

SURVEY OF PROCEDURES TO CONTROL EXTREME SAMPLING WEIGHTS

Frank Potter, Research Triangle Institute
Research Triangle Park, NC 27709

KEY WORDS: trimming

1. Introduction

In survey sampling practice, unplanned or extreme variation in the sampling weights may result from the sample selection procedure, inaccuracies or errors in frame data, the nonresponse compensation procedures, or other sources. This unplanned or extreme variation in sampling weights can result in inflated sampling variances and a few extreme weights can offset the precision gains from an otherwise well-designed and executed survey design.

Unequal sampling weights can result from probability proportional to size sampling, from inaccurate or out-dated sampling information in multi-stage surveys, or from nonresponse compensation procedures. Various weight trimming procedures are currently in use by government and private survey research organizations, the purpose of this article is to describe some of the procedures currently in use for trimming extreme sampling weights.

In practice, several procedures are used to limit or reduce the number and size of extreme sampling weights. The practices and procedures fall into two categories:

1. procedures used to avoid or minimize the number and size of extreme weights; and
2. procedures used to identify, trim, and compensate for extreme sampling weights.

The most notable use of procedures used to avoid or minimize the number and size of extreme weights is the Census Bureau's Current Population Survey (CPS) and the Consumer Expenditure Survey (CES). These procedures are characterized by the specification of limits on component factors of the weights. These limits are specified prior to the computation of the weights and, therefore, are established without inspection of the current final weight distribution from the survey. In the CPS, the Census Bureau limits the size of the noninterview adjustment factor and the first-stage ratio adjustment factor so that extreme weights are less likely to occur (Hanson 1978, Bailar et al. 1978). In the CES, the Census Bureau also sets a limit on an intermediate weighting factor (Alexander 1986). The Census Bureau compensates for this weight trimming by post-stratification.

The second set of procedures include procedures to identify the individual extreme weights as well as to specify limits for trimming of the final sampling weights. In most survey situations, the final adjusted sampling weights distribution is analyzed for extremely large sampling weights. In some of these situations, the survey statistician may impose a trimming strategy for excessively large weights. These trimming strategies generally include a procedure to determine excessive weights and a

method to distribute the trimmed portion of the weights among the untrimmed weights. Because of the weight trimming, the survey statistician will generally expect an increased potential for a bias in the estimate and a decrease in the sampling variance. In sum, a trimming strategy may reduce the sampling variance for an estimate but increase the mean squared error. The ultimate goal of weight trimming is to reduce the sampling variance more than enough to compensate for the possible increase in bias and, thereby, to reduce the mean squared error.

For this paper, I will provide an overview of sample weighting, sources of unequal weights, and describe current procedures.

2. Sources of Unequal Sampling Weights

Unequal sampling weights may result from design features or from unplanned or unexpected occurrences experienced during the study. The unequal weights can be both beneficial or deleterious in the same survey or in repeated samples from the same population. The planned unequal weighting can be beneficial for some estimates and may be deleterious for others. Statistical surveys are often designed for one objective but are used to provide information to address multiple topics, sometimes not directly related to the original objective. In addition, in some samples from a given population, a sample design can result in samples with minor or negligible weight variation and in other samples from the same population, the weights may exhibit substantial variation. Some possible sources of unequal sampling weights are the following:

1. Probability proportional to size (pps) sampling;
2. Disproportionate stratified sampling;
3. Adjustments for differential nonresponse or poststratification;
4. Inaccurate frame information in multi-stage designs;
5. Sampling weights for analysis units defined at intermediate stages in a multi-stage design.

More specifically, for pps sampling and disproportionate stratified sampling, the variation in the sampling can enable increased precision for data corresponding to the underlying models for these sampling designs. However, the variation can result in decreased precision for other data. Adjustments for differential nonresponse or poststratification can also increase the variation in weights, resulting in extreme weights. The ratio adjustment of sampling weights to account for nonresponse generally entails the assignment of sampling units into classes of 'similar' units (Oh & Scheuren 1983). The nonresponse adjusted sampling weights are the product of the sampling weight and the inverse of the weighted or

unweighted response rate within the class. Unequal weights will result from differences in response rates among classes. Similarly, poststratification adjustment of the sampling weight can also result in unequal weights.

The fourth potential source is inaccurate frame information in multi-stage designs. In multi-stage sampling designs (such as area household surveys), when first stage units are selected pps, inaccurate or outdated frame information can result in unequal weights for final stage units.

A fifth source is from sampling weights for analysis units defined at intermediate stages in a multi-stage design. In multi-stage designs when pps selection is used in all but the last stage of sampling, the sampling weights for all but the last stage will generally be unequal by design. Estimates computed at an intermediate stage using data not proportional to the size measure can suffer substantial loss of precision from the unequal weights.

There are other possible sources of variation in the sampling weights. This variation in the sampling weights can be both beneficial or deleterious.

3. Effects of Unequal Weights

Kish (1963) discusses the concept of the design effect (DEFF) for a sample survey. The DEFF is defined as the ratio of the actual variance ($Var_d(Y)$) of a sample to the variance ($Var_s(Y)$) of a simple random sample of the same size. That is,

$$DEFF = Var_d(Y) / Var_s(Y).$$

In Williams, Folsom, and LaVange (1983), the DEFF was partitioned into components representing design features such as multi-stage clustering, without replacement sampling, stratification, and unequal weights. In a less rigorous fashion (assuming constant unit variance), the usual formulation for providing an estimate of the design effect attributable to unequal weighting is given by

$$DEFF = Var_d(\hat{Y}) / Var_s(\hat{Y}) = n \sum w_i^2 / (\sum w_i)^2,$$

where w_i is the sampling weight for the i th unit and n is the sample size.

Analogously, for stratified designs with n_h sample members in stratum h and equal weights in each stratum, then the DEFF for the esti-

mator (\hat{Y}_s) can be represented as (assuming again constant unit variance)

$$\begin{aligned} DEFF &= Var_d(\hat{Y}_s) / Var_s(\hat{Y}_s) \\ &= n \left[\sum_h \sum_{n_h} w_i^2 / (\sum_h \sum_{n_h} w_i)^2 \right] \\ &= n \sum_h n_h w_h^2 / (\sum_h n_h w_h)^2 \end{aligned}$$

where w_h = the sampling weight in stratum h . These representations of the design effect attributable to unequal weighting will be useful in assessing extreme weights.

4. Current Practice and Procedures

In recent years, a number of articles and books (especially the three volume Incomplete Data in Sample Surveys) have addressed the issues of missing data, response errors, weight adjustment, and imputation. However, a literature search and personal contact with various major survey research organizations have identified little documentation of weight trimming procedures currently in use. The general trimming strategies identified are as follows.

A. Procedures to Minimize Number and Size of Extreme Sampling Weights

The Bureau of the Census uses procedures to reduce the variation of sampling weights in the Current Population Survey (CPS) and the Consumer Expenditure Survey (CES) (CPS - Hanson 1978, Bailar et al. 1978, Little 1986b, Scheuren 1986; CES - Alexander 1986). In the CPS methodology report (Hanson 1978) and in Alexander's (1986) discussion of the weighting methods used for Consumer Expenditure Surveys (CES), the sampling weights for the CPS and the CES are controlled by limiting the size of a components of the sampling weight. In the CES, at one stage of the sampling weight computation, the sampling weights are composed of the product of the base weight and a weighting control factor. To paraphrase the rather involved Census Bureau weighting procedure, the base weight is a first-stage weight for a area unit. The weighting control factor takes into account changes to the sampling rates that result (a) from changes in the size of the second stage unit since the last census (for example, substantial growth in a second stage unit) and (b) from deviations from an overall sampling rate. In these areas, subsampling may be required to maintain the desired workload. The weighting control factor takes this subsampling into account. For the CES, Alexander reports that the weighting control factor was arbitrarily limited to a magnitude of 8 until 1984 and since 1985, limited to a value of 4. The excess weight that was 'lost' was accounted for by poststratification in a later step of the weight computation procedure. Alexander states that, in other Census Bureau surveys, similar limitations are set on intermediate weighting factors.

The Census Bureau also limits the size of some noninterview adjustments in some surveys. For the CPS (Hanson 1978), if noninterview weighting classes contain either less than 20 cases or the noninterview adjustment factor is greater than 2.0, then a restricted form of weighting class collapsing is used until these requirements are achieved. The restriction allows for the collapsing across race (white and not-white) but not across area of residence categories. In some instances, these requirements can not be achieved and the Census Bureau limits the weighting class noninterview adjustment to a value of 3.0. In the CPS, the Census Bureau also limits the first-stage ratio adjustment factor to a value of 1.3.

The use of such limits sometimes introduces some debate (Little 1986b, Scheuren 1986). Little (1986b) questions the use of ad hoc cut-

off values. Scheuren (1986), in a discussing Little's paper, indicates that cut-off values are usually developed by considering the cost/benefit tradeoffs between reduced bias for some data items and increased variances for other items. Because of periodic nature of the CPS, CES and other Census Bureau surveys, one can conjecture that the Census Bureau can develop procedures using a historical perspective generally unavailable to survey statisticians conducting one-time surveys.

To provide insight into the effect of these procedures, let us consider only trimming of the weighting control factor in a simplified situation with complete response. For the kth sampled unit denote

w_{bk} as the base weight,

w_{ck} as the untrimmed weighting control factor,

w'_{ck} as the trimmed weighting control factor

$$\text{(that is } w'_{ck} = \begin{cases} w_{ck} & \text{if } w_{ck} \text{ less than 4, or} \\ 4 & \text{if } w_{ck} \text{ greater than 4),} \end{cases}$$

r as the ratio adjustment factor when no trimming is imposed, and

r' as the ratio adjustment factor when trimming is used.

Then for the full trimmed weight, w_{tk} ,

$$w_{tk} = w_{bk} * w'_{ck} * r'$$

and the untrimmed weight, w_{uk} ,

$$w_{uk} = w_{bk} * w_{ck} * r$$

To estimate a total, Y , from a sample of sample of size n , Y_k , $k=1,2,.. n$, the Horvitz-

Thompson estimator, \hat{Y}_t , using the trimmed weights is

$$\hat{Y}_t = \sum w_{tk} * Y_k = \sum w_{bk} * w'_{ck} * r' * Y_k$$

and using the untrimmed weights is

$$\hat{Y}_u = \sum w_{uk} * Y_k = \sum w_{bk} * w_{ck} * r * Y_k$$

The difference between \hat{Y}_t and \hat{Y}_u is

$$\begin{aligned} \hat{Y}_t - \hat{Y}_u &= \sum (w_{tk} - w_{uk}) Y_k \\ &= \sum w_{bk} Y_k (w_{ck} r' - w_{ck} r) \end{aligned}$$

Define $A = \{k: w_{ck} \text{ is not trimmed}\}$ and

$B = \{k: w_{ck} \text{ is trimmed}\}$.

Then

$$\begin{aligned} \hat{Y}_t - \hat{Y}_u &= \sum_{k \in A} w_{bk} Y_k w_{ck} (r' - r) + \\ &\quad \sum_{k \in B} w_{bk} Y_k (4r' - w_{ck} r). \end{aligned}$$

From this representation, the following observations can be made.

1. The w_{ck} will be greater than 4 in areas of high growth. In areas of very high growth, the w_{ck} may be substantially larger than 4.

2. The ratio adjustment factors r and r' will compensate for a portion of the trimmed value. However, the ratio adjustment strata are generally broad classes of the population so the trimmed excess of w_{ck} will be distributed across a broad portion of the sample.

3. By viewing the persons in these high growth areas as a domain, this domain will be underestimated as the result of these procedures.

4. These procedures require the availability of external data for developing the ratio adjustments.

In CPS Design and Methodology Report (Hanson 1978), the Census Bureau describes the potential for bias from some of these procedures.

In summary, the procedures used by the Census Bureau provide various examples of how and where extreme weights can be avoided by limiting factors included in the weights. The Census Bureau sets limits on subsampling weights in areas experiencing growth in number of housing units, redefines weighting classes to limit adjustment factor, sets limits on nonresponse adjustment factors, and sets limits of ratio adjustment factors. Some possible effects of these procedures include those discussed above and the following:

1. redefining and collapsing weighting classes to reduce adjustment factors can result in combining across groups with substantially different response experience;
2. limiting nonresponse adjustment factors imply that persons with relatively low response propensity are not being represented appropriately; and
3. because poststratification is used to account for the weight and weight adjustment factor limitation, persons and households in areas of high response propensity may be over-represented.

Some of these issues related to the nonresponse adjustments may result in negligible bias effect because of the high response rates generally achieved in Census Bureau surveys.

B. Procedures to Trim and Compensate for Extreme Sampling Weights

Procedures for trimming and compensating for this trimming differ on the basis of the amount and type of information used to determine a level of trimming. Three specific strategies were identified: (1) an inspection strategy, (2) a strategy involving the computation of an estimated mean squared error for selected items, and (3) a strategy involving the relative

contribution of extreme weights to the overall variance.

1. Inspection Strategies

Some form of inspection of the weight distribution is generally conducted regardless of whether trimming procedures are planned. At RTI, the sum, mean, variance, coefficient of variation and selected percentiles are usually computed to describe the weight distribution. The coefficient of variation (CV) is a useful descriptive measure because of its relationship to the constant unit variance model for the design effect for unequal weighting (DEFF_w). That is, for a sample of n cases

$$\begin{aligned} \text{DEFF}_w &= 1 + ((n - 1) / n) * (CV)^2. \\ &= 1 + (CV)^2 \text{ for large } n. \end{aligned}$$

In addition to these descriptive statistics, the 25 largest and 25 smallest weights are listed along with components to the weight (such as the initial or prior stage weight, the nonresponse adjustment factor, and post-stratification adjustment factors). This listing generally identifies a sufficient portion of the tails to ascertain the essential characteristics of the weight distribution. The largest 25 weights will include all weights that are likely to be trimmed and the 25 smallest weights can provide an indication of the symmetry of the distribution.

In most cases, the large weights that can effect (increase) the sampling variance can be easily identified because these weights will differ from the other weights by a substantial amount. In some cases, inspection of the listing of the largest 25 weights can identify logical trimming limits.

The problems associated with the inspection strategy are easily apparent. First, the procedure is subjective and the choice of a trimming limit is arbitrary. Second, the effects on the sampling variance and bias of the estimates are unknown. However, this procedure can be implemented both inexpensively and quickly. When data for computing the effect of the trimming may not be available and because of time or funding constraints exist, this procedure may be justified.

2. Estimated Mean Squared Error (MSE)

Trimming

A method used by some of the major survey research organizations is the evaluation of an estimate of the mean squared error for selected data items at various trimming levels to empirically determine the trimming level (Cox & McGrath 1981, Cox 1988, Heeringa 1988). In this procedure for determining cut-off values for weights, the statistician conducts a visual inspection of the distribution of the sampling weights. A set of key data items are identified and an estimate of the mean squared error is calculated for the key data items at different candidate cut-off values. The values of the estimated mean squared error can be plotted versus the cut-off values to determine a cut-off value that achieves adequate reductions in the estimated mean squared error for all or most of

the data items. In this procedure, the trimmed excess is distributed across the untrimmed weights to reproduce the original weight sum.

The assumption underlying this procedure is that for a set of weights and data, a point exists at which the reduction in the sampling variance resulting from the trimming is offset by the increase in the square of the bias introduced into the estimate. The mean squared error (MSE) for an estimator, \hat{Y} , is

$$\text{MSE}(\hat{Y}) = \text{Var}(\hat{Y}) + (\text{Bias}(\hat{Y}))^2.$$

where $\text{Var}(\hat{Y})$ is the sampling variance of the estimator \hat{Y} and

$\text{Bias}(\hat{Y})$ is the bias of $E(\hat{Y})$ relative to the true value of Y .

In the implementation of this procedure, an estimate of the mean squared error, $\text{MSE}(Y_c)$ is computed for each data variable at each cut-off value c . To estimate the $\text{MSE}(\hat{Y}_c)$, first note that

$$\begin{aligned} E(\hat{Y}_c - \hat{Y})^2 &= \text{Var}(\hat{Y}_c) + \text{Var}(\hat{Y}) - \\ &2 \text{Cov}(\hat{Y}_c \hat{Y}) + (E(\hat{Y}_c) - E(\hat{Y}))^2 \\ &= (\text{BIAS}(\hat{Y}_c))^2 + \text{Var}(\hat{Y}_c) + \text{Var}(\hat{Y}) - 2 \text{Cov}(\hat{Y}_c \hat{Y}) \end{aligned}$$

where $\text{Cov}(\hat{Y}_c \hat{Y})$ is the covariance between the two estimates. Therefore, assuming that the estimator using the untrimmed weights (\hat{Y}) is an unbiased estimator of Y , an unbiased estimator of the $\text{MSE}(\hat{Y}_c)$ is given by

$$\hat{\text{MSE}}(\hat{Y}_c) = (\hat{Y}_c - \hat{Y})^2 - \hat{\text{Var}}(\hat{Y}) + 2 \hat{\text{Cov}}(\hat{Y}_c \hat{Y}).$$

The procedure is implemented by repeatedly computing the estimate of the MSE for selected set of data items at differing levels of weight truncation. The estimated MSEs are plotted relative to the cut-off values (for example, the percentage of truncation) to determine reasonable truncation level. Under the assumption of the offsetting effects of the reductions in the sampling variance and increases in the bias, the plots are expected to be U-shaped. The 'optimal' level of truncation is the point that minimizes estimated MSE (i.e., minimizes sampling variance and estimated squared bias) for the set of key data items.

The advantages of this procedure are that the effect of trimming on the estimates and the sampling variances is utilized to determine the extent of trimming and the selection of the trimming value. In addition, the actual design effect can be computed for each data variable at each trimming level. However, the disadvantages of this approach are as follows.

1. This approach is a computer intensive requiring repeated computation of estimates and sampling variances for each data variable at each trimming level.

2. Key data items must be determined and the data available.
3. The trimming levels are determined by trial and error.
4. The plots may not show a single 'best' truncation level for all data items.

If the statistician has the available resources (time, computer resources, and data), this procedure is superior to the inspection strategy.

3. Comparison of the squared weight to the mean squared weight

This procedure uses the comparison of the contribution of each weight to the sampling variance of an estimate by systematically comparing all weights to a value computed from the sum of the squared weights for the sample. If a weight is above the computed value, the weight is assigned this value and the other weights are adjusted to have the new weights sum to the original weight total. The sum of the squared adjusted weights is computed again and used in a second comparison of each individual adjusted weight. The procedure is repeated until all adjusted weights are below or equal the value based on the sum of the adjusted squared weights.

The basis for this procedure can be described as evaluating the contribution of each weighted observation to the overall variance of the weighted estimate. For a sample of size n , let

Y_k denote the observed for the k th unit and

w_k denote the original weight for the k th unit and

assume that the true unit variance (under a superpopulation model) for Y_k is σ^2 . Under a superpopulation model,

$$\text{Var}(w_k Y_k) = w_k^2 \text{Var}(Y_k) = w_k^2 \sigma^2;$$

For an estimate of a total, $\hat{Y} = \sum w_k Y_k$, a similar representation of the variance of \hat{Y} under a superpopulation model is

$$\begin{aligned} \text{Var}(\hat{Y}) &= \text{Var}(\sum w_k Y_k) = \sum w_k^2 \text{Var}(Y_k) \\ &= \sum w_k^2 \sigma^2. \end{aligned}$$

The relative contribution of the variance associated to the k th unit, $\text{Var}(w_k Y_k)$, to $\text{Var}(\hat{Y})$ is

$$\begin{aligned} \text{Var}(w_k Y_k) / \text{Var}(\hat{Y}) &= w_k^2 \sigma^2 / \sum w_k^2 \sigma^2 \\ &= w_k^2 / \sum w_k^2. \end{aligned}$$

In this procedure, the relative contribution is limited to a specific value by comparing the square of each weight to a multiple of the the sum of the squared weights. That is,

$$w_k^2 < K (\sum w_k^2)$$

To take into account the sample size, K is set

as a function of n ,

$$K = c / n.$$

The final form of the algorithm is

$$\begin{aligned} w_k^2 &\leq c \sum w_k^2 / n \text{ or} \\ w_k &\leq (c \sum w_k^2 / n)^{1/2} \text{ or} \\ w_k &\leq K_n \end{aligned} \tag{1}$$

where $K_n = (c \sum w_k^2 / n)^{1/2}$.

The value for c is arbitrary and can be chosen empirically by looking at values of

$$n w_k^2 / \sum w_k^2.$$

In implementing the algorithm, each weight is compared using equation (1). Each weight in excess of K_n is given this value and the other weights are adjusted to reproduce the original weight sum. The sum of squared adjusted weights is computed and each weight is again compared using equation (1). The procedure is performed repeatedly until none of the weights exceed this criterion.

This trimming procedure facilitates the identification of trimming levels through a systematic and relatively objective algorithm. This procedure also provides an indication of the effect of the trimming of the sampling variance because the average of the squared weights is the numerator of the design effect attributable to unequal weighting. That is,

$$\begin{aligned} \text{DEFF}_w &= (S w_k^2 / n) / (S w_k / n)^2 \\ &= (K_n^2 / c) / (S w_k / n)^2. \end{aligned}$$

The primary disadvantage of this procedure is that the survey outcome data are not utilized. The actual effects of the trimming on the sampling variance and bias for the estimates are unknown. Complete reliance on the algorithm to select the trimming value may impose more extensive trimming than can be justified.

This procedure has been used for the sampling weights of the National Assessment Educational Progress (NAEP) for over 10 years. The procedure was alluded to in an RTI NAEP methodology report (Benrud, et al. 1978). A variation of this procedure that used some external data has been reported in a more recent NAEP methodology report (Johnson, et al. 1987). The initial version of this procedure is attributed to John Tukey but no specific reference has been found.

5. Summary and Discussion

In survey sampling practice, unplanned or extreme variation in the sampling weights occasionally occur. This variation may result from the sample selection procedure, inaccuracies or errors in frame data, the nonresponse compensation procedures, or other sources. This unplanned or extreme variation in sampling weights can result in inflated sampling variances and a few extreme weights can offset the precision gains from an otherwise well-designed and executed survey design. Various weight trimming procedures are currently in use by government and private survey research

organizations. In addition, a procedure using empirical Bayes methods was recently proposed to smooth response adjusted sampling weights (Little 1986a).

Of the procedures described in this paper, no single procedure offers a comprehensive strategy to identify extreme weights and to assess the effect of trimming of the extreme weights. One should be cautious in the use of procedures that do not assess the effect of trimming on the bias and sampling variances on estimates. However, there is often considerable pressure to complete the computation of the survey data analysis weights before the associated data files are available. In such surveys, the survey data analysis weights may be required to compute the key analysis variables when, for example, weighted factor analyses and weighted distributions of the respondent data are used to develop indices and scores.

The procedures to control the number and size of extreme weights used by the Census Bureau are useful in the situation of repeated surveys of the same population. For these procedures, the historical perspective and current population counts for poststratification are essential. The historical perspective is needed to determine the levels of trimming and the population counts must be used to compensate for the trimming. In many surveys, one or both of these items may not be available.

In one-time surveys, when data are available, a combination of the three trimming strategies provides the most comprehensive approach. A descriptive analysis and inspection of the weight distribution is an essential component. The use of the procedure that evaluates the contribution of each weight to the sum of the squared weights can facilitate the search for candidate trimming levels. When data are available, the statistician should conduct an assessment of the effect of trimming on estimates and sampling variances.

REFERENCES

- Alexander, C.H. (1986). The present consumer expenditure surveys weighting method. In *Population Controls in Weighting Sample Units*, papers from conference sponsored by Bureau of Labor Statistics. March 1986. pp. 1-32. Washington, D.C.: Government Printing Office.
- Bailar, B.A., Bailey, L. & Corby, C. (1978). A comparison of some adjustment and weighting procedures for survey data. in *Survey Sampling and Measurement*. Ed. N. K. Namboodiri. pp. 175-198. New York: Academic Press.
- Benrud, C.H., et al. (1978). Final report on national assessment of educational progress: sampling and weighting activities for assessment year 08. Prepared for National Assessment of Educational Progress. Research Triangle Park, N.C.: Research Triangle Institute.
- Cox, B.G. & McGrath, D.S. (1981). An examination of the effect of sample weight truncation on the mean square error of survey estimates. Presented at Biometrics Society ENAR meeting, Richmond, VA. March 1981.
- Cox, B.G. (1988). Weighting survey data. Submitted to the *American Statistician*.
- Heeringa, S. (1988) Personal communication. University of Michigan Survey Research Center.
- Hanson, R.H. (1978). The current population survey, design and methodology. Technical Paper 40. Washington D.C.: Bureau of the Census.
- Johnson, E.G. et al. (1987). Weighting procedures. in *Implementing the new design: The NAEP 1983-1984 technical report*. by A. E. Beaton. pp. 493-504. Princeton, N.J.: National Assessment of Educational Progress.
- Kish, L. (1963) *Survey sampling*. New York: John Wiley & Sons, Inc.
- Little, R.J.A. (1986a) Survey nonresponse adjustments for estimates of means. *International Statistical Review* (54), pp. 139-157.
- Little, R.J.A. (1986b) Missing data in Census Bureau surveys. In *Proceedings of the Second Annual Research Conference*, March 1986. Washington, D.C.: Bureau of the Census. pp. 442-454.
- Oh, H.L. & Scheuren, F.S. (1983). Weighting adjustment for unit nonresponse. In *Incomplete Data in Sample Surveys*, 2, Eds. W.G. Madow, I. Oklin, and D.B. Rubin, pp. 143-184. New York: Academic Press.
- Scheuren, F. (1986) Discussion of "Missing data in Census Bureau surveys". In *Proceedings of the Second Annual Research Conference*, March 1986. pp. 455-462. Washington, D.C.: Bureau of the Census.
- Williams, R. L., Folsom, R. E., & LaVange, L. M. (1983). The implications of sample design on survey analysis. in *Statistical Methods and the Improvement of Data Quality*. Ed. T. Wright. pp. 267-295. New York: Academic Press.