RTI Press

Methods Report

ISSN 2378-7813

April 2022

Calibration-Weighting a Stratified Simple Random Sample with SUDAAN

Phillip S. Kott

RTI INTERNATIONAL

RTI Press publication MR-0048-2204

RTI International is an independent, nonprofit research organization dedicated to improving the human condition. The RTI Press mission is to disseminate information about RTI research, analytic tools, and technical expertise to a national and international audience. RTI Press publications are peer-reviewed by at least two independent substantive experts and one or more Press editors.

Suggested Citation

Kott, P. S. (2022). *Calibration-Weighting a Stratified Simple Random Sample with SUDAAN*. RTI Press Publication No. MR-0048-2204. Research Triangle Park, NC: RTI Press. https://doi.org/10.3768/rtipress.2022.mr.0048.2204

This publication is part of the RTI Press Methods Report series.

RTI International
3040 East Cornwallis Road
PO Box 12194
Research Triangle Park, NC
27709-2194 USA

Tel: +1.919.541.6000
E-mail: rtipress@rti.org
Website: www.rti.org

Contents

About the Author

Phillip S. Kott, PhD, is a senior research statistician in RTI International's Center of Excellence for Complex Data Analysis.

RTI Press Associate Editor

Carolina Barbosa

Acknowledgments

I would like to thank Carolina Barbosa, the RTI Press associate editor who managed peer-review, and the reviewers whose thoughtful comments made this a stronger paper.

Abstract

This report shows how to apply the calibration-weighting procedures in SAS-callable SUDAAN (Version 11) to a stratified simple random sample drawn from a complete list frame for an establishment survey. The results are calibrated weights produced via raking, raking to a size variable, and pseudo-optimal calibration that potentially reduce and appropriately measure the standard errors of estimated totals. The report then shows how to use these procedures to remove selection bias caused by unit nonresponse under a plausible response model. Although unit nonresponse is usually assumed to be a function of variables with known population or full-sample-estimated totals, calibration weighting can often be used when nonresponse is assumed to be a function of a variable known only for unit respondents (i.e., not missing at random). When producing calibrated weights for an establishment survey, one advantage the SUDAAN procedures have over most of their competitors is that their linearization-based variance estimators can capture the impact of finite-population correction.

Introduction

This report demonstrates how to apply calibration-weighting procedures in SAS-callable SUDAAN Version 11 (RTI International, 2012) to an establishment survey employing a stratified simple random sample. The focus is on generating code and the resulting output rather than on calibration-weighting theory, although some discussions of theory are unavoidable. For more theoretical treatments, the reader is referred to Kott (2014), Kott and Liao (2012), and the references therein.

The fictional sample used in this demonstration consists of manufactured data like those produced in past years by the annual Drug Abuse Warning Network (DAWN) survey (Center for Behavioral Health Statistics and Quality, 2013). It mimics a stratified simple random sample drawn from a frame of hospital emergency departments (EDs) in the United States. The goal is to use this sample to estimate the annual number of drug-related ED visits in the nation. The variables in the sample data include the following:

`RECORD`	
`STRATUM`	
`BIG_N`	Population size in the stratum
`N`	Sample size in stratum
`W`	Design weight (`BIG_N/N`)
`REGION`	`East = 1; South = 2;` `Midwest = 3; West = 4`
`PUBLIC`	`Yes = 1` (a public hospital); `NO = 0`
`METRO`	`Yes = 1` (the hospital is in an urban area); `NO = 0`
`FRAME_VISITS`	Number of previous-year ED visits, which has been recorded on the frame for all EDs
`DR_VISITS`	Annual *drug-related* ED visits collected on the survey

We assume that the frame from which the sample has been drawn is complete and without duplications. Furthermore, the variables `REGION`, `PUBLIC`, `METRO`, and `FRAME_VISITS` on the frame are correct, and various associated frame totals are known.

Let y_k denote the `DR_VISITS` for ED k and d_k its design weight W. The simple expansion estimator for the annual number of drug-related ED visits, $T_y = \sum_U y_k$, is

$$t_y = \sum_S d_k y_k$$

where $\sum_U$ and $\sum_S$ denote summation over the EDs in the population and sample, respectively. As is well known, this estimator is unbiased under probability-sampling theory. In the absence of nonresponse (either at the unit or item levels), calibration weighting can be used to produce nearly unbiased estimators with less variance than t_y. I show how the SUDAAN's calibration-weighting procedures produce those estimates in the next section.

One popular calibration-weighting method, raking, is not effective with this sample, but raking to a size variable (a parallel technique designed for use with an establishment survey having a size variable associated with every element in the frame) is effective. Moreover, even that calibration-weighting method can be improved with quasi-optimal calibration weighting. The SUDAAN calibration weighting procedures, `WTADJUST` and `WTADJX`, not only produce calibrated weights but also measure the standard errors of the calibrated estimates in a nearly unbiased fashion under mild conditions. Consequently, these assertions about raking, raking to a size variable, and quasi-optimal calibration weighting can be demonstrated.

Calibration weighting can be used to compensate for unit nonresponse while retaining the near unbiasedness of the estimator under what is assumed to be a correctly specified logistic response model. SUDAAN also allows one to generate pseudo-maximum-likelihood weights under a logistic response model, but these are often inferior to calibrated weights when the goal is estimating a total or a mean rather than the parameters of the response model.

Calibration to Population Totals

Raking and Raking to a Size Variable

One starts by downloading the data set (DAWN) in the Appendix and creating some additional variables in a new data set (R) with this SAS code:

```
DATA R; SET DAWN;
Q = FRAME_VISITS/1000;
W1 = W – 1;
QW1 = Q*W1;
PUBLICQ = PUBLIC*Q;
PUBLICQW1 = PUBLIC*Q*W1;
METROQ = METRO*Q;
METROQW1 = METRO*Q*W1;
```

One can estimate totals for the United States and within the four regions with this (SAS-callable SUDAAN) code for the expansion estimator:

```
PROC DESCRIPT DATA = R;
    DESIGN = STRWOR; /* the design is stratified
      simple random sampling without replacement */
    NEST STRATUM;
    WEIGHT W;
    TOTCNT BIG_N; /* these stratum population sizes
      are needed for finite-population correction */
    CLASS REGION; /* estimates will be generated by
      region */
    VAR DR_VISITS; /* the survey variable we are
      estimating */
    OUTPUT TOTAL SETOTAL/FILENAME = OUT0
      REPLACE;
RUN;
```

Rather than looking at the output, which the DESCRIPT procedure usually produces, we save it (i.e., the estimated regional and US totals and their standard errors) in OUT0. If one runs SAS STUDIO, then standard SUDAAN output will not appear. Instead, the user can output results into a SAS data set as is in the code above.

The strata in the sample are almost completely cross-classified by region, public/private, and metro/non-metro—but not quite. There is only one hospital in the sample in the East (REGION = 1) that is private (PUBLIC = 0) and not urban (METRO = 0), but it is *not* in its own stratum.

Raking is an iterative procedure to try to get the weighted estimates of the numbers of EDs in each category to equal the frame numbers in those categories. From the frame, we know the total number of public EDs (1642), metro EDs (856), and EDs in each of the four regions (489, 1636, 3124, and 1051). One rakes the weights to those totals using the WTADJUST procedure with the following code and outputs the results (estimated totals and their standard errors) into OUT1:

```
PROC WTADJUST DATA = R DESIGN = STRWOR;
    ADJUST = POST; /* calibration will be to totals
      produced outside the sample */
    NEST STRATUM; WEIGHT W;
    TOTCNT BIG_N;
    CLASS REGION;
    VAR DR_VISITS;
    MODEL _ONE_ = PUBLIC METRO REGION/
      NOINT; /* _ONE_ = 1 */
    POSTWGT 1642 856 489 1636 3124 1051;
    OUTPUT TOTAL SE_TOTAL/FILENAME=OUT1
      REPLACE;
RUN;
```

Note that here, SUDAAN treats PUBLIC and METRO as continuous variables whereas REGION is a categorical variable with four levels. Because it is a categorical variable, it appears in the CLASS statement, and WTADJUST will produce estimated means (drug-related visits per ED) and totals at the regional and US (all regions) levels.

WTADJUST treats raking like a regression with the raking variables as the explanatory variables and 1 (_ONE_) as the dependent variable. NOINT is added to the model statement because there is no intercept.

The zeros in the sixth column in the output table (Table 1) tell us that calibration worked: The weighted numbers metro, public, and EDs in each region (the "sum[s] of the final adjusted weights over respondents") match the numbers in the frame (the "control totals").

Alternatively, the statement

```
OUTPUT TOTORIG TOTTRIM TOTFINAL
  CNTLTOTAL DIFFWT/ FILENAME DIFFWT
  REPLACE;
```

will put the second through sixth column from the table into an output data set DIFFWT.

Expressed mathematically, calibration weighting uses row vectors of calibration variables (PUBLIC, METRO, and REGION) denoted by $\mathbf{x}_k$ for each k and the row vector $\mathbf{T_x}$ for the frame totals of components of $\mathbf{x}_k$ to convert the design weights (d_k) into calibrated

Table 1. Output table from raking

Independent Variables and Effects	Sum of Original Weights Over Respondents	Sum of Trimmed Weights Over Respondents	Sum of Final Adjusted Weights Over Respondents	Control Totals	Final Weight Sum Minus Controls	Original Unequal Weighting Effect	Trimmed Unequal Weighting Effect
PUBLIC	1647.11	1647.11	1642.00	1642.00	0.00	-	-
METRO	856.00	856.00	856.00	856.00	0.00	-	-
REGION							
1	489.00	489.00	489.00	489.00	0.00	2.7157	2.7157
2	1636.00	1636.00	1636.00	1636.00	–0.00	1.1379	1.1379
3	3124.00	3124.00	3124.00	3124.00	–0.00	1.0423	1.0423
4	1051.00	1051.00	1051.00	1051.00	–0.00	1.4846	1.4846

weights (w_k) by finding a column vector **g**—if one exists—that solves the calibration equation. Here, it is

$$\sum_S d_k \exp(\boldsymbol{x}_k \boldsymbol{g}) \boldsymbol{x}_k = \boldsymbol{T_x}, \quad (1)$$

where $w_k = d_k \exp(\boldsymbol{x}_k \boldsymbol{g})$ are the calibrated weights for the k in S.

WTADJUST produces two other tables. One displays the components of **g**, which it calls "beta coefficients." In this case, they are "estimates" of 0 and uninteresting in themselves. Another table produced by WTADJUST displays the estimates of the total ($\Sigma w_k y_k$) and mean ($\Sigma w_k y_k / \Sigma w_k$) of DR_VISITS by region and across all regions (the summations are over the sample in the relevant domain).

The hope is that by calibrating the weights to the numbers of public, metro, and regional EDs standard errors will decrease relative to the expansion estimator. It turns out, as we will see later in the section, they do not. Instead, when the goal is to estimate DR_VISITS, it makes more sense to find a **g** satisfying the following calibration equation:

$$\sum_S d_k \exp(\boldsymbol{x}_k \boldsymbol{g}) z_k = \boldsymbol{T_z}, \quad (2)$$

where $\mathbf{z}_k = \mathbf{x}_k \cdot$ FRAME_VISITS (of ED k), $\mathbf{T_z}$ is the population total of $\mathbf{z}_k$, and, as before, $w_k = d_k \exp(\boldsymbol{x}_k \boldsymbol{g})$. The number of components in $\mathbf{z}_k$ and $\mathbf{x}_k$ need to be equal.

Equation (2) forces the number of weighted estimates of public, metro, and regional frame visits to match their full-population targets. Satisfying the calibration equation in (2) is more sensible than the one in (1) because annual drug-related ED visits are more closely related to previous annual ED visits than the number of EDs. Observe that if T_y were exactly equal to $\mathbf{T_z}\,\mathbf{b}^*$ for some vector $\mathbf{b}^*$, $t_y^{(w)} = \sum_S w_k y_k$ with the calibrated weights satisfying Equation (2) would estimate T_y perfectly.

The following code produces the calibrated weights and resulting estimates (in OUT2) from calibrating with Equation (2):

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = POST;
   NEST STRATUM;
   WEIGHT W;
   TOTCNT BIG_N;
   CLASS REGION; VAR DR_VISITS;
   MODEL _ONE_ = PUBLIC METRO REGION/
    NOINT;
   CALVARS PUBLICQ METROQ REGION*Q/
    NOINT;
   POSTWGT 58000 44000 22000 43000 33000
    36000;
   OUTPUT TOTAL SE_TOTAL/FILENAME=OUT2
    REPLACE;
RUN;
```

The MODEL statement is the same as in the previous WTADJUST, but a CALVARS statement has been added that contains the new calibration variables: PUBLICQ, METROQ, and REGION * Q (note that FRAME_VISITS has been divided by 1000 in each of these variables). The associated population totals, which have been computed using frame information, appear in the POSTWGT statement.

Solving for **g**, which the procedure does implicitly, is akin to a logistic regression without an intercept. The

NOINT in both the MODEL and CALVARS statements indicates that there is no intercept.

A Short Divergence Into Theory

The $\exp(\boldsymbol{x}_k \boldsymbol{g})$ in Equations (1) and (2) and special cases of the following *weight-adjustment function* are as follows:

$$\alpha\left(\boldsymbol{x}_k \boldsymbol{g}; L, U\right) = \frac{L + \exp(\boldsymbol{x}_k \boldsymbol{g})}{1 + \frac{\exp(\boldsymbol{x}_k \boldsymbol{g})}{U}}, \tag{3}$$

where $L = 0$ and $U =$ infinity (actually, it equals 10^{20}, which is close enough to infinity for all practical purposes). Those are the defaults for WTADJUST and WTADJX. The values LOWERBD (L) and UPPERBD (U) bound the size of the weight-adjustment function between L and U. LOWERBD cannot be negative, and UPPERBD must exceed LOWERBD. There is often a centering parameter, CENTER, in what SUDAAN calls "the general exponential model" or "GEM." For our purposes, CENTER need only be defined when UPPERBD takes on its near-infinite default value and LOWERBD is not less than 1. GEM also allows the LOWEBD, UPPERBD, and CENTER to vary across the k, but that wrinkle does not concern us here.

Observe that with calibrated weights satisfying Equation (2),

$$t_y^{(w)} = \sum_S w_k y_k = \boldsymbol{T}_z \boldsymbol{b}^* + \sum_S w_k\left(y_k - \boldsymbol{z}_k \boldsymbol{b}^*\right), \tag{4}$$

where $\mathbf{b}^*$ is the probability limit of

$$\boldsymbol{b} = \left(\sum_S w_k \boldsymbol{x}_k{}^T \boldsymbol{z}_k\right)^{-1} \sum_S w_k \boldsymbol{x}_k{}^T y_k \tag{5}$$

as the sample and population size grow arbitrarily large. We assume that $\mathbf{b}^*$ exists.

Now $\mathbf{T_z b}^*$ is a constant, and under mild conditions (that we assume to hold) $\mathbf{b} \approx \mathbf{b}^*$ and $w_k \approx d_k$, so the variance of $t_y{}^{(w)}$ is nearly equal to the variance of $t_e = \sum_S w_k e_k$ where $e_k = y_k - \mathbf{z}_k \mathbf{b}$; b is treated as if it were a constant. This is what WTADJX computes. So does WTADJUST when $\mathbf{z}_k = \mathbf{x}_k$. It appears that to keep standard errors low, one should find calibration variables that produce e_k with small absolute values.

Two Examples of Pseudo-Optimal Calibration

Although **b** in Equation (5) looks like an estimated regression coefficient, it is not necessarily attached to a linear model. In fact, were the sample drawn using Poisson sampling, a pseudo-optimal version of **b** given the calibration vector $\mathbf{z}_k$ sets $\mathbf{x}_k = \mathbf{z}_k(d_k - 1)$ (see, for example, Kott, 2011) and will usually produce smaller standard errors for estimated totals.

In the previous run of WTADJX the weight adjustment was a function of the model variables, PUBIC, METRO, and the four regions. Replacing each with a model variable of the form Variable * Q * W – 1 will usually result in smaller standard errors. This pseudo-optimal calibration is coded below with estimated totals and their estimated standard errors outputted into data set OUT3.

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = POST;
   NEST STRATUM;
   WEIGHT W;
   TOTCNT BIG_N;
   CLASS REGION; VAR DR_VISITS;
   MODEL _ONE_ = PUBLICQW1 METROQW1
    REGION*QW1/NOINT;
   CALVARS PUBLICQ METROQ REGION*Q/
    NOINT;
   POSTWGT 58000 44000 22000 43000 33000
    36000;
   OUTPUT TOTAL SE_TOTAL/FILENAME = OUT3
    REPLACE;
RUN;
```

An intercept is added to the code below by inserting W1 into the model statement and _ONE_ into the CALVARS statement and the number of EDs on the frame (6300) into the POSTWGT statement, while retaining the NOINTs. The estimated totals and their estimates standard errors outputted into data set OUT4.

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = POST;
   NEST STRATUM; WEIGHT W;
   TOTCNT BIG_N;
   CLASS REGION; VAR DR_VISITS;
   MODEL _ONE_ = W1 PUBLICQW1 METROQW1
    REGION*QW1/NOINT;
```

```
    CALVARS _ONE_ PUBLICQ METROQ
     REGION*Q/NOINT;
    POSTWGT 6300 58000 44000 22000 43000
     33000 36000;
    OUTPUT TOTAL SE_TOTAL/FILENAME = OUT4
     REPLACE;
RUN;
```

We are now ready to compare the estimated totals and their coefficients of variation (CVs) computed in the five sets of alternative weights (i.e., original design, raked, raked to a size variable, quasi-optimal calibrated without an intercept, and quasi-optimal calibrated with an intercept).

```
DATA OUT0; SET OUT0;
DESCV = SETOTAL/TOTAL;
DESTOT = TOTAL; RUN;

DATA OUT1; SET OUT1;
RAK1CV = SE_TOTAL/TOTAL;
RAK1TOT = TOTAL; RUN;

DATA OUT2; SET OUT2;
RAK2CV = SE_TOTAL/TOTAL;
RAK2TOT = TOTAL; RUN;

DATA OUT3; SET OUT3;
QO1CV = SE_TOTAL/TOTAL;
QO1TOT = TOTAL; RUN;

DATA OUT4; SET OUT4;
QO2CV = SE_TOTAL/TOTAL;
QO2TOT = TOTAL; RUN;

DATA C; MERGE OUT0 OUT1 OUT2 OUT3 OUT4;
 BY VARIABLE REGION;

DESCV = ROUND(DESCV * 100, .01);
RAK1CV= ROUND(RAK1CV * 100, .01);
RAK2CV= ROUND(RAK2CV * 100, .01);
QO1CV = ROUND(QO1CV * 100, .01);
QO2CV = ROUND(QO2CV * 100, .01);

PROC PRINT; ID REGION; VAR DESTOTAL
 RAK1TOTAL RAK2TOTAL QO1TOTAL QO2TOTAL;
PROC PRINT; ID REGION; VAR DESCV RAK1CV
 RAK2CV QO1CV QO2CV; RUN;
```

The results are shown in Table 2.

The CVs from raking (second column of CV results) are no better than those of the expansion estimator (first column). Raking to a size variable (third column) decreases the CVs considerably, and quasi-optimal calibration (fourth and fifth columns) usually decreases CVs even further but not so dramatically.

Calibration Weighting for Unit Nonresponse

Some More Theory

`WTADJUST` and `WTADJX` can be used to adjust weights to compensate for unit nonresponse. For example, setting L in Equation (3) at 1 and U at its near-infinite default (10^{20}) treats the unit response/nonresponse mechanism as virtually a logistic function of the components of $\mathbf{x}_k$ with the signs on the regression coefficients reversed.

More generally, let $r_k = 1$ when k is a unit respondent and 0 otherwise. Assume the probability of k responding when sampled is independent of whether any other k' responds; $p_k = E(r_k) = 1/\alpha(\boldsymbol{x}_k\boldsymbol{\gamma}; L, U)$ can range from $1/U$ to $1/L$.

The $\mathbf{g}$ satisfying the calibration equation:

$$\sum_S d_k r_k \alpha(\boldsymbol{x}_k \boldsymbol{g}; L, U)\boldsymbol{z}_k = \boldsymbol{T_z} \tag{6}$$

or

$$\sum_S d_k r_k \alpha(\boldsymbol{x}_k \boldsymbol{g}; L, U)\boldsymbol{z}_k = \boldsymbol{t_z} \tag{7}$$

Table 2. Comparison of the estimated totals and their CVs computed with the five sets of alternative weights

0	5376256.13	5371556.77	5526307.12	5520046.13	5532234.55	6.47	6.48	2.16	1.91	1.87
1	732956.71	732264.61	785406.82	788148.53	787600.83	5.67	5.71	3.32	3.27	3.28
2	1750451.22	1746312.59	1836788.16	1832863.12	1834005.28	13.92	13.94	3.49	2.02	1.95
3	1369022.76	1369134.56	1425517.22	1426590.44	1433664.63	7.55	7.55	3.23	3.22	3.26
4	1523825.45	1523845.00	1478594.92	1472444.04	1476963.82	14.58	14.58	5.77	5.69	5.61

Note: REGION = 0 is the United States (i.e., all regions).

is a consistent estimator for $\boldsymbol{\gamma}$ under mild conditions (i.e., its mean squared error tends to zero as the sample and population grow arbitrarily large). In Equation (7), $\boldsymbol{t_z} = \sum_S d_k \boldsymbol{z}_k$ is a vector of estimated totals based on the full sample before unit nonresponse. For simplicity, we are assuming no item nonresponse.

Equation (6) is called "calibration to the population" and is coded `ADJUST` = `POST` in the SUDAAN calibration procedures. Equation (7) is called "calibration to the full sample" as is coded `ADJUST` = `NONRESPONSE`. Either way, the estimated population total is

$$t_y{}^{(w)} = \sum_S w_k y_k = \sum_S (d_k r_k \alpha_k) y_k,$$

where $\alpha_k = \alpha(\boldsymbol{x}_k \boldsymbol{g}; L, U)$ is called the adjustment factor for ED k. Note that $w_k \approx d_k r_k / p_k$.

When calibrating to the full sample with Equation (7),

$$\begin{aligned} t_y{}^{(w)} &= \sum_S (d_k r_k \alpha_k) y_k \\ &= \sum_S d_k \boldsymbol{z}_k \boldsymbol{b}^* + \sum_S d_k r_k \alpha_k (y_k - \boldsymbol{z}_k \boldsymbol{b}^*) \\ &= \sum_S d_k [\boldsymbol{z}_k \boldsymbol{b}^* + r_k \alpha_k (y_k - \boldsymbol{z}_k \boldsymbol{b}^*)], \end{aligned} \tag{8}$$

where, for technical reasons explained elsewhere in the literature (e.g., Kott & Liao, 2012), $\mathbf{b}^*$ is the probability limit of

$$\begin{aligned} \boldsymbol{b} &= [\sum_S d_k r_k \alpha'(\boldsymbol{x}_k \boldsymbol{g}; L, U) \boldsymbol{x}_k{}^T \boldsymbol{z}_k]^{-1} \\ &\quad \times \sum_S d_k r_k \alpha'(\boldsymbol{x}_k \boldsymbol{g}; L, U) \boldsymbol{x}_k{}^T y_k. \end{aligned}$$

Unlike $\mathbf{b}$, $\mathbf{b}^*$ does not depend on which k are in the sample.

In estimating the variance of $t_y{}^{(w)}$, one cannot treat the square-bracketed term in Equation (8) as a constant because r_k is a random variable (when calibrating to the population, $r_k \alpha_k (y_k - \mathbf{z}_k \mathbf{b}^*)$ replaces the term in the squared brackets). This can cause difficulty when finite-population correction matters, and it often does in establishment surveys based on stratified simple random samples. The SUDAAN calibration weighting procedures handle this by adding the statement `VARNONADJ`. This adds $\sum_S d_k r_k \alpha_k^2 (1 - 1/\alpha_k)(y_k - \mathbf{z}_k \mathbf{b})^2 (n/n_r)$ to the variance estimator, where n is the original sample size, and n_r is the respondent sample size (n/n_r is an ad hoc adjustment for replacing $\mathbf{b}^*$ with $\mathbf{b}$). The addition removes the impact of the original sample finite-population-correction factors ($1 - 1/d_k$, which is a constant within each stratum) on that part of the original sample that did not respond. Note that $1 - 1/\alpha_k$ estimates the probability that sampled ED k is a nonrespondent. It is analogous to the second-stage finite-population-correction factors in a two-stage sample.

Fitting a Logistic Response Model

The SAS data set R contains a variable `RESPONDENT` that is equal to 1 if the ED is a unit respondent to the survey and 0 otherwise. Assuming the probability that an ED responds to the survey is a logistic function of its frame visits (`FRAME_VISITS`) and that the estimated number of frame visits can be determined for the full sample before nonresponse, there are (at least) two ways that SAS-callable SUDAAN can estimate the annual number of drug-related ED visits in the United States (and within census regions) in a nearly unbiased fashion. The first way uses `PROC RLOGIST`, and the second uses `PROC WTADJUST`. The former fits a weighted maximum-likelihood equation ($\sum_S d_k (r_k - 1/[1 + \exp(-\mathbf{z}_k \mathbf{g})]) \mathbf{z}_k{}^T = \mathbf{0}$) rather than the calibration equation. The respective codes follow.

```
PROC RLOGIST DATA = R;
   DESIGN = STRWOR VARNONADJ;
   NEST STRATUM;
   WEIGHT W;
   TOTCNT BIG_N;
   CLASS REGION;
   VAR DR_VISITS;
   MODEL RESPONDENT = LOG_FRAME;
RUN;

PROC WTADJUST DATA = R ADJUST =
 NONRESPONSE;
DESIGN = STRWOR VARNONADJ;
NEST STRATUM;
WEIGHT W;
TOTCNT BIG_N;
CLASS REGION;
VAR DR_VISITS;
LOWERBD 1; CENTER 2; /* UPPERBD is set at its
  near-infinite default; adding CENTER 2 assumes
  virtually the same response function fit by RLOGIST */
MODEL RESPONDENT = LOG_FRAME;
```

```
RUN;
```

Tables 3 and 4 show the key `RLOGIST` results. A 1 percent increase in ED *k*'s frame visits results in an estimated 0.27 percent increase (the value in the `LOG_FRAME` row and Beta Coeff. column) in its odds of response, $p_k/(1\text{-}p_k)$.

Tables 5 and 6 show the analogous `WTADJUST` results.

`WTADJUST` models the adjustment factor (α_k), which is the inverse of the estimated response probability. The `WTADJUST` code above estimates the percent decrease in the odds of an ED responding caused by a one percent increase in its frame visits as –0.30, which is not exactly –0.27. Nevertheless, both values are statistically consistent estimates of the same parameter value. The estimated means and totals from the two procedures are likewise not the same. In all cases, using `RLOGIST` appears to produce smaller standard errors.

Despite this, an advantage of using `WTADJUST` over `RLOGIST` is that one can directly output the nonresponse-adjusted weights by adding a statement to `WTADJUST` like the following:

```
OUTPUT IDVAR WTFINAL ADJFACTOR/
  FILENAME=OUT REPLACE;
```

`WTFINAL` are the adjusted weights (w_k), `ADJFACTOR` the adjustment factors ($r_k\alpha(\mathbf{x}_k{}^T\mathbf{g})$), and `OUT` is the data set containing both and other variables listed on a separate `IDVAR` statement. There is no parallel

Table 3. RLOGIST results for the model variables

Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	*t* test (B = 0)	P-value *t* test (B = 0)
Intercept	–2.80	1.35	–5.46	–0.14	–2.07	0.0393
LOG_FRAME	0.27	0.14	–0.01	0.55	1.91	0.0575

Table 4. RLOGIST results for the drug-related ED visits

Variable		REGION				
		Total	1	2	3	4
DR_VISITS	Mean	858.02	2190.32	1159.92	426.00	1540.93
	SE Mean	62.93	299.67	227.26	36.20	310.60
	Total	5395035.89	775779.95	1821447.05	1499939.77	1297869.12
	SE Total	404227.38	126787.02	372572.36	171359.14	333987.63

Table 5. WTADJUST results for the model variables

Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	*t* test (B = 0)	P-value *t* test (B = 0)	Respondent Sample Size	Nonrespondent Sample Size
Intercept	3.13	1.63	–0.07	6.33	1.92	0.0555	154	192
LOG_FRAME	–0.30	0.17	–0.64	0.03	-1.78	0.0760	-	-

Table 6. WTADJUST results for drug-related ED visits

Variable		REGION				
		Total	1	2	3	4
DR_VISITS	Mean	843.90	2170.05	1135.54	424.14	1528.21
	SE Mean	70.07	303.76	224.86	35.59	312.18
	Total	5316600.95	755474.88	1786181.15	504703.25	1270241.65
	SE Total	441439.17	131260.91	377767.89	171569.81	335005.73

statement in RLOGIST. Moreover, one cannot bound the probabilities of response by $1/U$ and $1/L$ with RLOGIST like one can with WTADJUST.

A third method of estimating drug-related ED visits, under the same logistic response model employs WTADJX. It again features LOG_FRAME as the sole non-intercept in the model statement but adds the statement CALVARS FRAME_VISITS, which means WTADJX attempts to calibrate on frame visits rather than on the log of frame visits (LOG_FRAME) even though response is assumed to be a logistic function of the latter. By calibrating on FRAME_VISITS rather than LOG_FRAME, one is attempting to reduce the size of the terms within the squared brackets of Equation (8) but not necessarily the precision of the regression coefficient.

The code for the WTADJX procedure described above is as follows:

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = NONRESPONSE VARNONADJ;
   NEST STRATUM; WEIGHT W; TOTCNT BIG_N;
   CLASS REGION;
   VAR DR_VISITS;
   LOWERBD 1; CENTER 2;
   MODEL RESPONDENT = LOG_FRAME;
   CALVARS FRAME_VISITS;
RUN;
```

Unfortunately, the output table with the "Final Weight Sum Minus Control Totals" column reveals that calibration fails (not shown). When the numbers in that column are not all zeros (or very close to it), then the estimated totals and means are specious.

Simply replacing FRAME_VISITS by Q (which, recall, is FRAME_VISITS/1000) in the CALVARS statement fixes things. That is why we created Q.

Tables 7 and 8 show the key results after the correction.

Observe that the estimated standard error for the US-level estimated total (and mean) is the least across the three methods. At the same time, the p-value of the estimated LOG_FRAME coefficient is the highest across the methods, suggesting that although it is the most efficient method among the three in estimating means and totals, it is least efficient in estimating the response model.

Calibration Weighting When Nonresponse Is a Function of the Survey Variable

One can also use WTADJX, but not RLOGIST, when nonresponse is assumed to be a function of DR_VISITS itself—that is, when nonresponse is not assumed to be missing at random (Kott & Liao, 2017). In the code below, response is assumed to be a logistic function of an intercept and the log of DR_VISITS, which is denoted LOG_DR:

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = NONRESPONSE VARNONADJ;
   NEST\\WEIGHT W; TOTCNT BIG_N;
   CLASS REGION;
   VAR DR_VISITS;
```

Table 7. WTADJX results for the model variables

Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	*t* test (B = 0)	P-value *t* test (B = 0)
Intercept	2.67	1.48	−0.25	5.59	1.80	0.0734
LOG_FRAME	−0.25	0.15	−0.56	0.05	−1.64	0.1010

Table 8. WTADJX results for drug-related ED visits

Variable		REGION				
		Total	1	2	3	4
DR_VISITS	Mean	863.98	2198.93	1170.11	426.77	1546.26
	SE Mean	55.20	299.70	219.21	37.26	307.20
	Total	5443069.22	786582.43	1841426.54	1501971.84	1313088.42
	SE Total	347760.02	131421.13	345990.63	171716.81	332202.08

```
    LOWERBD 1; CENTER 2;
    MODEL RESPONDENT = LOG_DR;
    CALVARS Q;
    OUTPUT TOTAL SE_TOTAL/FILENAME=OUT0
     REPLACE;
RUN;
```

Rather than printing the total here, the totals and their standard errors under the model have been outputted into OUT0 for future comparison.

As in the previous section, standard errors are likely reduced by calibrating to population totals (for the population size and FRAME_VISITS) rather than full-sample-estimated totals:

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = POST VARNONADJ;
    NEST STRATUM; WEIGHT W; TOTCNT BIG_N;
    CLASS REGION; VAR DR_VISITS;
    LOWERBD 1; CENTER 2;
    MODEL RESPONDENT = LOG_DR;
    CALVARS Q;
    POSTWGT 6300 134000;
    OUTPUT TOTAL SE_TOTAL/FILENAME=OUT1
     REPLACE;
RUN;
```

These totals and their standard errors under the model have been outputted into OUT1.

One can likely reduce the standard errors further by adding more population targets akin to raking to a size variable. At the same time, this adds potential dummies for public, metro, and region to the assumed response model (even when they are not needed in response modeling, which appears to be the case here):

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = POST VARNONADJ;
    NEST STRATUM; WEIGHT W; TOTCNT BIG_N;
    CLASS REGION; VAR DR_VISITS;
    LOWERBD 1; CENTER 2;
    MODEL RESPONDENT = LOG_DR PUBLIC
     METRO REGION;
    CALVARS Q PUBLICQ METROQ REGION*Q;
    POSTWGT 6300 134000 58000 44000 22000
     43000 33000 36000;
    OUTPUT TOTAL SE_TOTAL/FILENAME=OUT2
     REPLACE;
    OUTPUT ADJFACTOR/FILENAME = AFNR
     REPLACE;
```

These totals and their standard errors under the model have been outputted into OUT2.

The addition of the line "OUTPUT ADJFACTOR/ FILENAME = AFNR REPLACE;" in the code above creates data set AFNR, with the adjustment factors. A PROC UNIVARIATE of the adjustment factors (not shown) suggests we can set the upper bound at 2.75 for the adjustment factors in the code below with the hope or containing the variability of the weights and thus standard errors:

```
PROC WTADJX DATA = R DESIGN = STRWOR
 ADJUST = POST VARNONADJ;
    NEST STRATUM; WEIGHT W; TOTCNT BIG_N;
    CLASS REGION; VAR DR_VISITS;
    LOWERBD 1; CENTER 2; UPPERBD 2.75;
    MODEL RESPONDENT = LOG_DR PUBLIC
     METRO REGION;
    CALVARS Q PUBLICQ METROQ REGION*Q;
    POSTWGT 6300 134000 58000 44000 22000
     43000 33000 36000;
    OUTPUT TOTAL SE_TOTAL/FILENAME=OUT3
     REPLACE; /* When running SAS STUDIO,
     we make sure calibration is successful with the
     following two optional lines */
    OUTPUT DIFFWTZ/FILENAME = DIFF
     REPLACE;
PROC PRINT DATA = DIFF; RUN; /* All DIFFWTZ
 should be 0 */
```

The new totals and their standard errors under the model have been outputted into OUT3.

The following code compares the various estimates of the totals. Results are shown in Table 9.

```
DATA OUT0; SET OUT0;
NONCV = SE_TOTAL/TOTAL;
NONTOTAL = TOTAL;
RUN;

DATA OUT1; SET OUT1;
POSTCV = SE_TOTAL/TOTAL;
POSTTOTAL = TOTAL;
RUN;

DATA OUT2; SET OUT2;
RSCV = SE_TOTAL/TOTAL;
```


https://doi.org/10.3768/rtipress.2022.mr.0048.2204

Table 9. Results of alternative estimates of the totals

REGION	NONRTOTAL	POSTTOTAL	PRSTOTAL	PRS2TOTAL	NONRCV	POSTCV	PRSCV	PRS2CV
0	5380683.57	5525234.55	5616497.38	5610019.75	6.50	3.65	3.05	2.75
1	778121.80	813097.48	849767.71	829900.80	16.74	19.07	5.47	8.46
2	1815199.62	1878727.90	1839374.44	1847249.90	18.54	16.97	4.78	4.08
3	1490449.25	1489152.34	1413915.27	1412291.56	11.34	11.54	4.12	4.23
4	1296912.89	1344256.82	1513439.96	1520577.49	25.07	23.67	6.44	8.27

```
RSTOTAL = TOTAL;
RUN;

DATA OUT3; SET OUT3; ;
RS2CV = SE_TOTAL/TOTAL;
RS2TOTAL = TOTAL;
RUN;

DATA C; MERGE OUT0 OUT1 OUT2 OUT3; BY
 VARIABLE REGION;
NONCV = ROUND(NONCV * 100, .01);
POSTZCV= ROUND(POSTCV * 100, .01);
RSCV= ROUND(RSCV * 100, .01);
RS2CV = ROUND(RS2CV * 100, .01);

PROC PRINT; ID REGION; VAR NONTOTAL
 POSTTOTAL RSTOTAL RS2TOTAL;
PROC PRINT; ID REGION; VAR NONCV POSTCV
 RSCV RS2CV; RUN;
```

where

- `NONRxx` has been computed with `ADJUST = RESPONSE` and `CALVARS Q`,
- `POSTxx` with `ADJUST = POST` and `CALVARS Q`,
- `PRSxx` with `ADJUST = POST` and raking to size with default upper bound, and
- `PRS2xx` with `ADJUST = POST` and raking to size with an upper bound of 2.75.

As expected, calibrating to the population tends to reduce standard errors compared with calibrating to the sample, although not in every region. Adding some raking-to-size variables reduces standard errors even more. Constraining the weight adjustments to be no greater than 2.75 decreases the estimated US-level standard error by around 10 percent but increases the estimated standard error in Region 1 by roughly 35 percent. There are small estimated standard error increases in two of the three other regions as well. This shows that smaller upper bounds (and presumably less-variable calibrated weights) do not necessarily translate into smaller standard errors.

One can check whether the response model with and without the upper bound of 2.75 produce significantly different totals in the following manner. Make two copies of each record in newly created data set R2. Place one copy into `DOMAIN` = 1 and the other into `DOMAIN` = 2. `DOMAIN` = 1 has no upper bound (i.e., the default is used), whereas `DOMAIN` = 2 has an upper bound of 2.75. (So long as it is finite, the value assigned to `CENTER` only affects the intercept.) `DOMAIN` replaces the previously missing intercept in the `CALVARS` statement, so there is now a `NOINT` at the end of both the `MODEL` and `CALVARS` statements. When comparing estimates, it is common to treat the design as with replacement:

```
DATA R2; SET R;
PSU + 1;
U = .; DOMAIN = 1; OUTPUT;
U = 2.75; DOMAIN = 2; OUTPUT;
PROC WTADJX DATA = R2 DESIGN = WR;
ADJUST = POST; /* One can also use ADJUST=
  NONRESPONSE for this for test */
NEST STRATUM PSU; WEIGHT W;
CLASS DOMAIN REGION;
VDIFFVAR DOMAIN =(1 2); /* This tells SUDAAN
  to compare estimates between domains */
VAR DR_VISITS;
LOWERBD 1; CENTER 2; UPPERBD U;
MODEL RESPONDENT = DOMAIN LOG_DR*DOMAIN
  PUBLIC*DOMAIN;
METRO*DOMAIN REGION*DOMAIN/NOINT;
CALVARS DOMAIN Q*DOMAIN PUBLICQ*DOMAIN;
METROQ*DOMAIN REGION*DOMAIN*Q/NOINT;
POSTWGT 6300 6300 134000 134000 58000
  58000;
44000 44000 22000 22000 43000 43000
  33000 33000 36000 36000;
```

```
OUTPUT TOTAL SE_TOTAL /FILENAME=TEST
  REPLACE;
```

We compute *t*-values for the differences, but they are not significant. In fact, all the *t*-values are well below 1 in absolute value. Similar results (Table 10) obtain when the ADJUST = NONRESPONSE option is used:

```
DATA TEST; SET TEST;
T_TOTAL = ROUND(TOTAL/SE_TOTAL, .01);
PROC PRINT; ID REGION; VAR TOTAL SE_
  TOTAL T_TOTAL;
RUN;
```

Table 10. Results using the ADJUST = NONRESPONSE option

REGION	TOTAL	SE_TOTAL	T_TOTAL
0	6477.63	330062.23	0.02
1	19866.90	117306.69	0.17
2	−7875.46	162475.88	−0.05
3	1623.70	116702.29	0.01
4	−7137.52	217347.29	−0.03

A Few Additional Comments

If we add this OUTPUT statement to the previous PROC WTADJX:

```
OUTPUT BETA P_BETA/FILENAME=BETA
  REPLACE; then
PROC PRINT DATA =BETA; VAR BETA P_BETA;
```

produces the results shown in Table 11.

The odd observations are for the model with a near-infinite upper bound and the even for the model with an upper bound of 2.75. Observe that even though the estimates for the total number of drug-related ED visits are improved by adjusting for nonresponse using either version of the model, none of the estimated coefficients are significant at less than the 0.4 level.

The nonresponse model fit is better when we replace ADJUST = POST with ADJUST = NONRESPONSE. In addition, we remove the VDIFFVAR statement (and ignore the POSTWGT statement). The PROC PRINT of the OUTPUT BETA statement is shown in Table 12.

The coefficients for both the intercept and LOG_ER are significant at the 0.1 level when the upper bound is near infinite. Even though the corresponding estimated coefficients have larger absolute values with an upper bound of 2.75, neither is significant at the 0.3 level.

The added code

```
DATA TEST; SET TEST; IF DOMAIN > 0;
```

Table 11. Results for OUTPUT BETA P_BETA statement

Obs	BETA	P_BETA
1	1.79	0.5379
2	3.52	0.7421
3	−0.20	0.6050
4	−0.41	0.7546
5	0.48	0.8175
6	0.53	0.8228
7	−1.11	0.4113
8	−3.69	0.9929
9	0.43	0.8680
10	3.00	0.9942
11	−0.28	0.9015
12	2.47	0.9953
13	−0.32	0.7008
14	−0.69	0.7890
15	0.00	.
16	0.00	.

Table 12. PROC PRINT of the OUTPUT BETA statement

Obs	BETA	P_BETA
1	3.43	0.0241
2	6.10	0.3198
3	−0.42	0.0512
4	−0.74	0.3343
5	0.90	0.1378
6	1.02	0.3005
7	−0.61	0.4092
8	−0.41	0.8022
9	−0.55	0.6112
10	−0.82	0.7150
11	−1.04	0.1407
12	−1.17	0.5546
13	−0.74	0.1651
14	−1.35	0.3931
15	0.00	.
16	0.00	.

```
CV_TOTAL = ROUND(100*SE_TOTAL/TOTAL,
 .01);
PROC PRINT; ID DOMAIN; VAR TOTAL SE_
 TOTAL CV_TOTAL; RUN;
```

reveals that the two versions produce very similar estimates for the drug-related ED visits (Table 13).

Observe that the CVs for the estimated totals are much larger here than when `ADJUST` = `POST` was used (3.05 and 2.75). This suggests that although employing `ADJUST` = `NONRESPONSE` is the better option for fitting a nonresponse model, `ADJUST` = `POST` is better for estimating means and totals. To see why, recall from Equation (8) that when calibrating to the full sample

$$t_y^{(w)} = \sum_S d_k[\mathbf{z}_k \mathbf{b}^* + r_k \alpha_k(y_k - \mathbf{z}_k \mathbf{b}^*)],$$

and so,

$$Var(t_y^{(w)}) = Var\{\sum_S d_k[\mathbf{z}_k \mathbf{b}^* + r_k \alpha_k(y_k - \mathbf{z}_k \mathbf{b}^*)]\}.$$

When calibrating to the population (analogous to Equation (4))

$$t_y^{(w)} = \sum_U \mathbf{z}_k \mathbf{b}^* + \sum_S d_k r_k \alpha_k(y_k - \mathbf{z}_k \mathbf{b}^*), \text{and}$$

$$Var(t_y^{(w)}) = Var[\sum_S d_k r_k \alpha_k(y_k - \mathbf{z}_k \mathbf{b}^*)].$$

The former expression for the variance must account for the contribution coming from the random variable $\sum_S d_k \mathbf{z}_k \mathbf{b}^*$, whereas the latter does not.

In calibration weighting, fitting a weight-adjustment function or, equivalently, a nonresponse model, is simply a means to an end—producing better estimated means and totals. Sometimes the SAS log of a SUDAAN calibration-weighting procedure announces that convergence has not been reached, referring to a requirement needed to properly estimate the variances of the components of **g** in Equation (2), Equation (6), or Equation (7) (the `BETA` coefficients in the SUDAAN output). What is relevant to the success of calibration weighting, however, is whether the calibration equation is satisfied; that is, if the weighted sum of the calibration variables among respondents equals the designated control totals (whether from the population or the full sample). If it does, then calibration weighting succeeds.

Table 13. Comparison of estimates for drug-related ED visits

DOMAIN	TOTAL	SE_TOTAL	CV_TOTAL
1	5414966.10	579766.14	10.71
2	5412607.41	584681.51	10.80

Some Concluding Remarks

The main purpose of this report is to demonstrate several ways the calibration procedures in SAS-callable SUDAAN (Version 11) can be applied to a stratified simple random sample. Employing a fictional sample of a hospital ED, I computed estimates of drug-related visits. In the absence of nonresponse, design weights were raked using the `WTADJUST` procedure, but that did not lead to decreases in estimated standard errors. Instead, raking to a size variable using `WTADJX` did. Another technique using `WTADJX`, quasi-optimal calibration, decreased standard errors even more. The point here is to not to introduce raking to a size measure or quasi-optimal calibration to the literature but to demonstrate how to execute those helpful techniques with `WTADJX`.

Similarly, this report demonstrates how `WTADJUST` and `WTADJX` could be used to adjust for unit nonresponse if the probability of unit response was logistic or bounded logistic (i.e., bounded so that the resulting weight adjustment factor was no greater than 2.75). Moreover, these procedures produced statistically defensible standard error estimates.

References

Center for Behavioral Health Statistics and Quality. (2013). *Drug Abuse Warning Network methodology report, 2011 update*. Substance Abuse and Mental Health Services Administration. https://www.samhsa.gov/data/sites/default/files/DAWN2k11ED/DAWN2k11ED/rpts/DAWN2k11-Methods-Report.htm

Kott, P. (2011). A nearly pseudo-optimal method for keeping calibration weights from falling below unity in the absence of nonresponse or frame errors. *Pakistan Journal of Statistics*, *27*4391–396.

Kott, P. (2014). Calibration weighting when model and calibration variables can differ. In F. Mecatti, P. L. Conti, & M. G. Ranalli (Eds.), *Contributions to sampling statistics: ITACOSM 2013 selected papers* (pp. 1–18). Springer.

Kott, P., & Liao, D. (2017). Calibration weighting for nonresponse that is not missing at random: Allowing more calibration than response-model variables. *Journal of Survey Statistics and Methodology*, *5*(2), 159–174. https://doi.org/10.1093/jssam/smx003

Kott, P. S., & Liao, D. (2012). Providing double protection for unit nonresponse with a nonlinear calibration-weighting routine. *Survey Research Methods*, *6*2105–111. https://doi.org/10.18148/srm/2012.v6i2.5076

RTI International. (2012). *SUDAAN language manual, release 11.0*. RTI International.

Appendix. The Drug Abuse Warning Network (DAWN) SAS Data Set

```
DATA DAWN;
INPUT RECORD STRATUM BIG_N N W REGION
 PUBLIC METRO;
FRAME_VISITS ER_VISITS RESPONDENT;
CARDS;
1 1 289 8 36.125 2 0 0 1462 57 0
2 1 289 8 36.125 2 0 0 1477 57 0
3 1 289 8 36.125 2 0 0 1630 63 0
4 1 289 8 36.125 2 0 0 1764 69 1
5 1 289 8 36.125 2 0 0 2149 84 1
6 1 289 8 36.125 2 0 0 2547 151 1
7 1 289 8 36.125 2 0 0 2561 151 0
8 1 289 8 36.125 2 0 0 4189 248 0
9 2 939 22 42.6818 3 0 0 3849 183 0
10 2 939 22 42.6818 3 0 0 3877 184 0
11 2 939 22 42.6818 3 0 0 4105 195 1
12 2 939 22 42.6818 3 0 0 4119 195 1
13 2 939 22 42.6818 3 0 0 4152 197 1
14 2 939 22 42.6818 3 0 0 4277 203 0
15 2 939 22 42.6818 3 0 0 5075 241 0
16 2 939 22 42.6818 3 0 0 6310 338 1
17 2 939 22 42.6818 3 0 0 6331 339 1
18 2 939 22 42.6818 3 0 0 6342 340 1
19 2 939 22 42.6818 3 0 0 6347 340 1
20 2 939 22 42.6818 3 0 0 6381 342 0
21 2 939 22 42.6818 3 0 0 6400 343 1
22 2 939 22 42.6818 3 0 0 6432 345 0
23 2 939 22 42.6818 3 0 0 7233 388 0
24 2 939 22 42.6818 3 0 0 7429 398 1
25 2 939 22 42.6818 3 0 0 7541 404 1
26 2 939 22 42.6818 3 0 0 11702 324 0
27 2 939 22 42.6818 3 0 0 11799 327 0
28 2 939 22 42.6818 3 0 0 12755 353 1
29 2 939 22 42.6818 3 0 0 14862 411 0
30 2 939 22 42.6818 3 0 0 16379 453 1
31 3 545 18 30.2778 4 0 0 7477 165 0
32 3 545 18 30.2778 4 0 0 7498 165 1
33 3 545 18 30.2778 4 0 0 8161 180 0
34 3 545 18 30.2778 4 0 0 8979 198 0
35 3 545 18 30.2778 4 0 0 30604 1883 0
36 3 545 18 30.2778 4 0 0 30669 1887 1
37 3 545 18 30.2778 4 0 0 31097 1913 0
38 3 545 18 30.2778 4 0 0 32841 625 0
39 3 545 18 30.2778 4 0 0 34741 2138 0
40 3 545 18 30.2778 4 0 0 43127 2327 0
41 3 545 18 30.2778 4 0 0 46862 2529 0
42 3 545 18 30.2778 4 0 0 49275 2659 1
43 3 545 18 30.2778 4 0 0 50600 1680 1
44 3 545 18 30.2778 4 0 0 50968 1692 0
45 3 545 18 30.2778 4 0 0 51405 979 0
46 3 545 18 30.2778 4 0 0 52879 1756 1
47 3 545 18 30.2778 4 0 0 54787 1043 0
48 3 545 18 30.2778 4 0 0 55965 1858 1
49 4 170 9 18.8889 1 1 0 6196 225 0
50 4 170 9 18.8889 1 1 0 6351 387 0
51 4 170 9 18.8889 1 1 0 6365 231 0
52 4 170 9 18.8889 1 1 0 6660 406 0
53 4 170 9 18.8889 1 1 0 7307 266 0
54 4 170 9 18.8889 1 1 0 9029 508 0
55 4 170 9 18.8889 1 1 0 9592 540 1
56 4 170 9 18.8889 1 1 0 10400 622 0
57 4 170 9 18.8889 1 0 0 10602 634 0
58 5 153 10 15.3 2 0 1 915 41 0
59 5 153 10 15.3 2 0 1 1020 45 0
60 5 153 10 15.3 2 0 1 1465 40 0
61 5 153 10 15.3 2 0 1 3757 150 0
62 5 153 10 15.3 2 0 1 3763 150 1
63 5 153 10 15.3 2 0 1 8149 383 0
64 5 153 10 15.3 2 0 1 18566 1030 1
65 5 153 10 15.3 2 0 1 19375 578 1
66 5 153 10 15.3 2 0 1 20519 1059 0
67 5 153 10 15.3 2 0 1 206559 3205 1
68 6 2142 51 42 3 0 0 4712 289 1
69 6 2142 51 42 3 0 0 4753 291 0
70 6 2142 51 42 3 0 0 4938 303 1
71 6 2142 51 42 3 0 0 5016 307 0
72 6 2142 51 42 3 0 0 5182 317 0
73 6 2142 51 42 3 0 0 5369 329 0
74 6 2142 51 42 3 0 0 5408 331 0
75 6 2142 51 42 3 0 0 5504 337 1
76 6 2142 51 42 3 0 0 5536 339 0
77 6 2142 51 42 3 0 0 5738 156 0
78 6 2142 51 42 3 0 0 5820 158 0
79 6 2142 51 42 3 0 0 9308 365 1
80 6 2142 51 42 3 0 0 9313 365 0
81 6 2142 51 42 3 0 0 9314 365 0
82 6 2142 51 42 3 0 0 9360 367 1
83 6 2142 51 42 3 0 0 9386 368 0
84 6 2142 51 42 3 0 0 9406 351 0
85 6 2142 51 42 3 0 0 9406 368 1
```

```
86 6 2142 51 42 3 0 0 9415 369 1
87 6 2142 51 42 3 0 0 9428 352 1
88 6 2142 51 42 3 0 0 9429 352 1
89 6 2142 51 42 3 0 0 9442 352 0
90 6 2142 51 42 3 0 0 9444 352 1
91 6 2142 51 42 3 0 0 9469 371 1
92 6 2142 51 42 3 0 0 9576 375 0
93 6 2142 51 42 3 0 0 9602 358 1
94 6 2142 51 42 3 0 0 9679 379 0
95 6 2142 51 42 3 0 0 9716 363 0
96 6 2142 51 42 3 0 0 9850 386 0
97 6 2142 51 42 3 0 0 10011 374 1
98 6 2142 51 42 3 0 0 10259 383 1
99 6 2142 51 42 3 0 0 10392 407 0
100 6 2142 51 42 3 0 0 10415 389 0
101 6 2142 51 42 3 0 0 10599 396 1
102 6 2142 51 42 3 0 0 10774 402 1
103 6 2142 51 42 3 0 0 11212 418 0
104 6 2142 51 42 3 0 0 11435 448 0
105 6 2142 51 42 3 0 0 11453 427 0
106 6 2142 51 42 3 0 0 11823 463 0
107 6 2142 51 42 3 0 0 12135 334 0
108 6 2142 51 42 3 0 0 12980 508 0
109 6 2142 51 42 3 0 0 15044 414 1
110 6 2142 51 42 3 0 0 16054 599 0
111 6 2142 51 42 3 0 0 17667 801 1
112 6 2142 51 42 3 0 0 19734 884 1
113 6 2142 51 42 3 0 0 20102 900 1
114 6 2142 51 42 3 0 0 22777 1316 1
115 6 2142 51 42 3 0 0 24782 1124 1
116 6 2142 51 42 3 0 0 25166 1455 0
117 6 2142 51 42 3 0 0 28437 1610 0
118 6 2142 51 42 3 0 0 29099 1647 0
119 7 410 13 31.5385 4 0 0 1085 50 0
120 7 410 13 31.5385 4 0 0 1162 53 0
121 7 410 13 31.5385 4 0 0 14869 570 0
122 7 410 13 31.5385 4 0 0 16875 647 0
123 7 410 13 31.5385 4 0 0 17538 680 1
124 7 410 13 31.5385 4 0 0 18152 703 1
125 7 410 13 31.5385 4 0 0 21148 819 1
126 7 410 13 31.5385 4 0 0 21666 975 0
127 7 410 13 31.5385 4 0 0 26040 1172 1
128 7 410 13 31.5385 4 0 0 66555 3025 0
129 7 410 13 31.5385 4 0 0 73509 3341 1
130 7 410 13 31.5385 4 0 0 95355 4658 1
131 7 410 13 31.5385 4 0 0 104067 5083 0
132 8 53 53 1 1 1 1 7684 490 0
133 8 53 53 1 1 1 1 14151 711 0
134 8 53 53 1 1 1 1 17314 1020 0
135 8 53 53 1 1 1 1 22046 257 0
136 8 53 53 1 1 1 1 22731 894 0
137 8 53 53 1 1 1 1 28789 1008 0
138 8 53 53 1 1 1 1 30435 1636 0
139 8 53 53 1 1 1 1 31629 1179 0
140 8 53 53 1 1 1 1 31884 676 0
141 8 53 53 1 1 1 1 32820 548 0
142 8 53 53 1 1 1 1 36782 211 1
143 8 53 53 1 1 1 1 39885 1474 0
144 8 53 53 1 1 1 1 40117 2151 1
145 8 53 53 1 1 1 1 41550 1524 0
146 8 53 53 1 1 1 1 42176 1841 1
147 8 53 53 1 1 1 1 45664 1728 1
148 8 53 53 1 1 1 1 45950 2509 0
149 8 53 53 1 1 1 1 47827 2764 1
150 8 53 53 1 1 1 1 49665 1986 1
151 8 53 53 1 1 1 1 49747 1188 1
152 8 53 53 1 1 1 1 53478 2251 1
153 8 53 53 1 1 1 1 54539 1133 0
154 8 53 53 1 1 1 1 63966 1076 0
155 8 53 53 1 1 1 1 64634 2306 0
156 8 53 53 1 1 1 1 65209 2773 0
157 8 53 53 1 1 1 1 66183 1691 1
158 8 53 53 1 1 1 1 67133 3649 1
159 8 53 53 1 1 1 1 76112 910 0
160 8 53 53 1 1 1 1 78285 3815 1
161 8 53 53 1 1 1 1 80419 2238 0
162 8 53 53 1 1 1 1 81135 2101 1
163 8 53 53 1 1 1 1 95208 2968 0
164 8 53 53 1 1 1 1 96712 4998 0
165 8 53 53 1 1 1 1 101635 3610 1
166 8 53 53 1 1 1 1 105077 2457 1
167 8 53 53 1 1 1 1 105373 1877 0
168 8 53 53 1 1 1 1 105910 2630 0
169 8 53 53 1 1 1 1 109037 4119 1
170 8 53 53 1 1 1 1 113538 5424 0
171 8 53 53 1 1 1 1 114935 3774 0
172 8 53 53 1 1 1 1 117437 1365 1
173 8 53 53 1 1 1 1 119386 5279 1
174 8 53 53 1 1 1 1 120051 6781 1
175 8 53 53 1 1 1 1 124218 3745 1
176 8 53 53 1 1 1 1 133837 3390 1
177 8 53 53 1 1 1 1 140558 5168 0
178 8 53 53 1 1 1 1 148193 2210 1
179 8 53 53 1 1 1 1 150120 1455 0
```

```
180 8 53 53 1 1 1 1 175166 780 0
181 8 53 53 1 1 1 1 183897 4352 1
182 8 53 53 1 1 1 1 233971 4220 1
183 8 53 53 1 1 1 1 283909 8624 1
184 8 53 53 1 1 1 1 291052 4057 0
185 9 98 49 2 1 1 1 9432 585 1
186 9 98 49 2 1 1 1 12187 326 1
187 9 98 49 2 1 1 1 13545 371 0
188 9 98 49 2 1 1 1 22365 708 1
189 9 98 49 2 1 1 1 23343 782 0
190 9 98 49 2 1 1 1 25163 1107 1
191 9 98 49 2 1 1 1 30245 1326 0
192 9 98 49 2 1 1 1 31631 921 0
193 9 98 49 2 1 1 1 35969 1075 0
194 9 98 49 2 1 1 1 38158 229 0
195 9 98 49 2 1 1 1 45450 1807 0
196 9 98 49 2 1 1 1 47629 1809 0
197 9 98 49 2 1 1 1 52231 2217 1
198 9 98 49 2 1 1 1 53038 2023 0
199 9 98 49 2 1 1 1 53150 2682 0
200 9 98 49 2 1 1 1 53210 1662 1
201 9 98 49 2 1 1 1 53480 1740 0
202 9 98 49 2 1 1 1 54552 1426 1
203 9 98 49 2 1 1 1 60750 2038 0
204 9 98 49 2 1 1 1 60853 2544 0
205 9 98 49 2 1 1 1 63084 2030 0
206 9 98 49 2 1 1 1 66136 1916 0
207 9 98 49 2 1 1 1 66631 1975 1
208 9 98 49 2 1 1 1 67626 2460 1
209 9 98 49 2 1 1 1 67876 3083 0
210 9 98 49 2 1 1 1 69932 1845 0
211 9 98 49 2 1 1 1 71836 1270 0
212 9 98 49 2 1 1 1 72207 2153 0
213 9 98 49 2 1 1 1 73657 2762 1
214 9 98 49 2 1 1 1 78028 1475 1
215 9 98 49 2 1 1 1 78186 3450 1
216 9 98 49 2 1 1 1 80634 4185 0
217 9 98 49 2 1 1 1 84764 2790 0
218 9 98 49 2 1 1 1 85557 4594 1
219 9 98 49 2 1 1 1 88181 5212 1
220 9 98 49 2 1 1 1 99230 2281 0
221 9 98 49 2 1 1 1 104684 3018 1
222 9 98 49 2 1 1 1 119296 3567 1
223 9 98 49 2 1 1 1 119394 3358 1
224 9 98 49 2 1 1 1 127206 8251 1
225 9 98 49 2 1 1 1 133515 1370 0
226 9 98 49 2 1 1 1 141280 6672 1
227 9 98 49 2 1 1 1 155399 5542 0
228 9 98 49 2 1 1 1 156405 2267 0
229 9 98 49 2 1 1 1 160619 9528 1
230 9 98 49 2 1 1 1 174293 1750 1
231 9 98 49 2 1 1 1 181686 7061 1
232 9 98 49 2 1 1 1 184821 4921 1
233 9 98 49 2 1 1 1 192856 2641 0
234 10 108 36 3 1 1 1 5288 224 0
235 10 108 36 3 1 1 1 13102 785 1
236 10 108 36 3 1 1 1 13578 339 0
237 10 108 36 3 1 1 1 16928 707 1
238 10 108 36 3 1 1 1 17507 720 1
239 10 108 36 3 1 1 1 17701 413 1
240 10 108 36 3 1 1 1 17877 819 1
241 10 108 36 3 1 1 1 19740 1332 0
242 10 108 36 3 1 1 1 19798 824 1
243 10 108 36 3 1 1 1 21604 193 1
244 10 108 36 3 1 1 1 21723 1498 1
245 10 108 36 3 1 1 1 22141 639 0
246 10 108 36 3 1 1 1 22380 858 0
247 10 108 36 3 1 1 1 25776 801 1
248 10 108 36 3 1 1 1 26772 1136 0
249 10 108 36 3 1 1 1 30921 900 1
250 10 108 36 3 1 1 1 33197 1520 1
251 10 108 36 3 1 1 1 33317 458 0
252 10 108 36 3 1 1 1 33356 797 0
253 10 108 36 3 1 1 1 38470 1797 1
254 10 108 36 3 1 1 1 38989 317 0
255 10 108 36 3 1 1 1 51918 649 0
256 10 108 36 3 1 1 1 54784 1859 1
257 10 108 36 3 1 1 1 55005 2206 1
258 10 108 36 3 1 1 1 58819 2911 1
259 10 108 36 3 1 1 1 59138 2926 1
260 10 108 36 3 1 1 1 65557 1383 0
261 10 108 36 3 1 1 1 68189 1643 1
262 10 108 36 3 1 1 1 69492 207 0
263 10 108 36 3 1 1 1 72581 2658 1
264 10 108 36 3 1 1 1 81998 4211 1
265 10 108 36 3 1 1 1 85741 4469 0
266 10 108 36 3 1 1 1 102223 6381 1
267 10 108 36 3 1 1 1 114720 6793 0
268 10 108 36 3 1 1 1 157354 9142 1
269 10 108 36 3 1 1 1 159676 6858 1
270 11 96 4 4 4 0 1 678 5 0
271 11 96 4 4 4 0 1 1741 51 0
272 11 96 24 4 4 0 1 4070 151 1
273 11 96 24 4 4 0 1 5384 25 1
```

```
274 11 96 24 4 4 0 1 6161 88 1
275 11 96 24 4 4 0 1 9943 247 1
276 11 96 24 4 4 0 1 11897 681 0
277 11 96 24 4 4 0 1 13509 426 0
278 11 96 24 4 4 0 1 16698 578 1
279 11 96 24 4 4 0 1 16943 185 1
280 11 96 24 4 4 0 1 18089 348 1
281 11 96 24 4 4 0 1 19939 509 0
282 11 96 24 4 4 0 1 21927 289 0
283 11 96 24 4 4 0 1 24372 966 0
284 11 96 24 4 4 0 1 27575 1029 1
285 11 96 24 4 4 0 1 28188 169 0
286 11 96 24 4 4 0 1 28243 1016 1
287 11 96 24 4 4 0 1 29243 657 0
288 11 96 24 4 4 0 1 29460 1146 0
289 11 96 24 4 4 0 1 30122 1028 0
290 11 96 24 4 4 0 1 31083 449 0
291 11 96 24 4 4 0 1 35277 616 1
292 11 96 24 4 4 0 1 35450 978 0
293 11 96 24 4 4 0 1 191550 3263 0
294 12 60 12 5 1 0 1 3990 96 0
295 12 60 12 5 1 0 1 6775 175 0
296 12 60 12 5 1 0 1 6907 41 0
297 12 60 12 5 1 0 1 15723 422 0
298 12 60 12 5 1 0 1 20912 354 1
299 12 60 12 5 1 0 1 23934 397 0
300 12 60 12 5 1 0 1 24742 441 1
301 12 60 12 5 1 0 1 30665 1705 1
302 12 60 12 5 1 0 1 31621 804 0
303 12 60 12 5 1 0 1 34571 923 1
304 12 60 12 5 1 0 1 41262 1943 0
305 12 60 12 5 1 0 1 66897 3113 1
306 13 35 5 7 2 1 1 4756 261 0
307 13 35 5 7 2 1 1 16293 453 0
308 13 35 5 7 2 1 1 17371 447 1
309 13 35 5 7 2 1 1 33085 2177 0
310 13 35 5 7 2 1 1 41119 2130 0
311 14 43 5 8.6 3 1 1 6847 99 0
312 14 43 5 8.6 3 1 1 6887 326 0
313 14 43 5 8.6 3 1 1 7609 106 0
314 14 43 5 8.6 3 1 1 7931 117 1
315 14 43 5 8.6 3 1 1 8988 464 0
316 15 949 25 37.96 2 1 0 4333 201 0
317 15 949 25 37.96 2 1 0 4433 206 0
318 15 949 25 37.96 2 1 0 4454 207 0
319 15 949 25 37.96 2 1 0 4498 209 0
320 15 949 25 37.96 2 1 0 10289 489 1
321 15 949 25 37.96 2 1 0 10735 511 0
322 15 949 25 37.96 2 1 0 11190 532 1
323 15 949 25 37.96 2 1 0 12618 600 1
324 15 949 25 37.96 2 1 0 13136 354 0
325 15 949 25 37.96 2 1 0 13623 367 0
326 15 949 25 37.96 2 1 0 13817 373 0
327 15 949 25 37.96 2 1 0 14616 394 0
328 15 949 25 37.96 2 1 0 14711 397 1
329 15 949 25 37.96 2 1 0 26288 1160 0
330 15 949 25 37.96 2 1 0 26420 1166 0
331 15 949 25 37.96 2 1 0 26551 1172 1
332 15 949 25 37.96 2 1 0 28427 1490 1
333 15 949 25 37.96 2 1 0 28457 1491 1
334 15 949 25 37.96 2 1 0 30251 1452 0
335 15 949 25 37.96 2 1 0 30787 1613 1
336 15 949 25 37.96 2 1 0 30897 1619 0
337 15 949 25 37.96 2 1 0 31498 1390 0
338 15 949 25 37.96 2 1 0 32187 1545 0
339 15 949 25 37.96 2 1 0 34398 1651 1
340 15 949 25 37.96 2 1 0 37233 1788 0
341 16 210 6 35 2 1 1 12744 430 1
342 16 210 6 35 2 1 1 12929 437 1
343 16 210 6 35 2 1 1 89732 3974 1
344 16 210 6 35 2 1 1 96459 4272 1
345 16 210 6 35 2 1 1 113486 5163 1
346 16 210 6 35 2 1 1 143865 6545 0
;
```

www.rti.org/rtipress **RTI Press publication MR-0048-2204**