals for any domain can be found by summing the nonresponse-adjusted sample weights over records that fall within the domain.

- Computing an adjustment factor within each cell in a complex, multidimensional contingency table may lead to an increase in the disparity of the adjusted weights between respondents because some cells may yield a significantly larger (or smaller) nonresponse adjustment compared to other cells. One statistic that is used to monitor the magnitude of the weight disparity in a study after applying a weight adjustment is the unequal weighting effect. This is defined next.

In general, if one ignores clustering and the sometimes beneficial effects of stratification, as sample weights become more unequal, the variance of estimates produced from the weighted data in a study will increase. This increase is measured by what some statisticians refer to as the **unequal weighting effect**. This statistic can be computed for the total respondent sample or for any subgroup of interest and is defined as follows:

$$UWE = \sum_{i=1}^{n} \frac{n}{\hat{N}^2} \alpha_i^2 d_i^2 \qquad (1)$$

Where n is the sample size, α_i is the weight adjustment for record i, d_i is the original design weight and $\hat{N} = \sum_{i=1}^{n} \alpha_i d_i$ is an estimate of the population total. If all the adjusted sample weights $(\alpha_i d_i)$ were equal in some domain then the UWE would equal one. As the adjusted weights become more unequal, the UWE increases beyond one. A good discussion about the adverse effects of unequal weighting can be found in (Kish, 1965, Section 11.7b).

The problem of a weighting class cell having no respondents (and one or more nonrespondents) and/or the issue of widely varying, adjusted sample weights can be addressed by collapsing weighting class cells within the contingency table. And this is often what is done. However, if cells are collapsed across some dimension in the table then the adjusted sample weights will not sum to the correct margins associated with that dimension. Having sample weights that do not sum to the desired margins of a dimension can be undesirable, particularly if the adjusted weight sums deviate significantly from the desired totals.

3.2 Computing Weight Adjustments by Determining Maximum Likelihood Estimates for the Cells within a Contingency Table

Another way to address the two main disadvantages of the weighting class approach (instead of cell collapsing) is to use an algorithm that will derive adjusted weights that sum to desired <u>margins</u> within a contingency table, but not necessarily to the correct totals within each cell of the contingency table. This type of approach to deriving weight adjustments is generically referred to as **raking**. For a general discussion of raking, see for example (Deming, 1943) or (Kalton, 1983).

Numerous algorithms exist to deriving raking weight adjustments and in many cases, statisticians will simply develop an iterative raking algorithm in SAS's matrix language procedure (IML) or perhaps even in the DATA step in SAS in order to derive the weight adjustments. One procedure that already exists in SAS and that will perform raking is the CALL IPF function in the IML procedure. This routine performs an iterative proportional fit of a contingency table using the method described in Bishop, Fienberg and Holland (2007). An example is provided next.

Suppose we are interested in computing a nonresponse adjustment with four categorical variables in the NIS example (AGEGRP, INCPOV1, M_AGEGRP and EDUC1). And suppose we would like the nonresponse adjusted weights to sum to the correct population totals by the main effects of each of these categorical variables and by the interaction of AGEGRP and INCPOV1.

If we were using a weighting class approach, we would create a nonresponse adjustment in each cell defined by the full interaction of these four variables. But as noted earlier, the disadvantage of this is that some cells may have no respondents and other cells may yield a weight adjustment that is comparatively large (or small) compared to other cells. This contingency table approach addresses these problems to a large extent.

When computing nonresponse adjustments, it is generally desirable to obtain adjustments that allow each respondent to represent themselves (in a weighted sense) as well as some portion of the nonrespondents. In other words, it is desirable to obtain nonresponse adjustments that are greater than or equal to 1.00 for each respondent. Given this restriction on the magnitude of the nonresponse adjustment, we used the CALL IPF raking function to determine the number of nonrespondents that each weight-adjusted respondent should account for, within each cell of the contingency table. So if the function returns a maximum likelihood estimate within a cell that equals zero, the nonresponse adjustment for respondents in that cell would be 1.00. And if the function returned a value greater than zero, the nonresponse adjustment would be some value greater than 1.00.

To state this approach to deriving a nonresponse adjustment using CALL IPF another way, we are going to obtain from the CALL IPF routine the maximum likelihood estimate within each cell of the four-way contingency table. The estimate is the number of nonrespondents that the respondents within the cell should represent. This number should be zero if there are no respondents in the cell. And we only require the sum of this number across the main effect and the one two-way interaction to equal some weighted control totals that are pre-specified. The control totals are the weighted number of nonrespondents.

Assuming the function returns a convergent solution, a nonresponse adjustment (α_{ic}) will be formed for each respondent record i within cell c of this contingency table as follows:

$$\alpha_{ic} = \frac{\hat{D}_c + m_c}{\hat{D}_c} \qquad\qquad (2)$$

Where $\hat{D}_c$ is the sum of the respondent sample weights within cell c and m_c is the weighted number of nonrespondents in cell c that is returned from the CALL IPF function.

From a programming perspective, the syntax for the CALL IPF function is illustrated in **Exhibit 3**.

Exhibit 3.   Computing Weight Adjustments Using the CALL IPF in PROC IML

```
1    /*  Order of variables        = agegrp incpov1 m_agegrp educ1
2        Variable Number           =    4       3       2       1
3        # of Levels of Each Variable =  3       4       3       4     */
4    PROC FREQ DATA= indata NOPRINT;
5    TABLES agegrp*incpov1*m_agegrp*educ1 / sparse out=resp(rename=(count=wresp));
6    WHERE  pdat = 1;  /*Weighted Respondents*/
7    WEIGHT rddwt;
8
9    PROC FREQ data=indata NOPRINT;
10   TABLES agegrp*incpov1*m_agegrp*educ1
11                              / sparse out=nonresp(rename=(count=wnonresp));
12   WHERE  pdat = 0;  /*Weighted Nonrespondents*/
13   WEIGHT rddwt;
14
15   PROC IML;
16   USE nonresp;  READ ALL var{wnonresp} INTO nonresp;  CLOSE nonrespondents;
17   USE resp;     READ ALL var{wresp}    INTO resp;     CLOSE respondents;
18   START=CHOOSE(resp>0,resp+nonresp,0);
19   DIM={4 3 4 3};      /*Reverse of "# of Levels of Each Variable"*/
20   CONFIG={ 1 3 4,
21           2 0 0 };
22   CALL IPF(fit,status,dim,nonresp,config,start);
23   CREATE OUTDSN from fit[colname={newcount}];  APPEND FROM fit;   CLOSE outdsn;
```

On line 22 of **Exhibit 3**, FIT is the returned contingency table. This contains the values m_c in **Equation (2)**. STATUS is a returned flag to let one know if convergence was achieved. DIM is a vector that specifies the number of levels of each variable, in reverse order compared to the order specified on the two PROC FREQ's that precede the call to PROC IML. And NONRESP is a matrix that allows SAS to compute the totals that should be controlled for, i.e. the weighted number of nonrespondents.

CONFIG on lines 20 and 22 of **Exhibit 3** is used to specify which marginal totals should be controlled for. This matrix can be difficult to understand and define. If one would like to control for just the main effects of the four variables, then CONFIG should be set equal to the vector {1 2 3 4}. If one would like to control for the main effect of the variables and the two-way interaction of variable #1 and #2, then CONFIG should be defined as the matrix:

$$CONFIG = \begin{bmatrix} 1\ 3\ 4 \\ 2\ 0\ 0 \end{bmatrix}$$

So the columns of CONFIG indicate which variables should be included in an n-way interaction. Note that we specified that the totals associated with the interaction of variable #1 and #2 be controlled for in the first column of CONFIG, and controlling for the main effect of these variables is already then implied.

START in **Exhibit 3** is the starting values in each cell. We set this equal to the weighted number of respondents plus the weighted number of nonrespondents in each cell so that

the procedure would return cell values that are distributed as close as possible to the original weighted sample distribution. If there are no respondents in a cell, then we set the START value for that cell to zero. This was accomplished via the CHOOSE statement in **Exhibit 3** (see line 18). Setting the start value in this manner was done to minimize the impact of the resultant weight adjustments on the overall unequal weighting effect associated with the adjusted weights.

Finally, the output dataset OUTDSN referenced in **Exhibit 3** (see line 23) contains the m_c values for each cell *c* [see **Equation (2)**].

3.3 Computing a Nonresponse Adjustment by Fitting a Logistic Regression Model

One common approach for computing a nonresponse adjustment is to fit a logistic regression model. The resulting fitted logistic model is used to predict a response propensity for each respondent. And the inverse of the model predicted response propensity is the nonresponse adjustment associated with each respondent.

In this application, the dependent variable is a 0/1 response indicator. And unlike the raking and weighting class approaches, both categorical and continuous variables can be used as predictors. The raking and weighting class approaches only allow one to use categorical variables in the adjustment.

As in the example presented in **Section 3.2**, again suppose one were interested in computing a nonresponse adjustment with four categorical variables in the NIS example (AGEGRP, INCPOV1, M_AGEGRP and EDUC1). And also suppose we would like to control for the interaction of AGEGRP and INCPOV1. Example SAS code that will yield the desired nonresponse adjustment is displayed in **Exhibit 4**.

Exhibit 4. Weighting Adjustments with PROC LOGISTIC

```
PROC LOGISTIC DATA= indata;
WEIGHT   rddwt;
CLASS    agegrp incpov1 m_agegrp educ1;
MODEL    pdat(event='1')=agegrp incpov1 m_agegrp educ1 agegrp*incpov1;
OUTPUT   out=sasout predicted=pred;
run;

DATA adjust(keep=idnos nonresp2);
SET sasout;
Nonresp_adjust=1/pred;
```

Computing a nonresponse adjustment by predicting the response propensity via a logistic model has one very important disadvantage. Specifically, unlike the CALL IPF approach (for example), the final adjusted weights resulting from the logistic model will not necessarily equal the desired control totals. The reason why this is true relates to the procedure used to estimate the logistic model parameters. Assuming one is interested in obtaining the maximum likelihood estimates of the model parameters, estimating the model parameters in a logistic model reduces to solving the following **score functions**:

$$\sum_{i=1}^{n} d_i (y_i - \rho_i) \mathbf{X}_i = \mathbf{0} \tag{3}$$

Where d_i is the design weight for record *i*, y_i is a 0/1 response indicator, $\rho_i = \dfrac{e^{\hat{\beta}'\mathbf{X}_i}}{1 + e^{\hat{\beta}'\mathbf{X}_i}}$

is the model-predicted response propensity, $\hat{\beta}$ is the vector of model coefficients and $\mathbf{X}_i$ is the vector of explanatory variables.

In contrast, when doing a nonresponse adjustment, one generally seeks to find weight adjustments $\alpha_i = \rho_i^{-1}$ that satisfy **calibration equations**:

$$\sum_{i=1}^{n} d_i \alpha_i y_i \mathbf{X}_i = \sum_{i=1}^{n} d_i \mathbf{X}_i \tag{4}$$

In other words, one would like to obtain weight adjustments that will force the adjusted weights [left side of **Equation (4)**] to equal the control totals [right side of **Equation (4)**] over the components of the vector $\mathbf{X}$. In a nonresponse application, the adjusted weights should sum to the population totals, which in this case is $\sum_{i=1}^{n} d_i \mathbf{X}_i$.

After setting $\alpha_i = \rho_i^{-1}$, **Equation (4)** can be re-written as:

$$\sum_{i=1}^{n} d_i \left(\frac{y_i - \rho_i}{\rho_i} \right) \mathbf{X}_i = \mathbf{0} \tag{5}$$

Comparing (5) with (3), it is easy to see the score functions are not the same thing as the calibrations equations. Consequently the score function will not yield nonresponse adjusted weights that satisfy the same equality dictated by the calibrations equations.

3.4 Compute Weight Adjustments by Solving an Nonlinear Optimization Problem (Quadratic)

In **Section 3.1** it was noted that the unequal weighting effect [**Equation (1)**] is a measure of how much the variance of estimates will increase due to the unequal weights (ignoring clustering and stratification). Therefore this is generally something one would like to minimize when computing weight adjustments. In addition, it was noted in **Section 3.2** that we would prefer nonresponse adjustments to be bounded below by 1.00 and it was noted in the previous two sections that we would prefer adjusted weights that will satisfy calibration equations [**Equation (4)**]. Another method of obtaining weight adjustments that will satisfy all of these criteria is to formulate the problem as a nonlinear optimization problem. Specifically, suppose the unknowns that we seek to solve for are the weight adjustments, α_i for $i = 1,...,n_r$. The nonlinear optimization problem we seek to solve is displayed in **Exhibit 5**.

Exhibit 5. Computing Weighting Adjustments by Solving a Nonlinear Optimization Problem

$$\text{Minimize:} \quad UWE = \frac{n_r}{\hat{N}^2} \sum_{i=1}^{n} y_i d_i^2 \alpha_i^2 \tag{6}$$

$$\text{Subject to:} \quad L_i \le \alpha_i \le U_i \tag{7}$$

$$\text{And:} \quad \sum_{i=1}^{n} y_i d_i \alpha_i \mathbf{X}_i = \mathbf{T_x} \tag{8}$$

Where $\hat{N} = \sum_{i=1}^{n} y_i d_i \alpha_i = \sum_{i=1}^{n} d_i$ is the population total, the L_i and U_i for each i are pre-determined constants used to bound the resultant weight adjustments, **Equations (8)** are the calibration equations and the vector $\mathbf{T}_x$ is the vector of control totals.

There are numerous procedures and functions in statistical packages like SAS that can be used to determine the solution to a constrained, quadratic optimization problem as noted

in **Exhibit 5**. One of the newer procedures in SAS is PROC OPTMODEL. An example using PROC OPTMODEL is illustrated in **Exhibit 6**.

Exhibit 6. Determining Weight Adjustments Using PROC OPTMODEL

```
1   PROC OPTMODEL;
2   number n=&resp;                          /*Number of respondents*/
3   var alpha{1..n} >= 1.00 <= 2.00;        /*Bounds on weight adjustments*/
4
5   /*Reading in the data for the constraints*/
6       set<number> ids;
7       number idnos{ids};
8       number rddwt{ids};
9       number intercept{ids};
10      number age1{ids};
11      number age2{ids};
12      ... (more number lines)
13      read data indata into ids=[_n_] idnos rddwt intercept age1 age2 ...;
14
15  min q=sum{i in 1..n} (&resp / &sumwt2) * uss[i] * (alpha[i]**2) ;
16
17  con Xintercept: (sum{i in 1..n} intercept[i]*rddwt[i]*alpha[i]) = 6025084;
18  con Xage1:      (sum{i in 1..n} age1[i]*rddwt[i]*alpha[i])     = 1811763;
19  con Xage2:      (sum{i in 1..n} age2[i]*rddwt[i]*alpha[i])     = 2065164;
20  ... (more con lines)
21
22  EXPAND; SOLVE: CREATE DATA outdsn FROM [i] alpha idnos;
```

Looking at the code in **Exhibit 6**, line 3 shows where one would define the bounds on the adjustments in **Equation (7)**. In this example, we set $L_i = 1.00$ and $U_i = 2.00$ for each i. Lines 5-13 define parameters and read the data in. Line 15 is the objective function, i.e. **Equation (6)**. Lines 17-20 are example calibration equations, i.e. **Equation (8)**. And line 22 tells OPTMODEL to solve the problem and put the solution in a dataset called *outdsn*.

4. Weight Adjustments Using WESVAR

One of the disadvantages of adapting existing functions and procedures from general purpose statistical software packages like SAS is that these are not specifically made for weight adjustments. So the default printout and test statistics will generally not provide a lot of useful information that one needs to evaluate the adjustments. In fact, one may not even be certain the calibration equations are satisfied with some of these procedures. In general, a check of the adjusted weight sums should be done using a procedure like PROC MEANS after each weight adjustment is created and applied to the weights.

Two established statistical packages that have procedures in place specifically for weight adjustments are WESVAR and SUDAAN. SUDAAN will be discussed next.

WESVAR is a statistical software package distributed by WESTAT. WESVAR is appealing because the software is free and extremely user-friendly.

In summary, with WESVAR:
- One can compute nonresponse adjustments using a weighting class approach.
- One can compute nonresponse and post-stratification adjustments using an iterative proportional fitting (IPF) raking approach.
- And this software is particularly useful if replicate weights were constructed for variance estimation. WESVAR will compute weight adjustments for the replicate weights in addition to the main analysis weights.

Exhibit 7 provides an example dialog box one would see when using WESVAR to compute weight adjustments. In this example, weight adjustments using WESVAR's raking approach is being requested. The control variables specified in the grid of this dialog box is where one would indicate which marginals to control for. The "Control Files" on the right side of the grid box is where one would specify text files that contain the control totals. In this example weight adjustments that control for the main effects of AGEGRP, CWIC_02, FRSTBRN and INCPOV1 in the NIS study are being requested. The control totals in the various text files were compute by summing the weights for respondents and nonrespondents by each of the control variables. And the RDDWT in this example was redefined and set equal to zero for nonrespondents so that the resultant weight adjustment would be applicable for just respondents. Finally, note the Absolute/Relative/Misc dialog box in the upper right corner of **Exhibit 7** allows one to set parameters that control the convergence of the raking process.
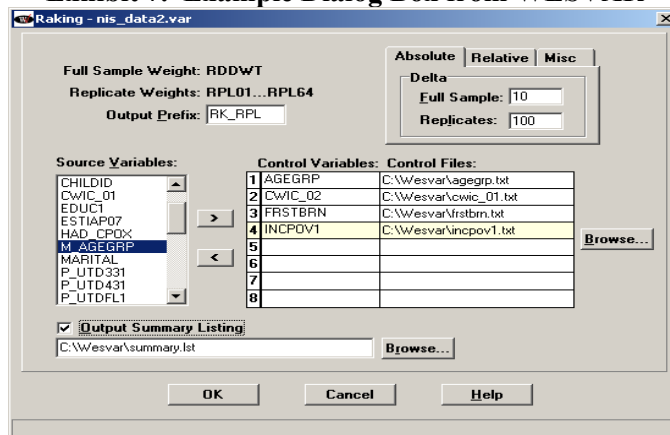
Exhibit 7. Example Dialog Box from WESVAR

Exhibit 8 shows example output from WESVAR (i.e. an example "Output Summary Listing"). This came from the request to do raking presented in **Exhibit 7**. Several lines of the output have been omitted for brevity purposes.

Exhibit 8. Example Output Summary Listing from WESVAR

```
Input WesVar Data File       : C:\Wesvar\nis_data2.var
Raking Converged             : Yes
Actual Number of Iteration   : 6
(more lines)

      AGEGRP    CWIC_02    FRSTBRN    INCPOV1   Factors   Sum of Weights
 1 1.000000  1.000000   1.000000   1.000000    3.2980    20832.3335294324643
 2 1.000000  1.000000   1.000000   2.000000    2.1386    208591.0012385853333
 3 1.000000  1.000000   1.000000   3.000000    1.3025    265625.0576440442819
(more lines)

         AGEGRP   Control Total   Sum of Weights
 1       1.000000    1811763.0119  1811766.6036133710300
 2       2.000000    2065164.0445  2065299.2473833742550
 3       3.000000    2148157.1675  2148018.3729032548140

         CWIC_02   Control Total   Sum of Weights
 1       1.000000    3207469.0465  3205879.6200621663590
(more lines)
```

Exhibit 8 shows the output from WESVAR will indicate whether convergence was achieved and how many iterations were needed to achieve convergence. And for each combination of the control variables, **Exhibit 8** indicates what the final adjustment factor was and what the sum of the weights were for that combination of variables. And for

each of the control variables specified, the output shows the controls totals as well as the sum of the adjusted weights.

In this particular example, **Exhibit 8** indicates the sum of the adjusted weights did not equal the control totals exactly. See for example, AGEGRP=2. The control total is 2,065,164 and the sum of the adjusted weights is 2,065,299. The difference between these numbers can be made smaller by decreasing the Delta's in **Exhibit 7**.

5. Weight Adjustments Using SUDAAN

SUDAAN is another software package that has a procedure specifically designed for doing weight adjustments. SUDAAN syntax is very similar to SAS and in fact, most people use SUDAAN directly within their SAS jobs. Some of the key features of the weight adjustment procedure (called WTADJUST) in SUDAAN include:

- The procedure can be used to compute nonresponse and post-stratification adjustments.
- It produces adjustments using a model-based approach. So all the benefits of model building are realized with this approach (e.g. testing the significance of variables to predict response propensity).
- The model parameters are estimated by solving calibration equations. So the adjusted final weights will equal the desire control totals.
- Pre-set bounds can be set on adjustments to control for the adverse effects of unequal weighting.
- SUDAAN has an option that will allow you to perform weight truncation. Any weight gain/loss resulting from the truncation will be accounted for in the weight adjustment that comes out of the procedure.

Exhibit 9 displays example code in SUDAAN. The procedure has a CLASS and MODEL statement that are very similar to the CLASS and MODEL statements associated with SAS's PROC LOGISTIC. The WTMIN and WTMAX statement tell SUDAAN what the minimum and maximum acceptable weights are before adjustment. Weights outside these bounds are trimmed, and any weight gain or loss from the trimming process is accounted for when computing the adjustment. And in this example the LOWERBD and UPPERBD statements are used. These statements are used to tell SUDAAN the desired upper and lower bounds on the final adjustment.

Exhibit 9. Weight Adjustments with SUDAAN

```
PROC WTADJUST DESIGN=WR ADJUST=nonresponse;
NEST     estiap07 hhid;
WEIGHT   rddwt;
CLASS    agegrp incpov1 m_agegrp educ1;
MODEL    pdat=agegrp incpov1 m_agegrp educ1;
WTMIN    1;
WTMAX    4000;
LOWERBD  1.00;
UPPERBD  2.20;
PRINT    beta sebeta p_beta;
PRINT    uweorig uwetrim uwefinal;
```

By default, the procedure will provide several useful pages of summary statistics on the weight adjustments and the final adjusted weights. In the example displayed in **Exhibit 9**, some information on the model parameters (i.e. the betas) and the unequal weighting effects are requested via the PRINT statements.

Exhibit 10 is an example of the printout one would get from the PRINT UWE line in the code displayed in **Exhibit 9**. This output displays the unequal weighting effects for the original weights; the weights after a weight truncation was applied and the unequal weighting effect after the final weight adjustments were applied.

Exhibit 10. Example Printout From SUDAAN

```
Variance Estimation Method: Taylor Series (WR)
Response variable PDAT: CHILD HAS ADEQUATE PROVIDER DATA (1=Y/0=N)
Nonresponse Adjustment
by: Independent Variables and Effects.

-----------------------------------------------------------
Independent
  Variables and        Original     Trimmed      Final
  Effects              Unequal      Unequal      Unequal
                       Weighting    Weighting    Weighting
                       Effect       Effect       Effect
-----------------------------------------------------------
Intercept                 2.9171       2.8673       2.8471
AGE CATEGORY OF
  CHILD
  1=19 - 23 MONTHS        2.8257       2.7788       2.7685
  2=24 - 29 MONTHS        3.0702       2.9731       2.9714
  3=30 - 35 MONTHS        2.8462       2.8397       2.7951
(more lines)
```

6. Summary and Conclusion

The various weight adjustment approaches considered in this paper were compared using three different public use files from national studies that reflect three different sample designs. These studies include:

- An RDD Study (National Immunization Survey),
- A Household Study (National Survey on Drug Use and Health),
- And an Establishment Study (Study of Fire Departments for NIOSH).

These three studies were summarized in **Exhibit 1**. For each of these studies, we computed a nonresponse-type adjustment and we were interested in controlling for the main effects of about 6 to 8 explanatory variables.

Exhibit 11. Average Rank of the Unequal Weighting Effect (UWE) by Weight Adjustment Approach
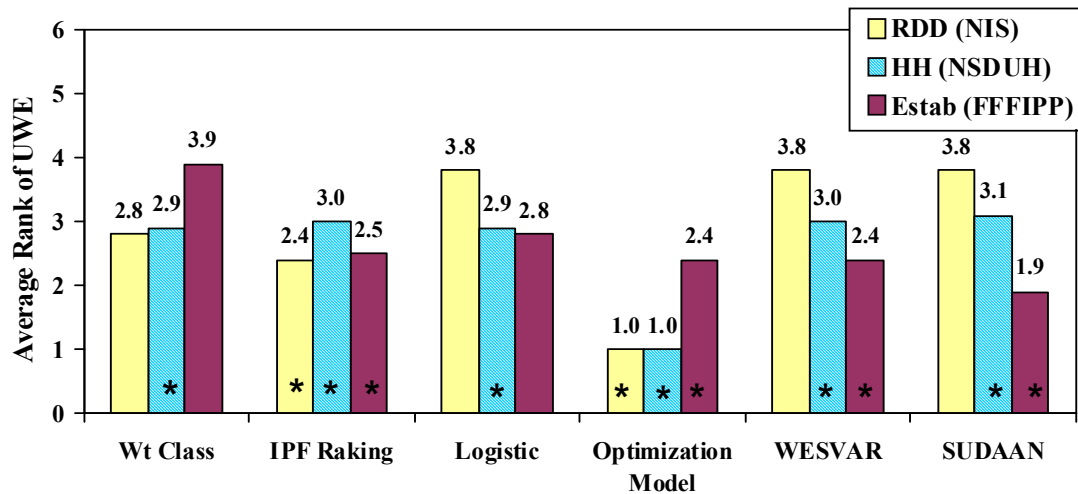
Note: Those procedures marked with an "*"identify those methods that yielded the smaller average UWE.

To compare the methods, we computed the unequal weighting effect (UWE) [see **Equation (1)**] among records for each level of the explanatory variables used in the weight adjustment process. Then within each level, we ranked the UWE across the six weight adjustment methods and then averaged the rank across the levels of all the explanatory variables. The resultant average ranks are displayed in **Exhibit 11.**

For the RDD (NIS) and HH (NSDUH) studies, **Exhibit 11** suggests the optimization model produced the smallest average UWEs and SUDAAN produced the smallest average UWE for the establishment study (FFFIPP). The fact that the optimization produced the smallest average UWE for the RDD and HH study is not a surprise given that the weight adjustments were derived from a model that is set up to explicitly minimize the UWE (see **Exhibit 5**). Seeing that the optimization model did not do as well for the establishment study is somewhat of a surprise. After looking at the results further, we found the optimization model did produce the smallest UWE for the establishment study at the <u>total sample</u> level. However, for some levels of the main effect variables the optimization method did not produce the smallest UWE. When the rank of the UWE was averaged over the different levels of the explanatory variables, SUDAAN happened to yield a smaller average ranking for the establishment study compared to the optimization model.

Those procedures marked with an "*" in **Exhibit 11** identify the methods that yielded the smallest UWEs, on average. From this, it appears the better weighting approaches are the optimization model and raking approaches. With WESVAR and SUDAAN coming in next.

Exhibit 12. Summary of Weight Adjustment Approaches Considered

Feature	Wt Class	IPF Raking	Logistic	Optimization Model	WESVAR[3]	SUDAAN
Can Do Post-Strata <u>and</u> Nonresp Adjust?	Yes	Yes	--	Yes	Yes	Yes
Yields Correct Wt Sums?	Some-times	Yes	--	Yes	Yes	Yes
Can Use More Vars in Adjust?[1]	--	Yes	Yes	Yes	Yes	Yes
Can Use Continuous Vars?[1]	--	--	Yes	Yes	--	Yes
Can Restrict Size of Wt Adjust?[2]	--	--	--	Yes	--	Yes
Diagnostic Info Provided	None	None	Model Fit Statistics	Convergence Information	Some Info, e.g. Wt Sums	Extensive

-- Indicates the feature is not available with the weight adjustment procedure.
[1]Can reduce nonresponse bias in estimates.
[2]Can reduce the UWE and consequently the variance of resultant estimates.
[3]The raking option for doing weight adjustments is being considered with WESVAR.

Exhibit 12 summarizes some of the qualitative results we found after using these weight adjustment methods. In summary:

- All methods can be used to create a nonresponse or a post-strata adjustment, except for the logistic modeling approach.

- All methods give adjustments that satisfy the calibration equations, except for the logistic and weighting class approaches. The weighting class approach is considered a "sometimes" for this feature because often weighting class cells will have no respondents to account for the weight represented by nonrespondents in the cell.

- The weighting class adjustment approach effectively requires one to include <u>all</u> variable interactions in the adjustment process. The other procedures allow one to control for just the main effect and lower order interactions if that is desired. Including only main effect and certain interactions can be advantageous for two reasons. It allows one to use a greater number of variables in the adjustment process and it allows one to include only the statistically significant interactions in the adjustment. Both of these benefits can ultimately lead to less bias in estimates computed from the adjusted weights.

- Continuous variables can be used only with the modeling approaches and the optimization model. Again, using continuous variables enables one to potentially reduce bias in estimates produced using the adjusted weights.

- The optimization model and SUDAAN allow one to pre-specify bounds on the weight adjustment process. Specifying bounds allow one to control the size of the final adjusted weight which in turn allows one to control for the adverse effects of unequal weighting.

- SUDAAN by far provides the most diagnostic information about the final weight adjustments and adjusted weights. WESVAR also provides a good amount of information, including the weight sums.

So after considering both the UWE and the qualitative results (**Exhibits 11** and **12**), the overall conclusion is that both the weighting class and logistic approaches provide fairly good solutions but both have various limitations that were already discussed. The raking approach and WESVAR are much better. The WESVAR software in particular is very easy to use. But overall, the best approaches for doing a weight adjustment for small to mid-size studies are to use either SUDAAN or the optimization model approaches. The optimization model works very well for smaller studies and with a smaller number of calibration equations. And the SUDAAN approach is ideal, really for any size study. It is easy to use a large number of variables with SUDAAN. And SUDAAN provides a good deal of diagnostic information to help evaluate the final weight adjustments.

References:

Bishop, Y. M., S. E. Fienberg, P. W. Holland (2007). <u>Discrete Multivariate Analysis</u>. New York: Springer.

Centers for Disease Control and Prevention (2007). National Immunization Survey. Accessed on 8/1/2009 at http://www.cdc.gov/nis/datafiles.htm

Centers for Disease Control and Prevention (2006). Fire Fighter Fatality Investigation and Prevention Program: Program Description. Accessed on 8/1/2009 at http://www.cdc.gov/niosh/fire/implweb.html

Chen, P., L. Dai, H. Gordek, W. Shi, A. Singh, M. Westlake (2005). "2003 National Survey on Drug Use and Health, Person-Level Sampling Weight Calibration". Report prepared for the Substance Abuse and Mental Health Services Administration,Rockville, Maryland 20857. Accessed on 8/1/2009 at http://www.oas.samhsa.gov/nhsda/2k3MRB/2k3SamplingWgt.pdf

Deming WE. (1943). <u>Statistical Adjustment of Data</u>. New York: Wiley.

Kalton G. (1983). <u>Compensating for Missing Survey Data</u>. Survey Research Center, Institute for Social Research, University of Michigan.

Kish, L. (1965). <u>Survey Sampling</u>. New York, Wiley.

Lessler, J. T. and W. D. Kalsbeek. (1992) <u>Nonsampling Errors in Surveys</u>. New York: Wiley.

SAS Release 9.1.3 (2003). Distributed by SAS Institute Inc., Cary, NC, USA.

Substance Abuse And Mental Health Services Administration (2007). National Survey on Drug Use and Health. Accessed on 8/1/2009 at https://icpsr.umich.edu/cocoon/SAMHDA/STUDY/23782.xml

SUDAAN Release 10.0 (2008). Distributed by Research Triangle Institute, Research Triangle Park, NC, USA.

WESVAR Version 5.1.17 (2008). Distributed by Westat, Rockville, MD, USA.