OPTIMALLY WEIGHTED MEANS IN STRATIFIED SAMPLING

Jiahe Qian, NORC, Bruce D. Spencer, Northwestern University
Jiahe Qian, 1136 S. Washington St., Falls Church, VA 22046

KEY WORDS: Weighting, Survey inference, Trimming, Shrinkage

1. Weights, Problems in Weighting, and Adjustments

This paper develops a minimum mean square error (MSE) estimator of the population mean in the context of stratified simple random sampling. Like the unbiased estimator, our estimator is a weighted average of the sample means for the strata; unlike the classical estimator, our estimator is biased. The justification for using a biased estimator is that it has smaller mean square error than the unbiased estimator and it has only a relatively small bias.

Weighting is usually used to obtain unbiasedness. The cost of bias reduction, however, is often the inflation of variance. Therefore, it has long been common practice to restrict the variability of the weights to prevent variance from becoming excessive, even though the adjustment of the weights introduces bias into the estimators (Kish 1990).

In practice, weights are adjusted by trimming of extreme weights and shrinkage of weights. Trimming may be used to avoid or to minimize the size of extreme weights, by setting pre-specified limits on the size of the weights prior to the computation of weights (Alexander 1978, Hanson 1978, Cox and Grath 1981, Johnson et al. 1987, Potter 1990). Although other criteria or assumptions can be used, these type of methods are usually computation-intensive and hard to employ in practice. Furthermore, these informal methods of adjusting the weights are not generally optimal. Another type procedure of adjustments shrinks weights toward each other to reduce the variability of weights rather than truncating extreme weights (Cohen and Spencer 1991). Instead of restricting ourselves to trimming or shrinking the weights, we will derive adjusted stratum weights that minimize the MSE of the weighted sample mean.

2. Optimal Adjustment of Weights in Stratified Sampling

2.1 Setup and Notations

Consider a set of N units partitioned into L disjoint groups or strata with $N_h > 0$ units in the h^{th} stratum, $h = 1,2,\ldots,L$. Let y_{hi} denote the value of the characteristic of interest for the i^{th} unit in stratum h. A stratified random sample is selected by independent simple random sampling of $n_h > 0$ units from stratum h, $h = 1,2,\ldots,L$. We also use the notation y_i and y_{hi} to denote values of the sampled units; there is no relation between a particular value of the index i in the population and the sample. For stratum h, we define ($\triangleq$ stands for define): the stratum weight $W_h \triangleq N_h/N$ which reflects the stratum population base, the sampling ratio $f_h \triangleq n_h/N_h$, the population mean: $\bar{Y}_h \triangleq 1/N_h \sum_{i=1}^{N_h} y_{hi}$, the sample mean: $\bar{y}_h \triangleq 1/n_h \sum_{i=1}^{n_h} y_{hi}$, the variance of stratum h: $S_h^2 \triangleq 1/(N_h-1)\sum_{i=1}^{N_h}(y_{hi}-\bar{Y}_h)^2$, the sample variance of stratum h: $s_h^2 \triangleq 1/(n_h-1)\sum_{i=1}^{n_h}(y_{hi}-\bar{y}_h)^2$.

Let $E(\cdot)$, $V(\cdot)$ and $Cov(\cdot,\cdot)$ denote expectation, variance and covariance. The variance of $\bar{y}_h$ is $V_h \triangleq V(\bar{y}_h) = (1-f_h)S_h^2/n_h$ and it is unbiasedly estimated by $v_h \triangleq v(\bar{y}_h) = (1-f_h)s_h^2/n_h$.

Matrices are indicated by bold font, like

$$\mathbf{1} \triangleq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \ \mathbf{Y} \triangleq \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \vdots \\ \bar{Y}_L \end{bmatrix}, \mathbf{V} \triangleq \begin{bmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & V_L \end{bmatrix};$$

the matrices of sample estimates are in bold lower case.

Rather than restrict ourselves to trimming or shrinking the weights, we will derive adjusted stratum weights A_h, $h=1,2,\ldots,L$, that minimize the mean squared error of the weighted sample mean $\bar{y}_A \triangleq \sum_{h=1}^{L} A_h \bar{y}_h$, or in its matrix form $\bar{y}_A = \mathbf{A}'\mathbf{y}$ where $\mathbf{A}$ is the vector of A_h's. That is, $MSE(\bar{y}_A) = \min_{\mathbf{a} \in R^L}(E(\bar{y}_a - \bar{Y})^2)$.

2.2 Lagrange multiplier approach and minimum MSE weights

The usual stratification weights, $W_h = N_h/N$, can be interpreted as the proportion of units in the various strata. Of course, they sum to 1. In modifying the weights to reduce the mean square error below that of $\sum W_h \bar{y}_h$, it may be desirable to retain the property that the weights are non-negative and sum to 1: $\sum A_h = 1$. The Lagrange multiplier approach will be used to derive the minimum MSE weights. Let λ be a real number and let

$$\begin{aligned} F(\mathbf{A}) &= V(\bar{y}_A) + \text{bias}^2(\bar{y}_A) + \lambda(\textstyle\sum A_h - 1) \\ &= \mathbf{A}'\mathbf{V}\mathbf{A} + (\mathbf{Y}'\mathbf{A} - \mathbf{Y}'\mathbf{W})^2 + \lambda(\mathbf{1}'\mathbf{A} - 1). \end{aligned} \tag{2.1}$$

We seek to minimize $F(\mathbf{A})$ subject to $\mathbf{1}'\mathbf{A} = 1$. Setting the partial derivative to zero yields,

$$\partial F/\partial \mathbf{A} = 2\mathbf{V}\mathbf{A} + 2\mathbf{Y}\mathbf{Y}'\mathbf{A} - 2\mathbf{Y}\mathbf{Y}'\mathbf{W} + \lambda\mathbf{1} = \mathbf{0}\ . \tag{2.2}$$

Substitute λ in the equation and note that the optimal weights **A** will satisfy

$$(\mathbf{I}-\mathbf{1}\mathbf{1}'/L)(\mathbf{V}+\mathbf{Y}\mathbf{Y}')\mathbf{A}=(\mathbf{I}-\mathbf{1}\mathbf{1}'/L)\mathbf{Y}\mathbf{Y}'\mathbf{W}\ .$$

Since $\mathbf{V}$ is nonsingular, so is $\mathbf{V}+\mathbf{Y}\mathbf{Y}'$. For any $\mathbf{X}\in\mathbf{R}^L$, $(\mathbf{I}-\mathbf{1}\mathbf{1}'/L)\mathbf{X}$ lies in the column space of $\mathbf{I}-\mathbf{1}\mathbf{1}'/L$, hence the equations are consistent. In addition, $\mathbf{I}-\mathbf{1}\mathbf{1}'/L$ is idempotent, so its Moore-Penrose inverse is itself. The general form of the solution satisfies

$$(\mathbf{V}+\mathbf{Y}\mathbf{Y}')\mathbf{A}=(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}'}{L})\mathbf{Y}\mathbf{Y}'\mathbf{W}+\frac{\mathbf{1}\mathbf{1}'}{L}\xi,$$

where $\xi\in\mathbf{R}^L$, or equivalently

$$\mathbf{A}=(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}'}{L})\mathbf{Y}\mathbf{Y}'\mathbf{W}+k(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1},$$

where $k=\mathbf{1}'\xi/L$. We can eliminate k by applying the constraint $\mathbf{1}'\mathbf{A}=1$. Simplifying the formula we get

$$\begin{aligned}\mathbf{A}&=(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}'}{L})\mathbf{Y}\mathbf{Y}'\mathbf{W}+\frac{(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}\mathbf{1}'\mathbf{W}}{\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}}\\&\quad-\frac{(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}'}{L})\mathbf{Y}\mathbf{Y}'\mathbf{W}}{\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}}\\&=(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{Y}'\mathbf{W}+\frac{(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}\mathbf{1}'\mathbf{W}}{\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}}\\&\quad-\frac{(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}\mathbf{Y}'\mathbf{W}}{\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}}.\end{aligned}\qquad(2.3)$$

Theorem 2.1. The minimum MSE solution of equation (2.2) is **A** given by (2.3).

Proof. We have shown that A solves (2.2). To show it has the minimum MSE, first note that $\mathbf{V}>0$ and $\mathbf{y}\mathbf{y}'\geq 0$. Now, the second derivative of equation F(A) is $\partial^2F/\partial\mathbf{A}^2=2(\mathbf{V}+\mathbf{Y}\mathbf{Y}')>0$, showing that F(A), as the function of weights A, is convex in $\mathbf{R}^L$. Under the constraint $\mathbf{1}'\mathbf{A}=1$, F(A), given by (2.1), has a strict local minimum at A given by (2.3). Since the equation has a unique solution under the constraint, A is the unique minimum MSE solution. □

3. Properties of the Minimum MSE Weights

In this section we will study the relationship among the minimum MSE weights, conventional unbiased weights and the minimum variance weights.

3.1 Basic Properties of the Minimum MSE Weights

The conventional weighted estimator of the mean in stratified sampling is $\bar{y}_W\triangleq\sum W_h\bar{y}_h$ with expected value $\bar{Y}_W\triangleq\sum W_h\bar{Y}_h=\bar{Y}$. We may rewrite this estimator in an alternative way, as a weighted average of the sampled units, with weights inversely proportional to the units' selection probabilities, $\pi_h=n_h/N_h$, in their respective strata: $\bar{y}_W=\sum_{h=1}^{L}W_h\bar{y}_h=1/N\sum_{h=1}^{L}\sum_{i=1}^{n_h}y_{h_i}/\pi_h$.

Therefore, a large contribution to the variance can by made by a unit with a value far from the mean and a small selection probability. This is why high fluctuation in weights can cause high variance.

Among all convex combinations of the sampled y's, the minimum variance is attained by $\bar{y}_c=\sum W_{ch}\bar{y}_h$, where $W_{ch}\triangleq(1/V_h)/(\sum 1/V_h)$. The variance of $\bar{y}_c$ is $V(\bar{y}_c)=1/(\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})$. Unless the stratum means are all equal, $\bar{y}_c$ is biased. Note that $\sum W_{ch}=1$ and define $\bar{Y}_c\triangleq\sum W_{ch}\bar{Y}_h$. A kind of generalized variance is

$$S_c^2\triangleq\sum W_{ch}(\bar{Y}_h-\bar{Y}_c)^2\geq 0.$$

The following theorem shows that the expected value of $\mathbf{y}'\mathbf{A}$, $\mathbf{Y}'\mathbf{A}$, is a weighted average of the expected values of $\bar{y}_c$, which minimizes the variance, and $\bar{y}_w$, which minimizes the bias. It reveals the fundamental relationship among estimators of $\bar{y}_A$, $\bar{y}_c$ and $\bar{y}_w$. It is shown that the sample based mean estimator with minimum-MSE stratum weights from a stratified sample is equivalent to a shrinkage estimator. The application of shrinkage estimator is justified as an optimal compromise between a design-unbiased approach and a model-dependent approach in survey data.

Theorem 3.1. Let $\gamma=(n_cS_c^2)/(1+n_cS_c^2)$ with $n_c=\sum 1/V_h$. The expected value of $\mathbf{y}'\mathbf{A}$ may be written as

$$\mathbf{Y}'\mathbf{A}=\gamma\bar{Y}_W+(1-\gamma)\bar{Y}_c\ ,\qquad(3.1)$$

$$\text{Bias}(\bar{y}_A)=\mathbf{Y}'\mathbf{A}-\mathbf{Y}'\mathbf{W}=(1-\gamma)(\bar{Y}_c-\bar{Y}_W)\qquad(3.2)$$

Proof. Let $\mathbf{1}_V\triangleq\mathbf{V}^{-1/2}\mathbf{1}$ and $\mathbf{Y}_V\triangleq\mathbf{V}^{-1/2}\mathbf{Y}$. Application of matrix algebra yields the following equations.

$$n_cS_c^2=\mathbf{Y}_V'\mathbf{Y}_V-n_c\bar{Y}_c^2.\qquad(3.3)$$

$$\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{1}=\frac{\mathbf{1}_V\mathbf{1}_V'}{1+\mathbf{Y}_V'\mathbf{Y}_V}(1+n_cS_c^2)\ .\qquad(3.4)$$

$$\mathbf{1}'(\mathbf{V}+\mathbf{Y}\mathbf{Y}')^{-1}\mathbf{Y}=\frac{\mathbf{1}_V'\mathbf{Y}_V}{1+\mathbf{Y}_V'\mathbf{Y}_V}\ .\qquad(3.5)$$

Substituting formulas (3.4) and (3.5) into (2.3) of the minimum MSE weights, we may verify that

$$\mathbf{Y}'\mathbf{A}=\left[\frac{\mathbf{Y}_V'\mathbf{Y}_V+\mathbf{Y}_V'\mathbf{Y}_Vn_cS_c^2-n_c\bar{Y}_c^2}{1+n_cS_c^2}\right]\frac{\bar{Y}_w}{1+\mathbf{Y}_V'\mathbf{Y}_V}+\frac{\bar{Y}_c}{1+n_cS_c^2}.$$

Using (3.3), we obtain (3.1).

The formula of bias in (3.2) is obvious. □

The form of (3.1) shows that, although $\mathbf{y}'\mathbf{A}$ itself is not a shrinkage estimator, its expected value equals a Stein-type weighted average. In other words, $\mathbf{y}'\mathbf{A}$ has the same bias as a Stein estimator represented by the right side of equation (3.2). However, $\mathbf{y}'\mathbf{A}$ has the smallest MSE, which is different from the MSE of any compromise or other weighted estimates satisfying $1'W=1$. The following result is proved in Qian (1993).

Theorem 3.2. For γ as in Theorem 3.1,

$$V(\bar{y}_A) = A'VA = \gamma(1-\gamma)(\bar{Y}_c - \bar{Y}_w)^2 + 1/n_c. \quad (3.6)$$

$\square$

It follows that

$$MSE(\bar{y}_A) = (1-\sigma)(\bar{Y}_c - \bar{Y}_w)^2 + 1/n_c. \quad (3.7)$$

Define shrinkage weights

$$W_A \triangleq \gamma W + (1-\gamma)W_c \ , \quad (3.9)$$

and note that any weights on line from A to W_A can be expressed as $W^* = \alpha A + (1-\alpha)W_A$, where $\alpha \in [0,1]$. The expectation of the weighted mean under weights W^* is the same as that under the minimum MSE weights A:

$$\begin{aligned} Y'W^* &= Y'(\alpha A + (1-\alpha)W_A) \\ &= \alpha Y'A + (1-\alpha)Y'W_A \\ &= Y'A, \end{aligned}$$

leading to the result that every weighted mean under weights on the line from A to W_A has the same bias:

Theorem 3.3. The weights W^* on line from A to W_A satisfy $Y'W^* = Y'A$. $\square$

3.2 Empirical Estimates for the Minimum MSE Weights and Related Estimators

The derivation of the matrix A has been of theoretical interest, but in practice A depends on unknown moments and cannot be calculated exactly. Using moment estimators, however, we may derive an estimate of A, say $\hat{A}$, which has a form like (2.3). Although the form of $\hat{A}$ is relatively complex, the formula for the estimator of the mean, $\bar{y}_A = y'\hat{A}$, is relatively simple:

$$\begin{aligned} y'\hat{A} &= \frac{1}{1+\hat{n}_c s_c^2}\bar{y}_c + \frac{\hat{n}_c s_c^2}{1+\hat{n}_c s_c^2}\bar{y}_W \\ &= (1-\hat{\gamma})\bar{y}_c + \hat{\gamma}\bar{y}_W, \end{aligned}$$

where we correspondingly define $\hat{n}_c \triangleq \sum 1/v_h, w_{ch} \triangleq 1/\hat{n}_c(1/v_h)$, $\bar{y}_c \triangleq \sum w_{ch}\bar{y}_h$, $s_c^2 \triangleq \sum w_{ch}(\bar{y}_h - \bar{y}_c)^2$, and $\hat{\gamma} \triangleq \hat{n}_c s_c^2/(1+\hat{n}_c s_c^2)$. Notice that the sample-based mean estimator with minimum-MSE stratum weights from a stratified sample has the form of a shrinkage estimator. The application of shrinkage estimator can be viewed as an optimal compromise between a design-unbiased approach ($\bar{y}_W$) and a model-dependent approach ($\bar{y}_c$).

In view of (3.6), a first-order estimator of the variance of $y'\hat{A}$ may be taken

$$\hat{A}'v\hat{A} = \hat{\gamma}(1-\hat{\gamma})(\bar{y}_c - \bar{y}_W)^2 + 1/\hat{n}_c.$$

We may estimate the bias by

$$y'\hat{A} - y'W = (1-\hat{\gamma})(\bar{y}_c - \bar{y}_W).$$

4. Results of Computer Simulation

A computer simulation was employed to assess the mean estimates with empirical weights studied in this paper. These empirical weights are conventional weights W, minimum variance weights W_c, minimum MSE weights W_A, and the optimal shrinkage weights W_β proposed by Cohen and Spencer (1991). (Cohen and Spencer defined a compromise between an unbiased estimator $\bar{y}_w$ and another, "model-based," estimator $\bar{y}_m$ is $\bar{y}_\beta = \beta\bar{y}_w + (1-\beta)\bar{y}_m$, where

$$\beta \triangleq \begin{cases} A_m/(A_u + A_m) & \text{if } A_m \geq 0 \text{ and } A_u \geq 0 \\ 1 & \text{if } A_u < 0 \\ 0 & \text{otherwise} \end{cases}$$

with $A_m \triangleq V(\bar{y}_m) + \mathrm{Bias}^2(\bar{y}_m) - \mathrm{Cov}(\bar{y}_m, \bar{y}_w)$ and $A_u \triangleq V(\bar{y}_w) - \mathrm{Cov}(\bar{y}_m, \bar{y}_w)$. An estimator $\bar{y}_\beta$ is obtained by substitution of sample moments for A_m and A_u.)

We drew stratified samples of U.S. counties and calculated alternative estimates of 1989 per capita income in the U.S., as measured in the 1990 census. The target value is $14,420, the national per capita income obtained from 1990 census data. The population was clustered into 3141 counties, and the counties were stratified into 4 regions: with 217, 1055, 1425, and 444 respectively. For a fixed sample size and sample allocation, we select a stratified simple random sample of clusters (Cochran 1977, 249) and repeated the sampling procedure 4150 times. For each sample, we can calculated each four mean estimates separately based on different types of weights. We calculated the empirical distributions of the estimators over the 4150 samples as a means of comparing their sampling distributions. Two sample sizes were used in the simulation, 160 and 2199, and proportional allocation was used in each case. The results of simulation are listed in the table.

For the smaller sample size (160), the biased estimators $\bar{y}_A$ and $\bar{y}_\beta$ have smallest RMSE, although their biases are small. In fact, the estimator $\bar{y}_\beta$ has smaller RMSE than $\bar{y}_A$. This shows that, although the RMSE of $\bar{y}_A$ is smaller than the RMSE of $\bar{y}_\beta$, this superiority need not hold when the weight matrix is estimated from the data. Thus, for the larger sample size (2199), $\bar{y}_A$ has smaller RMSE than $\bar{y}_\beta$. The payoff from biased estimators decreases as sample size increases (and bias contributes a relatively larger share to RMSE), and the reduction in RMSE provided by $\bar{y}_A$ compared to the usual unbiased estimator $\bar{y}_W$ is slight. Owing to the estimated weight matrix W_β, the estimator $\bar{y}_\beta$ actually has larger RMSE than $\bar{y}_W$. For both sample sizes, the "minimum variance" estimator based on W_ϵ has the largest RMSE; again, as a result of estimation of the weight matrix, its variance higher than $\bar{y}_\lambda$ for the smaller sample size and higher than $\bar{y}_\beta$ for both sample sizes. (Significance tests show that the number of replications was large enough that the differences among the moments are not due to the size of the simulation.)

Simulation Results
(based on 4150 replications; moments in thousands)

Weights		Sample Size	
		160	2199
W	BIAS	0.00	0.00
	SE	1.96	.64
	RMSE	1.96	.64
W_λ	BIAS	-.36	-.07
	SE	1.80	.63
	RMSE	1.84	.63
W_β	BIAS	-.52	-.15
	SE	1.73	.61
	RMSE	1.81	.65
W_ϵ	BIAS	-1.86	-.63
	SE	1.79	.65
	RMSE	2.58	.90

6. Conclusions

In this paper, we have derived adjusted stratum weights that minimize the MSE of the weighted sample mean. These methods may be desirable when the reduction in MSE is appreciable and the increase in bias is relatively small. Although not discussed here, generalizations of the methods used here lead to minimum MSE weights of a mean estimator subject to constraints on the amount of bias. The empirical estimates of minimum MSE weights and their related estimators are consistent and asymptotically normal. (Qian, 1993).

An important feature of estimators with minimum MSE is that they are design-consistent. This paper has shown that the empirical mean estimator with minimum MSE weights from a stratified sample are equivalent in expectation to a Stein-type estimator. This connection reveals the relationship among estimates based on minimum MSE weights, minimum variance weights, and the familiar unbiased weights (reciprocals of selection probabilities). In particular, the Stein-type estimator is justified as an optimal compromise between a design-unbiased approach and a model-dependent approach to survey data. An inspection of the properties of the weights shows that bias classifies them into equivalence classes that form lines on a plane defined by the weights.

Although our analysis focuses on the field of sampling, especially stratified sampling, the conclusions and methodology can be extended to other fields, such as panel data analysis, etc. The results have the potential to improve estimates as well as to provide guidance in trimming or shrinking weights.

References

Alexander, Charles H.1986. "The Present consumer Expenditure Surveys Weighting Method." *Papers for Conference Sponsored by Bureau of Labor Statistics*, 1-32.

Cochran, William G. 1977. *Sampling Techniques*, 3rd ed. John Wiley & Sons.

Cohen, Theodora, and Bruce Spencer. 1991. "Shrinkage Weights for Unequal Probability Samples." *Proceedings of the Section on Survey Research Methods*, 625-630.

Cox, Brenda G., and D. S. Mcgrath. March 1981. "An Examination of the Effects of Sample Weight Truncation on the Mean Square Error of Survey Estimates." Presented at Biometrics Society ENAR Meeting, Richmond,VA.

Haberman, Shelby J. 1990. "Mathematical Statistics." *Lecture Notes*, Northwestern University.

Haberman,Shelby J.1975."How much do Gauss-Markov and Least Square estimates differ? A Coordinate-free Approach." *Annals of Statistics* **3** 982-990.

Hansen, Morris H., William G. Madow, and Benjamin J. Tepping. 1983. "An Evaluation of Model -Dependent and Probability-Sampling Inferences in Sample Surveys." *JASA* **78** 776-779.

Johnson, Eugene, et al. 1987. "Weighting Procedures in Implementing the New Design: The NAEP 1983 -1984." Technical Report. Princeton, N.J.: National Assessment of Educational Progress.

Kish, Leslie. 1990. "Weighting: Why, When and How?" *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 121-130 .

Potter, Frank J. 1989. "Survey of Procedures to Control Extreme Sampling Weights." *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 453-458.

Qian, Jiahe. 1993. "Minimum Mean Square Error Estimation in Stratified Sampling." *Doctoral dissertation*, Northwestern University, Evanston IL.

Spencer, Bruce D. 1985. "Optimal Data Quality." *JASA* **80** 564-573.