Simultaneous Calibration and Nonresponse Adjustment

Eric V. Slud * Yves Thibaudeau †

Abstract
Single and joint inclusion probabilities are generally known for complex survey designs up to the point where survey weights are modified due to nonresponse and population controls. Best practice by sophisticated survey practitioners generally includes weight modifications, first by calibration, ratio adjustment or raking to correct for nonresponse, next by further steps to impose population survey controls; and often, by final steps involving weight truncation or cell-collapsing to constrain the modified weights, usually so that the largest and smallest weights do not differ by more than a designated multiplicative factor. These adjustments are sometimes made in successive stages, the order of which may differ from one survey to another. In this article, generalized-raking calibration methodology is adapted to allow all of these adjustments to be accomplished in a single stage, after which linearization-based large-sample variance formulas are available.

Key Words: consistency, inclusion probabilities, Lagrange multipliers, linearized variance, objective function, population controls, weight adjustment.

This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau.

1. Introduction

Survey weights in large complex surveys are frequently modified from their designed values, for at least three reasons: to correct for nonresponse, to compensate for frame deficiencies by enforcing population controls (often by demographic categories), and to prevent the adjusted unit weights from being too different from one another. As a result of these modifications, there are usually no meaningful joint inclusion probabilities when complex survey results are analyzed.

Modifications to survey weights are generally applied in multiple stages, and it is fair to say that in practice, the effect of propagating early-stage modifications through later adjustments is poorly understood. Moreover, the later adjustments (particularly the final population controls) are often repeated after trimming or compressing the most extreme weights until controls and moderate weights are achieved simultaneously. Generally, only the final population controls and weight-trimming criteria are imposed rigidly, with no assurance that the criteria used to adjust at earlier stages still hold at the final stage.

This paper begins by summarizing the existing methods to correct survey data both for nonresponse and population controls, while retaining overall bounds on the weights. While some theoretically based methods do exist for enforcing two out of the three of these types of weight constraints, there do not seem to be methods which simultaneously incorporate all of them. A new framework is presented which handles all of these weight adjustments *simultaneously* in a single stage, which

*Census Bureau, Stat. Res. Div., 4600 Silver Hill Road Rm 5K004, Washington DC 20233-9100, & Math. Dept., Univ. of Maryland, College Park, MD 20742

†Census Bureau, Stat. Res. Div., 4600 Silver Hill Road Rm 5K105, Washington DC 20233-9100

justifies linearized variance formulas for survey total estimates based on the adjusted weights, and which allows a tuning parameter to place more or less weight on initial nonresponse adjustment while strictly enforcing population controls.

1.1 Background Literature

There is a large literature on construction and modification of survey weights, much of which has been absorbed into standard survey methodology texts, like that of Särndal et al. (1992). The most important theoretical contributions include:

• nonresponse adjustment by cell-based ratios or raking (Oh and Scheuren 1983), or by models fitted by model-assisted 'pseudo-likelihood' (Kim and Kim 2007);
• weight modification via calibration leading up to generalized raking in Deville and Särndal (1992) and Deville, Särndal & Sautory (1993), papers which established linearized variance formulas for weighted survey totals;
• linear regression-based approaches to nonresponse weight adjustment surveyed in Fuller (2002), treated more fully in the monograph of Särndal and Lundström (2005) which discusses simultaneous calibration to benchmarks or controls along with calibrated nonresponse adjustment;
• (single-stage) methods combining weight-truncation with weight-adjusted calibration as in Singh and Mohl (1996) and Théberge (2000);
• methods for handling informative nonresponse (Sverchkov and Pfeffermann 2004, Chang and Kott 2008);
• nonresponse adjustment followed by a separate calibration stage, e.g., Yung and Rao (2000) treat jackknife variance estimation in such a setting, and Särndal and Lundström (2005) give design-plus-pseudo-randomization-based variance formulas.

1.2 Notation and Assumptions

Consider a sample survey with a frame $\mathcal{U}$ from which a probability sample $\mathcal{S}$ is drawn according to a plan with known single and double inclusion probabilities π_k, π_{kj}, for $k, j \in \mathcal{U}$. Assume that the total $Y = t_y = \sum_{k \in \mathcal{U}} y_k$ of a scalar attribute is of primary interest, and that $(y_k, x_k, \; k \in \mathcal{S})$ is (potentially) observable, i.e., the sample data include the auxiliary p-dimensional vector x_k. This setting corresponds to the *InfoS* sampling framework of Särndal and Lundström (2005), with auxiliary data available at sample but not frame level.

Assume that each sampled individual in the survey decides independently whether or not to respond. Without loss of generality, denote by r_k for all $k \in \mathcal{U}$ the indicator which is 1 if the k'th individual *would* have responded if sampled, and assume that these random variables are independent of each other and of the sample selection mechanism: this is the *pseudo-randomization* model of Oh and Scheuren (1983). (However, there certainly are surveys where this assumption could be applied only with 'individuals' defined as households.) The observable data are now taken to be $(y_k \cdot r_k, r_k, x_k, \; k \in \mathcal{S})$. No restriction is placed on the probabilities $P(r_k = 1) = \rho_k$ with which individual units respond. These quantities must be estimated in order to adjust weights for nonresponse, and this is typically done either by ratio-adjustment and raking (Oh and Scheuren 1983) or by using a working generalized-linear parametric model $1/\rho_k = \kappa(\lambda' x_k), \; k \in \mathcal{U}$ (Kim and Kim 2007), where λ is a p-dimensional parameter vector which is estimated from sample data through the solution $\hat{\lambda}$ to an estimating equation. The most important example of such a working nonresponse model is the case treated in this paper,

$$1/\rho_k \;\equiv\; (Er_k)^{-1} \;=\; \kappa(\lambda' x_k) \;=\; 1 + \lambda' x_k \tag{1}$$

This model motivates the estimation of ρ_k through $\hat{\lambda}$ defined from a nonresponse-adjustment constraint equation (Särndal and Lundström 2005) which requires that modified weights $w_k \equiv r_k \, w_k^o / \hat{\rho}_k$ defined from initial weights $w_k^o = 1/\pi_k$ satisfy

$$\sum_{k \in \mathcal{S}} r_k \, w_k \, \mathbf{x}_k \;=\; \sum_{k \in \mathcal{S}} w_k \, \mathbf{x}_k \;=\; t^*_{\mathbf{x}} \tag{2}$$

The constrained totals $t^*_{\mathbf{x}}$, which may not always be close to the true frame totals, might arise from a survey or census believed to be larger or more accurate than the current survey; but much more often, these totals are themselves estimated either from the same survey or another survey of comparable size with the same frame.

Nonresponse adjustment $w_k \;=\; r_k \, w_k^o / \hat{\rho}_k \;=\; r_k \, w_k^o \, (1 + \hat{\lambda}' \mathbf{x}_k)$ is often treated as a distinct weight-adjustment stage, providing input to further weight modification stages. A special case of the linear-calibration weight adjustment is the standard ratio adjustment derived from a set $C_1, \ldots, C_K$ of *adjustment cells* partitioning the frame $\mathcal{U}$, where the components of $\mathbf{x}_k$ are defined by the indicator variables $x_{k,j} = I_{[k \in C_j]}$.

The results of surveys designed to estimate totals and ratios of totals are often reported after controlling total numbers of units within designated population cells to be equal to the totals found in a more comprehensive survey or (updated) census, generally through constraints on final weights

$$\sum_{k \in \mathcal{S}} r_k \, \hat{w}_k \, \mathbf{z}_k \;=\; t^*_{\mathbf{z}} \tag{3}$$

Here $t^*_{\mathbf{z}}$ is a known vector approximating the frame total $t_{\mathbf{z}}$, for a vector $\mathbf{z}_k = (z_{1k}, \ldots, z_{qk})$ of survey variables defined for each unit $k \in \mathcal{U}$. The constraint (3) is imposed on any system of survey weights $\{\hat{w}_k\}_{k \in \mathcal{S}}$, however obtained — by modifications for nonresponse, population controls, and weight compression or truncation — starting from a designed system $w_k^o \;=\; 1/\pi_k$ of inverse inclusion probabilities. The final weights $\hat{w}_k$ are ultimately used in estimating population totals of survey variables y_k, $k \in \mathcal{U}$, by weighted totals

$$\hat{t}_{y,adj} \;=\; \sum_{k \in \mathcal{S}} r_k \, \hat{w}_k \, y_k \tag{4}$$

2. A New Weight-Adjustment Framework

In the present research, nonresponse adjustment, population-control calibration, and weight-truncation are done in a single step. This step general calculates iteratively the solution to an estimating equation, optimizing a single objective function. The framework is similar to that of Deville and Särndal (1992) and Deville, Särndal and Sautory (1993), in keeping a loss function quantifying weight changes as small as possible. Initial weights $w_k^o = 1/\pi_k$ arise from the design inclusion probabilities, and the transition to final weights is viewed as $\{w_k^o\} \mapsto \{w_k\} \mapsto \{\hat{w}_k\}$, with only the final weights appearing in the survey estimates (4), but with the two sets $\{w_k\}$, $\{\hat{w}_k\}$ of survey weights defined simultaneously to obey respective constraints (2) and (3), both contributing to an aggregated loss function. The auxiliary weight difference $w_k - w_k^o$ can be interpreted as the nonresponse adjustment component of the overall weight-modification $\hat{w}_k - w_k^o$.

The three systems of weights $\{w_k^o\}_k$, $\{w_k\}_k$, $\{\hat{w}_k\}_k$ are related through the desire to minimize simultaneously the losses

$$\sum_{k\in\mathcal{S}} r_k\, w_k^o\, G_1(w_k/\, w_k^o \,-\, 1) \qquad \text{and} \qquad \sum_{k\in\mathcal{S}} r_k\, w_k^o\, G_2((\hat{w}_k - w_k)/w_k^o)$$

where each of $G_1(z)$, $G_2(z)$ is a convex loss-function which is locally of the form $z^2/2$ plus a term of smaller order (bounded by a constant times z^3) near $z = 0$. The intermediate and final modified weights $\mathbf{w} = \{w_k\}_{k\in\mathcal{S}}$ and $\hat{\mathbf{w}} = \{\hat{w}_k\}_{k\in\mathcal{S}}$ are defined as positive only for indices $k \in \mathcal{S}$ for which $r_k = 1$. They are found together, subject to the constraints (2) and (3), i.e. to the the combined constraint

$$\sum_{k\in\mathcal{S}} r_k \begin{pmatrix} w_k\, \mathbf{x}_k \\ \hat{w}_k\, \mathbf{z}_k \end{pmatrix} = \begin{pmatrix} t^*_{\mathbf{x}} \\ t^*_{\mathbf{z}} \end{pmatrix} \tag{5}$$

by the objective-function minimization

$$\min_{\mathbf{w},\,\hat{\mathbf{w}}} \sum_{k\in\mathcal{S}} r_k\, w_k^o \Big\{ G_1(\frac{w_k}{w_k^o} - 1) + \alpha\, G_2(\frac{\hat{w}_k - w_k}{w_k^o}) + Q(\frac{\hat{w}_k}{w_k^o}) \Big\} \tag{6}$$

where $\alpha > 0$ is a constant chosen by the statistician and Q is a convex and piecewise smooth penalty term which is nonzero only for large or small weight ratios, and enters this single optimization step to enforce *weight-truncation* or *restricted weights* as in Singh and Mohl (1996) or Théberge (2000). The most important instances of (6) will have $G_1(z) = G_2(z) = z^2/2$ — which we assume from now on — and Q a piecewise smooth function such that $Q(z) \equiv 0$ on an interval $[c_1, c_2]$, for fixed constants $0 < c_1 < 1 < c_2 < \infty$, and $Q(z)$ is large when $\max(c_1 - z,\, z - c_2)$ is positive and not very small. We assume that when $Q \not\equiv 0$, Q satisfies $Q'(-\infty) = -\infty$ and $Q'(\infty) = \infty$.

In (6), the strict convexity of the objective-function implies that the weights $\mathbf{w}$, $\hat{\mathbf{w}}$ have a unique optimal solution. In several special and limiting cases, the solution relates simply to existing methods.

(Case 1. Full-response, $r_k \equiv 1$, and $t^*_{\mathbf{x}} = \hat{t}_{\mathbf{x},\pi}$). Here the $\mathbf{w}$ weights are unconstrained, and (6) becomes a 'generalized raking' problem with penalized weights, as in Singh and Mohl (1996) and Théberge (2000). If also $Q \equiv 0$, then the final weights $\hat{w}_k$ coincide with the calibrated 'g' weights arising in generalized regression (Särndal et al. 1992, Deville and Särndal 1992) subject to (3).

(Case 2. $\alpha \to \infty$) If α in (6) is taken very large, for fixed Q, then the limiting systems $\mathbf{w}$ and $\hat{\mathbf{w}}$ of weights are identical, i.e., in the large-α limit, $\hat{w}_k \equiv w_k$ for all $k \in \mathcal{S}$ such that $r_k = 1$ (and $\hat{w}_k = w_k = 0$ for all other indices $k \in \mathcal{U}$).

When α is large, the weight optimization problem (6) approximates the problem of finding $\{\hat{w}_k \,:\, k \in \mathcal{S},\, r_k = 1\}$, to solve

$$\min_{\hat{\mathbf{w}}} \sum_{k\in\mathcal{S}} r_k\, w_k^0\, \{\frac{\alpha}{2}\Big(\frac{\hat{w}_k - w_k^o}{w_k^o}\Big)^2 + Q(\frac{\hat{w}_k}{w_k^o})\} \quad \text{such that} \qquad \sum_{k\in\mathcal{S}} r_k\, \hat{w}_k \begin{pmatrix} \mathbf{x}_k \\ \mathbf{z}_k \end{pmatrix} = \begin{pmatrix} t^*_{\mathbf{x}} \\ t^*_{\mathbf{z}} \end{pmatrix}$$

This is a pure linear calibration problem which, although not explicitly considered by Särndal and Lundström (2005), falls directly into the framework of that book.

(Case 3. $Q \equiv 0$.**)** When there is no penalty Q for extreme weight-ratios, the optimal weights of formula (7) below resemble a ridge-regression form of linearly calibrated weights, with the population control condition (3) holding precisely but not (2). However, the particular form (7) has not arisen before.

2.1 Combined Nonresponse Adjustment and Calibration

Many large-scale complex surveys, such as the Census Bureau's American Community Survey (ACS) and Survey of Income and Program Participation (SIPP), are analyzed by first adjusting for nonresponse by a cell-based ratio or raking method and then later by imposing population controls. (In Census Bureau surveys, those controls usually require that population totals in certain demographic and geographic categories match those of the demographically updated decennial census.) This two-step approach is implicit in much of the survey sampling journal literature, and explicit in some sources, such as Yung and Rao (2000) and Särndal and Lundström (2005, Ch. 8 & Sec. 11.4 on 'Two-Step Methods'), which treat variance estimation at a realistic level of complexity.

The method of weight adjustment proposed here, in (6), does not exactly reproduce the known two-step method in any limiting case. Instead, the asymmetric role of the nonresponse adjustment contraints (2) and population controls is made explicit through choice of the parameter α. In accord with current practice, the population controls are required to hold exactly for the final weights actually used in survey estimation. If (2) is to hold at least approximately for w_k replaced by $\hat{w}_k$, then α should be chosen large. When that is done, and there is no penalty term Q in (6), then the proposed method (**Case 2** above) is a simultaneous calibration in the spirit of Särndal and Lundström (2005) or Deville and Särndal (1992) which relaxes the $\mathbf{x}$ constraints in (5).

3. Numerical Algorithms for the Weights

The simultaneous adjustment and calibration step described above is the minimization subject to (5) over $(\lambda, \mu, \mathbf{w}, \hat{\mathbf{w}})$ of

$$\sum_{k \in \mathcal{S}} r_k \Big[\frac{(w_k - w_k^o)^2}{2\, w_k^o} + \alpha\, \frac{(\hat{w}_k - w_k)^2}{2\, w_k^o} + w_k^o\, Q(\frac{\hat{w}_k}{w_k^o}) - \hat{w}_k\, \mu' \mathbf{z}_k - w_k\, \lambda' \mathbf{x}_k \Big] + \mu' t_{\mathbf{z}}^* + \lambda' t_{\mathbf{x}}^*$$

where $\lambda \in \mathbf{R}^p$ and $\mu \in \mathbf{R}^q$ are Lagrange multiplier vectors. After differentiation and some algebra, the solution equations take the form, for $k \in \mathcal{S}$ with $r_k = 1$,

$$\hat{w}_k \;=\; w_k^o \,\cdot\, h(1 + \,(1 + \alpha^{-1})\, \mu' \mathbf{z}_k \,+\, \lambda' \mathbf{x}_k) \tag{7}$$

where h is defined as the inverse function on the whole real line of the function $x + (\alpha^{-1} + 1)\, Q'(x)$, so that $h(1) = 1$ and $h(u) \,+\, (1 + \alpha^{-1})\, Q'(h(u)) \,\equiv\, u$, and the Lagrange multipliers are determined in terms of the matrix

$$M_\alpha \;=\; \sum_{k \in \mathcal{S}} \; w_k^o \begin{pmatrix} \mathbf{x}_k \mathbf{x}_k' & \mathbf{x}_k\, \mathbf{z}_k' \\ \mathbf{z}_k \mathbf{x}_k' & (1 + \alpha^{-1})\, \mathbf{z}_k \mathbf{z}_k' \end{pmatrix}$$

by the expression (the left-hand side of which is a nonsingular function of (λ, μ), with Jacobian bounded above and below by positive-definite matrices)

$$N^{-1}\, M_\alpha \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \;-\; N^{-1} \sum_{k \in \mathcal{S}} r_k\, w_k^o\, Q' \circ h(1 + \,\frac{1 + \alpha}{\alpha}\, \mu' \mathbf{z}_k \,+\, \lambda' \mathbf{x}_k) \begin{pmatrix} \mathbf{x}_k \\ (1 + \alpha^{-1}) \mathbf{z}_k \end{pmatrix}$$

$$= \; N^{-1} \left\{ \begin{pmatrix} t^*_{\mathbf{x}} \\ t^*_z \end{pmatrix} - \sum_{k \in \mathcal{S}} r_k \, w^o_k \begin{pmatrix} \mathbf{x}_k \\ \mathbf{z}_k \end{pmatrix} \right\} \tag{8}$$

When $Q \equiv 0$, these equations immediately give the Lagrange multipliers $\hat{\lambda}$, $\hat{\mu}$ and final weights $\hat{w}_k$ in closed form. Otherwise, for $Q \not\equiv 0$, we use these multipliers and weights as initial values and alternate between solving (7) for new weights and solving (8) — with the new weights substituted — for updated Lagrange multipliers, until convergence. This algorithm can be shown to converge to the unique solutions, and the method works well in the numerical examples we have tried.

4. Asymptotic Theory for Solutions

Following the approach of and regularity conditions similar to Deville and Särndal (1992), we sketch the theory of large-sample behavior of the solutions of the weight-equations just developed. Additional superpopulation regularity conditions needed here are that the response indicators r_i are independent of each other and the sample $\mathcal{S}$, and that the population-total constraints $t^*_{\mathbf{z}}$, $t^*_{\mathbf{x}}$ satisfy as $N \to \infty$:

$$\lim_N \frac{n^{1/2}}{N} \begin{pmatrix} t^*_{\mathbf{x}} - t_{\mathbf{x}} \\ t^*_{\mathbf{z}} - t_{\mathbf{z}} \end{pmatrix} = \begin{pmatrix} k_{\mathbf{x}} \\ k_{\mathbf{z}} \end{pmatrix} \quad \text{exists} \;, \quad k_{\mathbf{z}} \in \mathbf{R}^q \;, \quad k_{\mathbf{x}} \in \mathbf{R}^p \tag{9}$$

Since the functions G_1, G_2, and Q are all assumed convex, so is the objective function in (6), and there will be a unique set of minimizing weights $\mathbf{w}$, $\hat{\mathbf{w}}$ equating the gradient of the objective function to $\mathbf{0}$. For this reason, the limit along any convergent subsequence for the iterative scheme solving (7) and (8) yields the unique solution to these equations.

A proof given in Slud and Thibaudeau (2009), along the same lines as in Deville and Särndal (1992), establishes the large-sample design consistency of the calibrated weights. The solutions $(\hat{\lambda}, \hat{\mu})$ have finite in-probability limits, as do the calibrated weights $\hat{w}_k$, the latter being given by (7) with $(\hat{\lambda}, \hat{\mu})$ substituted for (λ, μ).

The parameters (λ_*, μ_*) consistently estimated by $(\hat{\lambda}, \hat{\mu})$ in large samples are

$$\begin{pmatrix} \lambda_* \\ \mu_* \end{pmatrix} = \phi^{-1} \Big(\lim_N N^{-1} \sum_{k \in \mathcal{U}} (1 - \rho_k) \begin{pmatrix} \mathbf{x}_k \\ \mathbf{z}_k \end{pmatrix} \Big) \tag{10}$$

where the invertible mapping ϕ on $\mathbf{R}^{p+q}$ is the limit of the mappings defined by the left-hand side of (8). Next define (for all $k \in \mathcal{U}$),

$$D_* \equiv \nabla \phi(\lambda_*, \mu_*) \;, \quad u^*_k \equiv 1 + \frac{1+\alpha}{\alpha} \, \mu'_* \, \mathbf{z}_k + \lambda'_* \, \mathbf{x}_k \;, \quad f^*_k \equiv h(u^*_k) \tag{11}$$

These derivations lead to estimators for the variances of survey-weighted totals (4). By linearization of (7) as a function of (λ, μ) around (λ_*, μ_*), there follows:

Proposition 1 *Under regularity conditions like those of Deville and Särndal (1992) along with (9),*

$$\frac{\sqrt{n}}{N} \Big[\sum_{k \in \mathcal{S}} \hat{w}_k \, y_k - \sum_{k \in \mathcal{U}} \rho_k \, f^*_k \, y_k \Big] = \begin{pmatrix} k_{\mathbf{x}} \\ k_{\mathbf{z}} \end{pmatrix}' D_*^{-1} \, b_* + \frac{\sqrt{n}}{N} \sum_{k \in \mathcal{U}} (r_k \, w^0_k \, I_{[k \in \mathcal{S}]} - \rho_k) \cdot$$

$$\cdot \Big[f^*_k \, y_k - b'_* D_*^{-1} \Big(\frac{1 + \lambda'_* \mathbf{x}_k}{1+\alpha} \begin{pmatrix} \mathbf{x}_k \\ \mathbf{0} \end{pmatrix} + \frac{\alpha \, f^*_k}{1+\alpha} \begin{pmatrix} \mathbf{x}_k \\ (1+\alpha^{-1}) \mathbf{z}_k \end{pmatrix} \Big) \Big] + o_P(1)$$

where b_* *is defined by*

$$b_* \;=\; \lim_N \, N^{-1} \sum_{k \in \mathcal{U}} \rho_k \, y_k \begin{pmatrix} \mathbf{x}_k \\ (1+\alpha^{-1}) \, \mathbf{z}_k \end{pmatrix}$$

4.1 Consistency of Weighted Survey Totals

The consistency of survey-weighted totals in the nonresponse and calibration framework must be justified by one or both of two essentially model-based assumptions, saying either that the weight terms $\rho_k \, f_k^*$ in the centering constant of Proposition 1 are almost uniformly close to 1 or that their differences from 1 are asymptotically orthogonal to the vector $\{y_k\}_{k \in \mathcal{U}}$ of survey attributes in the frame population. This twofold path to consistency is a known instance of the concept of *double robustness* (Kang and Schafer 2007). The underlying assumptions do require that response be noninformative for y_k, as specified in the following Proposition. The proof is given in Slud and Thibaudeau (2009).

Proposition 2 *Assume the regularity conditions of Prop. 1 plus (9), and in addition one of the following:*

(i). There is a subset $\mathcal{U}_1 \subset \mathcal{U}$ *and* $\lambda \in \mathbf{R}^p$ *such that for all* $k \in \mathcal{U}_1$, $\rho_k \equiv (1 + \lambda' \mathbf{x}_k)^{-1}$, *and the open interval* (c_1, c_2) *on which* $Q \equiv 0$ *contains the closed interval* $[\min_{k \in \mathcal{U}_1} (1 + \lambda' \mathbf{x}_k), \, \max_{k \in \mathcal{U}_1} (1 + \lambda' \mathbf{x}_k)]$, *and also*

$$\sum_{k \in \mathcal{U} \setminus \mathcal{U}_1} (|y_k| \;+\; \|\mathbf{z}_k\|^2 \;+\; \|\mathbf{x}_k\|^2) \;=\; o(N/\sqrt{n})$$

(ii). The function Q *is identically* 0, *and for some* $\beta \in \mathbf{R}^q$,

$$\lim_{N \to \infty} (\sqrt{n}/N) \sum_{k \in \mathcal{U}} \rho_k \, \mathbf{z}_k \, (y_k \;-\; \beta' \mathbf{z}_k) \;=\; 0$$

Then the left-hand side of the equality in Prop. 1 is equal to

$$(\sqrt{n}/N) \, (\sum_{k \in \mathcal{S}} \hat{w}_k \, y_k \;-\, t_y) \;+\; o_P(1)$$

Assumption (i) says that the 'working' quasi-randomization model is correct, while (ii) is a slight generalization (as long as n is much smaller than N) of the requirement that for some $\beta \in \mathbf{R}^{p+q}$, the vector of residuals $y_k - \beta' \mathbf{x}_k^*$ is asymptotically orthogonal to the columns of the predictor matrix $X_* = (x_{k,j}, \, z_{k,j})$ with row-index k. The second of these assumptions is the technical sense in which response should be noninformative for y_k. Other authors, such as Fuller (2002), prove consistency in superpopulation models by assuming that the residuals are independent mean-0 errors uncorrelated with the calibration variables $\mathbf{x}_k^*$.

4.2 Consequences of Propositions 1 and 2

It seems unavoidable that the consistency of the survey-weighted total estimators depends strongly on unverifiable model assumptions about the 'stochastic' mechanisms of unit nonresponse and omission from survey frames. Even the relatively slight (order of $1/\sqrt{n}$) discrepancies assumed between the calibration totals $t_{\mathbf{x}}^*$, $t_{\mathbf{z}}^*$ and the respective totals $t_{\mathbf{x}}, t_{\mathbf{z}}$ they approximate, imply through the first term on the right-hand side of the equality in Proposition 1 that the confidence intervals based on survey estimates (4) can be wrongly centered.

However, the clear positive consequence of the new single-stage approach to weight adjustment is a readily computable estimator for the linearized variance formula for $\hat{t}_{y,adj}$. In terms of the natural design-consistent estimators $\hat{f}_k$, $\hat{b}$, and $\hat{D}$ for the respective quantities f_k^*, b_*, D_*, and of the (ratio-adjustment or calibration based) estimator $\hat{\rho}_k$ for ρ_k, when the totals $t_{\mathbf{x}}^*$ and $t_{\mathbf{z}}^*$ are treated as fixed, this variance estimator $\hat{V}(\sqrt{n}\,\hat{t}_{y,adj}/N)$ is

$$\frac{n}{N^2}\Big[\sum_{k,l\in\mathcal{S}}(w_k^o\,w_l^o - \frac{1}{\pi_{kl}})\,\hat{R}_k\,\hat{R}_l\,r_k\,r_l + \sum_{k\in\mathcal{S}} r_k\,w_k^o\,\hat{R}_k^2\,(1-\hat{\rho}_k)\Big] \tag{12}$$

where

$$\hat{R}_k \;\equiv\; \hat{f}_k\,y_k - \hat{b}'\,\hat{D}^{-1}\Big(\frac{1+\hat{\lambda}'\mathbf{x}_k}{1+\alpha}\begin{pmatrix}\mathbf{x}_k\\ \mathbf{0}\end{pmatrix} + \frac{\hat{f}_k\,\alpha}{1+\alpha}\begin{pmatrix}\mathbf{x}_k\\ (1+\alpha^{-1})\mathbf{z}_k\end{pmatrix}\Big)$$

5. Implementation on SIPP 1996 Data

As an example, single-stage calibrated weights were fitted to the 1996 Wave 1 data from the SIPP survey, a large stratified complex longitudinal survey conducted by the U.S. Census Bureau. As described in Slud and Bailey (2009), SIPP nonresponse adjustment was based on poststratified ratio-adjustment using a system of $p = 149$ cells (involving demographics and some indicators of assets and of income compared with the poverty level), and SIPP population controls involved raking to CPS-based totals of race by family structure, race by age interval, and of Hispanic-origin persons by (coarser) age groups. The SIPP population-control constraints can be expressed in $q = 126$ linearly independent equations of the type (3). The weights w_k^o used in the SIPP file of 94444 individuals sampled and responding in 1996 Wave 1 are the base weights `GBASEWT` before nonresponse adjustment. The $t_{\mathbf{x}}^*$ constrained totals in (2) are the values $N\sum_{\mathcal{S}}\mathbf{x}_k\,w_k^o/\sum_{\mathcal{S}} w_k^o$ estimated with the base weights, and the population N and control totals $t_{\mathbf{z}}^*$ were, as in SIPP production estimates, derived from the 1990 census demographically updated to 1996.

In this example, the single-stage calibrated weights were fitted with $\alpha = 1$ and 100, using penalty functions Q defined on the interval $(0, 10)$ by

$$Q(u) \;=\; a\,\{(c_1-u)_+^3/u^4 \;+\; (u-c_2)_+^3/(10-u)^2\}$$

where $(x)_+ \equiv \max(x, 0)$, and $c_1 = .6, c_2 = 3, a = .2$ when $\alpha = 1$, while $c_1 = .8, c_2 = 2.5, a = .5$ when $\alpha = 100$. The fitted weight ratios $\hat{w}_k/w_k^o$ fell in the range $(.55, 3.86)$ for $\alpha = 1$, and in $(.62, 4.08)$ for $\alpha = 100$. The 126 fitted Lagrange multiplier components of $\hat{\mu}$ had range $(-1.13, 0.72)$ for $\alpha = 1$, and $(-0.86, 1.77)$ for $\alpha = 100$. Both sets of single-stage modified weights satisfied (3) accurately. The modified weights with $\alpha = 1$ already satisfied (2) to within several percent, while those with $\alpha = 100$ satisfied (2) to within a few tenths of a percent. As an indication of the similarity of estimated results $\hat{t}_{y,adj}$ using the estimated weights from (7), with $\alpha = 1$ or 100, to those obtained in SIPP 1996 by what amounted to two-stage weight adjustment (with raking rather than linear calibration in the second population-control stage), Table 1 gives the estimated totals (in thousands) for 11 of the important SIPP surveyed attributes. The labelling for the survey attributes is as in Slud and Bailey (2009), where further details can be found. Differences among the estimators based on the three sets of weights are quite small compared with the standard deviations of the two-stage estimated totals, given in the Table as the EHG (Ernst, Huggins and Grill 1986) estimated standard deviation, a slightly upwardly biased estimator of standard deviation which the VPLX Fay-method standard deviation estimator approximates.

Table 1: Estimated Totals and Std. Dev.'s in 1000's. '2-Stage' method and EHG standard deviation described in text; totals and standard deviations for $\hat{t}_{y,adj}$ based on (7) with $\alpha = 1$, 100, with variables $\mathbf{x}_k$, $\mathbf{z}_k$ and function Q given in text.

Item	2stage	$\hat{t}_{\alpha=1}$	$\hat{t}_{\alpha=100}$	EHG.sd	$SD_{\alpha=1}$	$SD_{\alpha=100}$
FOODST	27541	27454	26930	687.0	317.7	300.8
AFDC	14123	14089	13800	450.5	298.3	287.6
MDCD	28468	28399	27895	573.8	403.7	351.1
SOCSEC	36994	37071	37240	469.6	156.9	156.8
HEINS	194216	194475	195030	1625.1	438.8	422.6
POV	41951	41978	41475	747.3	360.1	357.1
EMP	190871	190733	190097	1477.3	254.8	239.8
UNEMP	6403	6379	6295	163.1	144.5	143.2
NILF	66979	67354	67864	626.7	231.4	216.7
MAR	111440	114457	114347	1088.1	159.1	157.7
DIV	18534	18542	18591	253.4	194.9	194.9

The VPLX or EHG variances calculated for Table 1 account for nonresponse adjustment only by scaling up the `GBASEWT` base weights so that their total is the known population size N. The Table also displays variances as calculated from formula (12), using $\alpha = 1$ or 100 and Q as specified above. These estimated variances are much smaller than the VPLX variances, but only because the variance formula (12) calculated here treats all calibration totals $t^*_{\mathbf{x}}$, $t^*_{\mathbf{z}}$ as though they were known from the outset and nonrandom. In fact, the values $t^*_{\mathbf{x}}$ were derived from sample estimates from the same SIPP 1996 survey.

6. Discussion and Future Research

This paper has developed asymptotic theory and computational tools for a new, single-stage approach to the adjustment of weights in large complex surveys for nonresponse, population-controls, and weight-trimming. Numerical experiments using functions created in the `R` statistical programming language show that the single-stage modified weights are readily computable iteratively, and are generally similar to the two-stage weights obtained first by nonresponse-adjustment and then by population controls and trimming. The single-stage calibrated weights with large α (of order 100) penalized via function Q to lie in intervals such as $(.4, 3)$, seem particularly satisfactory.

Future research will compare variance properties of estimators based on the new weights under various sources for the calibration totals $t^*_{\mathbf{x}}$, $t^*_{\mathbf{z}}$. In addition, the theory and algorithms will be extended to the case of nonlinear $G'_j(z) = z \log z - z + 1$, the so-called 'multiplicative' form of the adjustment-discrepancy loss-function which leads in the absence of weight-penalization to classical raking.

REFERENCES

Chang, T. and Kott, P. (2008), Using calibration weighting to adjust for nonresponse under a plausible model. *Biometrika* **95**, 555-571.

Deville, J. and Särndal, C. (1992), Calibration estimators in survey sampling, *Jour. Amer. Statist. Assoc.* **87**, 376-382.

Deville, J., Särndal, C. and Sautory, O. (1993), Generalized raking procedures in survey sampling. *Jour. Amer. Statist. Assoc.* **88**, 1013-1020.

Ernst, L., Huggins, V., and Grill, D. (1986) Two new variance estimation techniques. *ASA Proc. Survey Res. Meth.*, Washington DC, pp. 400-405.

Fuller, W. (2002), Regression estimation for survey samples. *Survey Methodology* **28**, 5-23.

Kang, J. and Schafer, J. (2007) Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data. *Statist. Sci.* **22**, 523-539.

Kim, Jae-Kwang and Kim, Jay J. (2007) Nonresponse weighting adjustment using estimated response probability. *Canad. Jour. Statist.* **35**, 501-514.

Oh, H. and Scheuren, F. (1983) Weighting adjustment for unit nonresponse. In: *Incomplete Data in Sample Surveys*, vol. 2, Eds. Madow, W., Olkin, I. and Rubin, D. New York: Academic Press, 143-184.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, ISBN 3-900051-07-0, `http://www.R-project.org`.

Särndal, C. and Lundström, S. (2005) *Estimation in Surveys with Nonresponse*. Wiley: Chichester.

Särndal, C.-E., Swensson, B. and Wretman, J. (1992) *Model Assisted Survey Sampling*. Springer: New York.

Singh, A. and Mohl, C. (1996), Understanding calibration estimation in survey sampling. *Survey Methoology* **22**, 107-115.

Slud, E. and Bailey, L. (2009), Evaluation and selection of models for attrition nonresponse adjustment, *Jour. Official Statist.*, to appear.

Slud, E. and Thibaudeau, Y. (2009), Simultaneous calibration and nonresponse adjustment, SRD Technical Report, Census Bureau.

Sverchkov, M. and Pfeffermann, D. (2004). Prediction of finite population totals based on the sample distribution. *Survey Methodology* **30**, 79-92.

Théberge, A. (2000) Calibration and restricted weights. *Survey Methodology* **26**, 99-107.

Yung and Rao, J.N.K. (2000) Jackknife variance estimation under imputation for estimators using poststratification information. *Jour. Amer. Statist. Assoc.* **95**, 903-915.