Journal of Official Statistics
Vol. 10, No. 2, 1994, pp. 153–165
© Statistics Sweden

Estimating Totals and Distribution Functions Using Auxiliary Information at the Estimation Stage

J.N.K. Rao[1]

Abstract: A general set-up for inference from survey data that covers the estimation of totals and distribution functions is provided, using auxiliary information at the estimation stage. Both probability sampling and model-assisted approaches are studied. A conditional probability sampling approach that provides conditionally valid repeated sampling inferences, under model misspecifications, is also considered. Finally, asymptotically efficient calibration estimators that satisfy certain consistency constraints are proposed.

Key words: Calibration estimators; conditional probability sampling approach; model-assisted approach.

1. Introduction

In sample surveys, supplementary population information is often used at the estimation stage to increase the precision of estimators of a population total. In particular, customary ratio and regression estimators make use of known population totals of auxiliary variables. Recently, several estimators of a population distribution function have also been proposed, using auxiliary information at the estimation stage. The main purpose of this article is to provide a general set-up that covers the estimation of totals and distribution functions, utilizing auxiliary information at the estimation stage.

It is often desirable to revise the basic survey weights to satisfy certain consistency constraints. In particular, the sample sum of a weighted auxiliary variable should equal the known population total for that auxiliary variable. Deville and Särndal (1992) named such revised weights as calibration weights and the resulting estimators of a total as calibration estimators. They proposed a general method of deriving calibration estimators by choosing a distance measure between the calibration weights and the basic weights and then minimizing this distance subject to specified consistency constraints, called the calibration equations. They have also shown that a "chi-square distance" leads to the generalized regression estimator (Särndal 1980; Bethlehem and Keller

[1] Department of Mathematics and Statistics, Carleton University, Ottawa, K1S 5B6, Canada.

Acknowledgements: This paper was presented at the Workshop on Uses of Auxiliary Information in Surveys, Örebro, Sweden, October 4–6, 1992. I am thankful to Phil Kott for some useful discussion on the model-assisted variance estimators and to a referee for constructive comments. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

1987). In this article, we provide alternative calibration estimators that are asymptotically efficient.

2. General Set-Up

The following theoretical framework is often assumed in estimating population parameters. A survey population U consists of N distinct elements identified through the labels $j = 1, \ldots, N$. The characteristic of interest y_j associated with element j is exactly known by observing the element j. A sample is a subset, s, of U and the associated y-values, i.e., $\{(i, y_i),\ i \in s\}$, selected according to a specified sampling design which assigns a known probability $p(s)$ to s such that $p(s) > 0$ for all $s \in S$, the set of possible samples s, and $\sum_{s \in S} p(s) = 1$.

We consider general parameters of interest

$$H = \sum_{j \in U} h(y_j) \text{ and } \bar{H} = N^{-1} H \qquad (2.1)$$

for a specified function h. The choice $h(y) = y$ gives the population total $H = Y$ and the population mean $\bar{H} = \bar{Y}$, while the choice $h(y) = \Delta(t - y)$ with $\Delta(a) = 1$ when a $\geqslant 0$ and $\Delta(a) = 0$ otherwise gives the distribution function

$$\bar{H} = F(t) = N^{-1} \sum_{j \in U} \Delta(t - y_j) \qquad (2.2)$$

for each t.

The problem is to estimate H or $\bar{H}$ by observing a sample selected according to the specified design and also using available auxiliary data. We assume that supplementary information $\mathbf{x}_j = (x_{j1}, \ldots x_{jp})'$ associated with population elements j is available at the estimation stage. The case where only the population total $\mathbf{X}$ is available is also considered.

There are essentially three different approaches to inference on H or $\bar{H}$: (i) design-based approach, also called probability sampling approach; (ii) model-dependent approach, also called prediction approach; (iii) hybrid approach, called model-assisted approach. An advantage of the model-assisted approach is that it provides valid inferences under an assumed model and at the same time protects against model misspecifications in the sense of providing valid repeated sampling inferences. In this paper, we will focus on (i) and (iii), but also consider a conditional probability sampling approach that provides conditionally valid repeated sampling inferences, under model misspecifications, given suitable ancillary statistics such as design-unbiased estimators, $\hat{\mathbf{X}}$, of the known totals $\mathbf{X}$.

Probability sampling approach refers to repeated sampling from the survey population U involving all samples $s \in S$ and associated probabilities $p(s)$. It provides valid inferences irrespective of the population y-values in the sense that the pivotals $t_1 = (\hat{H} - H)/s(\hat{H})$ and $t_2 = (\hat{\bar{H}} - \bar{H})/\ s(\hat{\bar{H}})$ are approximately N(0, 1), at least for large samples, where $(\hat{H}, \hat{\bar{H}})$ and $(s^2(\hat{H}),\ s^2(\hat{\bar{H}}))$ are design-consistent estimators of $(H, \bar{H})$ and $(\mathrm{Var}(\hat{H}),\ \mathrm{Var}(\hat{\bar{H}}))$ respectively.

We assume that the inclusion probabilities $\pi_i = \sum_{\{s: i \in s\}} p(s),\ i = 1, \ldots, N$ are positive, which permits unbiased and consistent estimation of H and $\bar{H}$. We also assume that the joint inclusion probabilities $\pi_{ij} = \sum_{\{s: i, j \in s\}} p(s),\ i < j = 1, \ldots, N$, are positive, which permits unbiased and consistent estimation of the variance of $\hat{H}$ and $\hat{\bar{H}}$.

A general class of estimators of H is given by

$$\hat{H} = \sum_{i \in s} d_i(s) h(y_i) \qquad (2.3)$$

where the basic weights $d_i(s)$ can depend both on s and $i (i \in s)$ and satisfy the design-unbiasedness condition: $\sum_{\{s: i \in s\}} p(s) d_i(s) = 1$ for $i = 1, \ldots, N$. The choice $h(y) = y$ in

(2.3) gives Godambe's (1955) class of estimators, $\hat{Y}$, of a total Y. The well-known Horvitz-Thompson (H-T) estimator is a special case of (2.3) with $d_i(s) = \pi_i^{-1}$. The well-known Murthy's estimator and Rao-Hartley-Cochran's estimator (Cochran, 1977, ch. 9A) also belong to the general class (2.3).

If the variance of $\hat{Y}$, $V(\hat{Y})$, becomes zero when $y_i \propto w_i$ for some known non-zero constants w_i, then a nonnegative unbiased quadratic estimator of $V(\hat{H})$ is necessarily of the form (Rao 1979)

$$v(\hat{H}) = -\sum_{\substack{i<j \\ i,j \in s}}\sum d_{ij}(s) w_i w_j (z_i - z_j)^2 \tag{2.4}$$

where $z_i = h(y_i)/w_i$ and the weights $d_{ij}(s)$ can depend both on s and $(i,j) \in s$, and satisfy the unbiasedness condition. The well-known Sen-Yates-Grundy (S-Y-G) estimator of variance of H-T estimator is a special case of (2.4) with $w_i = \pi_i$ and $d_{ij}(s) = (\pi_{ij} - \pi_i \pi_j)/(\pi_{ij}\pi_i\pi_j)$, for any fixed sample size, n, design. It is interesting to note that the original H-T estimator of variance does not belong to class (2.4), although it is valid both for fixed and non-fixed sample size designs. For the general estimator (2.3), a H-T type unbiased variance estimator is given by

$$\begin{aligned} v^*(\hat{H}) = {} & \sum_{i \in s} d_i(s)(d_i(s) - 1) w_i^2 z_i^2 \\ & + 2 \sum_{\substack{i<j \\ i,j \in s}}\sum \frac{(\alpha_{ij} - 1)}{\alpha_{ij}} \\ & \times d_i(s) d_j(s) w_i w_j z_i z_j \end{aligned} \tag{2.5}$$

where $\alpha_{ij} = \sum_{\{s:i,j \in s\}} p(s) d_i(s) d_j(s)$. If $d_i(s) = \pi_i^{-1}$, then $v^*(\hat{H})$ reduces to the H-T variance estimator. The H-T variance estimator is seldom used in practice since it can take negative values often and can lead to a large coefficient of variation.

Turning to $\bar{H}$, a general class of estimators of $\bar{H}$ is given by

$$\hat{\bar{H}} = \frac{\hat{H}}{\hat{N}} = \frac{\sum_{i \in s} d_i(s) h(y_i)}{\sum_{i \in s} d_i(s)}. \tag{2.6}$$

Note that if $h(y_i) = \Delta(t - y_i)$ in (2.6), then $\hat{\bar{H}}$ retains the properties of a distribution function, provided all the basic weights, $d_i(s)$, are nonnegative. A consistent estimator of variance of $\hat{\bar{H}}$ is obtained from (2.4) by replacing $h(y_i)$ by $(h(y_i) - \hat{\bar{H}})/\hat{N}$.

The estimators (2.3) and (2.6) do not utilize the auxiliary information $\mathbf{x}_j$ $(j = 1, \ldots, N)$ at the estimation stage. A ratio estimator of H, in the case of a single x-variable, can be obtained as

$$\hat{H}_r = (\hat{H}/\hat{G})G \tag{2.7}$$

where

$$\hat{G} = \sum_{i \in s} d_i(s) g(x_i), \quad G = \sum_{j \in U} g(x_j)$$

and $g(x_j)$ is positively related to $h(y_j)$ such that $\hat{H}_r$ reduces to the known total G when $y_j \propto x_j$ for all $j \in U$, and hence the variance becomes zero in the latter case. Note that $\hat{H}_r$ is a calibration estimator with respect to the auxiliary variable $g(x)$. If the population size, N, is known, a ratio estimator of $\bar{H}$ is given by

$$\hat{\bar{H}}_r = \hat{H}_r/N. \tag{2.8}$$

In the case of the total Y, we choose $g(x_j) = x_j$ and $\hat{H}_r$ reduces to

$$\hat{Y}_r = \frac{\sum_{i \in s} d_i(s) y_i}{\sum_{i \in s} d_i(s) x_i} X = \frac{\hat{Y}}{\hat{X}} X = \hat{R} X \tag{2.9}$$

where $X = \sum_{j \in U} x_j$ is the known total of the x_j's. The ratio estimator $\hat{Y}_r$ leads to significant reduction in the variance relative to the unbiased estimator $\hat{Y}$, when y_j is positively related to x_j. In the case of the distribution function $F(t)$, we can choose $g(x_j) = \Delta(t - \hat{R}x_j)$, provided the population total of the $g(x_j)$'s is known. The resulting estimator $\hat{\bar{H}}_r = \hat{\bar{F}}_r(t)$ ensures the above desirable property of zero variance when $y_j \propto x_j$. However, in general the correlation between $\Delta(t - y_j)$ and $\Delta(t - Rx_j)$ is likely to be weaker than the correlation between y_j and x_j, where $R = Y/X$. As a result, the gains in efficiency of $\hat{\bar{F}}_r(t)$ over the estimator $\hat{\bar{H}} = \hat{\bar{F}}(t)$ are likely to be smaller than those achieved by the ratio estimator $\hat{Y}_r$ over $\hat{Y}$.

A regression estimator of H can also be obtained as

$$\hat{H}_{\text{reg}} = \hat{H} + \hat{B}(G - \hat{G}) \tag{2.10}$$

where

$$\hat{B} = \text{cov}(\hat{H}, \hat{G})/v(\hat{G}) \tag{2.11}$$

and $\text{cov}(\hat{H}, \hat{G})$ and $v(\hat{G})$ are obtained from (2.4) by replacing $(z_i - z_j)^2$ with $(z_i - z_j)$ $(u_i - u_j)$ and $(u_i - u_j)^2$ respectively, where $u_i = g(x_i)/w_i$. Similarly, a regression estimator of $\bar{H}$ is given by

$$\hat{\bar{H}}_{\text{reg}} = \hat{H}_{\text{reg}}/N \tag{2.12}$$

provided N is known. In the case of $F(t)$ with $g(x_j) = \Delta(t - \hat{R}x_j)$, the regression estimator $\hat{\bar{H}}_{\text{reg}} = \hat{\bar{F}}_{\text{reg}}(t)$ retains the above desirable property of zero variance when $y_j \propto x_j$, but it also suffers from the same drawback as the ratio estimator $\hat{\bar{F}}_r(t)$. The regression estimator $\hat{\bar{H}}_{\text{reg}}$ is computationally more cumbersome than the ratio estimator $\hat{\bar{H}}_r$ since it involves the evaluation of $\text{cov}(\hat{H}, \hat{G})$ and $v(\hat{G})$. However, the latter evaluation can be simplified for some commonly used designs (see Section 5). The regression estimator $\hat{\bar{H}}_{\text{reg}}$ can be readily extended to multiple auxiliary variables.

3. Model-Assisted Approach

Probability sampling approach has been criticized on the grounds that the associated inferences, although assumption-free, refer to repeated sampling instead of just the particular sample, s, that has been drawn. Prediction approach, on the other hand, assumes that the population y-values are random and obey a model, and the model distribution leads to valid inferences referring to the particular s that has been drawn, irrespective of the sample design $p(s)$. Prediction inferences, in large samples, however, are very sensitive to model misspecifications, as illustrated by Hansen, Madow, and Tepping (1983). By considering only design-consistent estimators and variance estimators that are also model-unbiased (at least asymptotically) under an assumed model, the model-assisted approach attempts to provide valid conditional inferences under the assumed model and at the same time protects against model misspecifications in the sense of providing valid design-based inferences irrespective of the population y-values.

Although model-assisted estimators of a total, Y, can be obtained under general linear (or nonlinear) regression models, we will confine ourselves here, for simplicity, to a single x-variable and the following often-used simple linear regression model

$$E_m(y_j) = \beta x_j, \quad j = 1, \ldots, N \tag{3.1}$$

where E_m denotes the model expectation and β is an unknown parameter. It is further assumed that the y_j's are independent with model variance $V_m(y_j) = \sigma^2 x_j$ and $\sigma^2 (> 0)$ is an unknown parameter and V_m denotes the model variance. We assume that the population model (3.1) also holds for the sample, i.e., there is no sample selection bias (see Krieger and Pfeffermann (1992) for an illuminating discussion of the effects of sample selection). An estimator of Y, say $\tilde{Y}$, is model-unbiased for Y if $E_m(\tilde{Y} - Y) = 0$ for every $s \in S$. Under model (3.1), the best linear unbiased estimator of Y, in the sense of minimizing the model variance $V_m(\hat{Y} - Y)$, is the simple ratio estimator $(\bar{y}/\bar{x})X$ for any $p(s)$, where $\bar{y}$ and $\bar{x}$ are the sample means (Brewer 1963; Royall 1970). Since this estimator does not depend on the survey weights $d_i(s)$, it is generally design-inconsistent.

A model-assisted estimator of Y, under model (3.1), is given by

$$\hat{Y}_{ma} = \sum_{i\in s} d_i(s)y_i + \hat{R}\left(X - \sum_{i\in s} d_i(s)x_i\right)$$
$$= \hat{R}X \tag{3.2}$$

which is the same as the ratio estimator (2.9). The ratio estimator (3.2) can be motivated along the lines of Särndal (1980), noting that $\hat{R}x_j$ is a predictor of y_j under model (3.1) and that the total of prediction errors $e_j = y_j - \hat{R}x_j$ is estimated by $\Sigma_{i\in s}d_i(s)e_i$. It can also be written as

$$\hat{Y}_{ma} = \sum_{i\in s} d_i^*(s)y_i \tag{3.3}$$

where the revised weight $d_i^*(s)$ is the product of the basic weight $d_i(s)$ and the so-called g-weight, $g_i(s) = X/\hat{X}$, which converges in probability to 1.

A consistent estimator of variance of $\hat{Y}_{ma}$ is either given by the S-Y-G type variance estimator (2.4) or by the H-T type variance estimator (2.5) with z_i replaced by $(y_i - \hat{R}x_i)/w_i = e_i/w_i$, where $\hat{R}$ is a model-unbiased estimator of β. However, it is in general not model-unbiased (even approximately) for the model variance $V_m(\hat{Y}_{ma} - Y)$. In the case of the H-T estimator of Y with $d_i(s) = \pi_i^{-1}$, Särndal, Swensson, and Wretman (1989) proposed a model-assisted variance estimator that is both approximately model unbiased (when n/N is of the order $O(n^{-1/2})$ or less) and design-consistent. This is simply obtained by changing y_i to $g_i(s)e_i = (X/\hat{X})(y_i - \hat{R}x_i)$ in the H-T variance estimator (also see Hidiroglou, Fuller, and Hickman 1976).

We now extend the Särndal et al. (1989) result to the ratio estimator (3.3) with general weights $d_i(s)$. We show that the H-T type variance estimator (2.5) with $w_iz_i = y_i$ replaced by $g_i(s)e_i$ is both model-unbiased (approximately) and design-consistent. The latter property follows from the fact that $g_i(s)$ converges in probability to 1. (We assume that the design is such that $v^*(\hat{Y})$ is design-consistent.) Under model (3.1), it is straightforward to show that

$$V_m(\hat{Y}_{ma} - Y)$$
$$= \sigma^2\left[\sum_{i\in s}\{d_i(s)g_i(s)\}^2x_i - \sum_{i\in U}x_i\right]. \tag{3.4}$$

Also, the proposed variance estimator, $\tilde{v}(\hat{Y}_{ma})$ say, is approximately equal to (2.5) with w_iz_i changed to $g_i(s)\epsilon_i$ where $\epsilon_i = y_i - \beta x_i$ are independent errors with mean zero and variance $\sigma_i^2 = \sigma^2x_i$. Hence, its model expectation is given by

$$E_m\tilde{v}(\hat{Y}_{ma}) \doteq \sigma^2\left[\sum_{i\in s}\{d_i(s)g_i(s)\}^2x_i - \sum_{i\in s}d_i(s)g_i^2(s)x_i\right]. \tag{3.5}$$

Comparing (3.4) and (3.5), we note that the leading terms are identical. Assuming that n/N is of the order $O(n^{-1/2})$ or less, the lower order terms are also approximately equal by noting that

$$\sum_{i\in s}d_i(s)g_i^2(s)x_i = \frac{X^2}{\hat{X}} = \sum_{i\in U}g_i(s)x_i \tag{3.6}$$

and that $g_i(s)$ converges to 1 in probability.

Unfortunately, the above simple recipe of getting a variance estimator that is both design-consistent and approximately model-unbiased does not seem to work when applied to the more useful S-Y-G type variance estimator (2.4). This is also true in the special case of H-T weights $d_i(s) = \pi_i^{-1}$, unless certain restrictions are placed on the joint probabilities, π_{ij}. (Note that (2.4) reduces to the S-Y-G variance estimator in this case.) Nevertheless, we recommend the variance estimator (2.4) with $w_i z_i = y_i$ replaced by $g_i(s)e_i$ since it remains design-consistent and is expected to be more stable than the corresponding H-T type variance estimator, $\tilde{v}(\hat{Y}_{ma})$. Moreover, its model bias is likely to be smaller than that of the customary variance estimator (2.4) with y_i replaced by e_i, although it is not approximately model-unbiased.

Kott (1990) proposed an alternative variance estimator for the H-T estimator of Y which is also design-consistent and model-unbiased. Generalizing his approach, we get the following variance estimator

$$v_*(\hat{Y}_{ma}) = \frac{v(\hat{Y}_{ma})}{E_m v(\hat{Y}_{ma})} V_m(\hat{Y}_{ma} - Y) \qquad (3.7)$$

where

$$v(\hat{Y}_{ma}) = -\sum_{\substack{i<j \\ i,j\in s}}\sum d_{ij}(s) w_i w_j \left(\frac{e_i}{w_i} - \frac{e_j}{w_j}\right)^2, \qquad (3.8)$$

$V_m(\hat{Y}_{ma} - Y)$ is given by (3.4) and

$$E_m v(\hat{Y}_{ma}) = -\sigma^2 \sum_{\substack{i<j \\ i,j\in s}}\sum d_{ij}(s) w_i w_j \times \left(\frac{x_i}{w_i^2} + \frac{x_j}{w_j^2}\right). \qquad (3.9)$$

Note that the unknown parameter σ^2 cancels out in (3.7). An advantage of Kott's approach is that it is applicable to the more useful S-Y-G type variance estimator, unlike the Särndal et al. approach, but the resulting variance estimator (3.7) is somewhat more complicated.

Turning to the distribution function $F(t)$, a predictor of $\Delta(t - y_j)$ under model (3.1) is given by

$$\tilde{g}(x_j) = \left(\sum_{i\in s} d_i(s)\right)^{-1} \left\{\sum_{i\in s} d_i(s) \times \Delta\left[x_j^{-1/2}((t - \hat{R}x_j) - e_i)\right]\right\}. \qquad (3.10)$$

A model-assisted estimator of $F(t)$ based on (3.10) is then given by

$$\hat{F}_{ma}(t) = N^{-1}\left\{\sum_{j\in s} d_j(s)\Delta(t - y_j) + \left[\sum_{j\in U} \tilde{g}(x_j) - \sum_{j\in s} d_j(s)\tilde{g}(x_j)\right]\right\}. \qquad (3.11)$$

This estimator is asymptotically model-unbiased for $F(t)$, but its asymptotic design-bias is zero only for a subclass of sampling designs which, however, seems to cover a wide variety of sampling designs (see Godambe (1989) for details). Using estimation function theory, Godambe (1989) arrived at the estimator (3.11) for the special case of $d_i(s) = \pi_i^{-1}$. Rao, Kovar, and Mantel (1990) proposed an alternative model-assisted estimator, for the special case of $d_i(s) = \pi_i^{-1}$, which is asymptotically both model-unbiased and design-unbiased under all designs. Rao and Liu (1992) extended this estimator to the case of general weights $d_i(s)$.

A consistent estimator of variance of $\hat{F}_{ma}(t)$ is obtained from (2.4) by changing z_j to $\{\Delta(t - y_j) - \tilde{g}(x_j)\}/(Nw_j)$. It seems difficult, however, to construct a model-assisted variance estimator that is both asymptotically model-unbiased and design-unbiased. We are currently investigating this problem.

4. Conditional Probability Sampling Approach

As noted in Section 3, the model-assisted approach appeals to unconditional repeated sampling properties of estimators and variance estimators when the model is misspecified. In this section, we develop alternative model-assisted estimators of Y and $F(t)$ with good conditional repeated sampling properties under model misspecification. This is accomplished by conditioning on a suitable ancillary statistic. To simplify the discussion, we confine ourselves to simple random sampling with replacement and omit technical details and extensions which are given in Liu (1992). Under simple random sampling (SRS) and model (3.1), the model-assisted estimator (3.2) reduces to the simple ratio estimator $\hat{Y}_r = (\bar{y}/\bar{x})X$, noting that $d_i(s) = N/n$.

Employing real population data, Royall and Cumberland (1981) studied the conditional bias of estimators and variance estimators, given the sample mean $\bar{x}$ which may be treated as an ancillary statistic when the population mean $\bar{X}$ is known. They drew repeated samples of size n, arranged them in groups with approximately the same value of x, and computed the conditional bias of estimators and variance estimators within these groups. Robinson (1987) used the fact that, given the sample mean $\bar{x}$, the sample mean $\bar{y}$ is asymptotically normal with mean $\bar{Y}+ B(\bar{x}-\bar{X})$, where B is the population regression coefficient and $\bar{Y} = Y/N$ is the population mean, to show that the asymptotic conditional bias of $\hat{\bar{Y}}_r = \hat{Y}_r/N$ is

$$E\left(\hat{\bar{Y}}_r|\bar{x}\right) - \bar{Y} \doteq (R-B)(\bar{X}-\bar{x})\bar{X}/\bar{x}. \tag{4.1}$$

Thus, noting that $\bar{x}-\bar{X} = O_p(n^{-1/2})$ the conditional bias of $\hat{\bar{Y}}_r$ is of the order $O_p(n^{-1/2})$, unlike the unconditional bias of order $O(n^{-1})$. It also follows that the conditional relative bias of $\hat{Y}_r$ or $\hat{\bar{Y}}_r$, i.e., the ratio of the conditional bias to the conditional standard error, is of the order $O_p(1)$ unlike the unconditional relative bias of order $O(n^{-1/2})$ which is asymptotically negligible. Thus the ratio estimator $\hat{Y}_r$ may not lead to conditionally valid inferences in large samples, under model misspecification, although the inferences are asymptotically valid unconditionally. Note that $R \doteq B$ if model (3.1) holds in which case the conditional bias of $\hat{\bar{Y}}_r$ is approximately zero.

Using (4.1), Robinson (1987) obtained a bias-adjusted estimator

$$\hat{Y}_{ra} = \hat{Y}_r + (r-b)(\bar{x}-\bar{X})X/\bar{x} \tag{4.2}$$

with conditional relative bias of order $O_p(n^{-1/2})$, where $r = \bar{y}/\bar{x}$ and b is the sample regression coefficient. This estimator leads to conditionally valid inferences since the conditional relative bias is asymptotically negligible. It remains model-unbiased under model (3.1) since $E_m(r-b) = 0$. An alternative estimator with conditional relative bias of order $O_p(n^{-1/2})$ is given by the customary linear regression estimator

$$\hat{Y}_{\ell r} = N[\bar{y} + b(\bar{X}-\bar{x})] \tag{4.3}$$

noting that $E(\bar{y}|\bar{x}) \doteq \bar{Y} + B(\bar{x}-\bar{X})$, $\bar{x}-\bar{X} = O_p(n^{-1/2})$ and $E(b|\bar{x}) = B + O_p(n^{-1/2})$. Liu (1992) has shown that the conditional variances of $\hat{Y}_{\ell r}$ and $\hat{Y}_{ra}$ are approximately equal. Hence, the two estimators should perform similarly in the conditional framework. Note that $\hat{Y}_{\ell r}$ remains model-unbiased under model (3.1), and it has a smaller unconditional asymptotic variance than $\hat{Y}_r$. However, it has a larger model variance than $\hat{Y}_r$ under model (3.1) since $\hat{Y}_r$ is the best model-unbiased estimator.

Liu (1992) has shown that the customary

variance estimator of $\hat{Y}_{\ell r}$ is conditionally biased for the conditional variance, and derived a bias-adjusted variance estimator which together with $\hat{Y}_{\ell r}$ leads to conditionally valid inferences. By writing $\hat{Y}_{\ell r}$ in the form (3.3) with $d_i(s) = N/n$ and g-weight $g_i(s) = 1 + (x_i - \bar{x})(\bar{X} - \bar{x})/\sum_s (x_i - \bar{x})^2$, the bias-adjusted variance estimator is simply obtained from the customary variance estimator by changing $\tilde{e}_i = y_i - \bar{y} - b(x_i - \bar{x})$ to $g_i(s)\tilde{e}_i$

$$v_a(\hat{Y}_{\ell r}) = \frac{N^2}{n}\left(1 - \frac{n}{N}\right)\sum_{i\in s}\frac{(g_i(s)\tilde{e}_i)^2}{n-1}. \tag{4.4}$$

It is interesting to note that $v_a(\hat{Y}_{\ell r})$ is identical to the Särndal et al. (1989) variance estimator under the linear regression model $y_j = \alpha + \beta x_j + \epsilon_j$, $j = 1, \ldots, N$ with i.i.d. errors ϵ_j. The generalized regression estimator of Särndal et al. (1989) reduces to $\hat{Y}_{\ell r}$ under the latter model.

Turning to the estimation of the distribution function $F(t)$, Rao and Liu (1992) have shown that the model-assisted estimator $\hat{F}_{ma}(t)$, given by (3.11), is conditionally biased, given $\bar{x}$. They also obtained a bias-adjusted estimator given by

$$F^*_{ma}(t) = \hat{F}_{ma}(t) + s_x^{-2}(s_{xh} - s_{x\tilde{g}})(\bar{X} - \bar{x}) \tag{4.5}$$

where s_{xh} and $s_{x\tilde{g}}$ are the sample covariances of x and $h = \Delta(t - y)$ and x and $\tilde{g}(x)$ respectively, and s_x^2 is the sample variance of x. The adjusted estimator remains asymptotically model-unbiased under model (3.1). Properties of the estimator (4.5) are under investigation.

We now turn to the case where only the population mean $\bar{X}$ is known. The estimators $\hat{F}_{ma}(t)$ and $F^*_{ma}(t)$ cannot be implemented in this case since they require the knowledge of all the population values x_j. We therefore adjust the estimator $\hat{\bar{H}} = \hat{F}(t)$ to obtain the following bias-adjusted estimator of $F(t)$

$$\hat{F}_a(t) = \hat{F}(t) + (s_{xh}/s_x^2)(\bar{X} - \bar{x}). \tag{4.6}$$

The conditional bias of $\hat{F}_a(t)$ is of the order $O_p(n^{-1})$, and as a result $\hat{F}_a(t)$ leads to conditionally valid inferences in large samples, unlike $\hat{F}(t)$. However, $\hat{F}_a(t)$ is asymptotically model-biased.

5. Calibration Estimators

5.1. General results

As noted in Section 1, calibration estimators satisfy certain consistency constraints with respect to auxiliary population information. The well-known post-stratified estimator and raking ratio estimator are simple examples of a calibration estimator. If ${}_jN$ $(j = 1, \ldots, J)$ denote the known population counts in J cells (e.g., cells based on age and sex categories), then post-stratification adjusts the basic weights $d_i(s)$ to $d_i^*(s) = ({}_jN/{}_j\hat{N})d_i(s)$ if sample element i belongs to the jth cell, where ${}_j\hat{N} = \sum_{i\in_j s} d_i(s)$ and ${}_js$ is the set of sample elements belonging to the jth cell. The revised weights $d_i^*(s)$ guarantee that the estimated counts in each of the J cells equal the corresponding population counts. Similarly, raking ratio estimators ensure consistency with two or more sets of marginal population counts; for example, row and column margins $\{{}_{i\cdot}N\}$ and $\{{}_{\cdot j}N\}$ in an $I \times J$ table $(i = 1, \ldots, I; j = 1, \ldots, J)$ of cell counts $\{{}_{ij}N\}$.

We first extend the method of Deville and Särndal (1992) to the general class of estimators (2.3) with basic weights $d_i(s)$. For simplicity we restrict ourselves to the chi-square distance

$$\phi_s = \sum_{i\in s}\{d_i(s) - d_i^*(s)\}^2/q_i(s)d_i(s) \tag{5.1}$$

where $q_i(s)$ are known positive weights unrelated to $d_i(s)$. The uniform weights

$q_i(s) = 1$ are commonly used, but other types of weights can also be used; for example, weights related to the variance structure of the errors ϵ_k in a super-population model

$$y_k = \mathbf{x}'_k\beta + \epsilon_k, \quad k = 1, \ldots, N. \tag{5.2}$$

Minimizing ϕ_s subject to consistency constraints (or calibration equations)

$$\sum_{i\in s} d^*_i(s)\mathbf{x}_i = \mathbf{X} \tag{5.3}$$

where $\mathbf{X} = (X_1, \ldots, X_p)'$ are known population totals, we get the revised (or calibration) weights $d^*_i(s)$ and the resulting estimator $Y^* = \sum_{i\in s} d^*_i(s)y_i$ which reduces to

$$\hat{Y}_{gr} = \hat{Y} + (\mathbf{X} - \hat{\mathbf{X}})'\hat{\mathbf{B}} \tag{5.4}$$

with

$$\hat{\mathbf{B}} = \left(\sum_{i\in s} d_i(s)q_i(s)\mathbf{x}_i\mathbf{x}'_i\right)^{-1} \times \left(\sum_{i\in s} d_i(s)q_i(s)\mathbf{x}_i y_i\right).$$

For the special case $d_i(s) = \pi_i^{-1}$ and $q_i(s) = q_i$, (5.4) reduces to the calibration estimator of Deville and Särndal (1992) which is identical to the generalized regression estimator of Särndal (1980). Huang and Fuller (1978) also used the generalized regression estimator with basic weights $d_i(s) = \pi_i^{-1}$, and developed an algorithm that produced nonnegative revised weights or standardized revised weights, $d^*_i(s)/\sum_{j\in s} d^*_j(s)$, that fall within a specified range, say $[0.25, 1.75]$. Bankier (1992) used a two-step extension of $\hat{Y}_{gr}$ in the context of 1991 Canadian Census which satisfies several consistency constraints at the weighting area (WA) level and at the same time ensures close agreement at the enumeration area (EA) level for number of households and number of persons.

If the set of auxiliary variables $x_1, \ldots, x_p$ includes at least one mutually exclusive and exhaustive set of indicator variables (as in the case of calibrating on known marginal population counts of a three-way table) and $q_i(s) = 1$, then it is easy to see that $\hat{Y} - \hat{\mathbf{X}}'\hat{\mathbf{B}} = 0$ and $\hat{Y}_{gr}$ reduces to the generalized "projection" estimator

$$\hat{Y}_p = \mathbf{X}'\hat{\mathbf{B}}. \tag{5.5}$$

This estimator is currently being used in the Canadian Labour Force Survey. If we partition the population U into households U_j with individual values (y_{jt}, x_{jt}), $t = 1, \ldots, M_j$; $j = 1, \ldots, N_c$, then the revised weights associated with (5.5) may be written as

$$d^*_{jt}(s) = \tilde{d}_j(s)\mathbf{X}' \times \left(\sum_{j\in s_c} \tilde{d}_j(s) \sum_{t\in U_j} \mathbf{x}_{jt}\mathbf{x}'_{jt}\right)^{-1} \mathbf{x}_{jt} \tag{5.6}$$

where s_c is the sample of households and $d_{jt}(s) = \tilde{d}_j(s)$ is the common basic weight attached to all members, t, of the household, j. It is clear from (5.6) that the revised weights are different for each member of the household, but in practice it is desirable to use the same weight for estimating totals of both family and individual characteristics. This is easily accomplished by replacing $\mathbf{x}_{jt}$ in (5.6) with the household mean value $\mathbf{z}_j = M_j^{-1}\sum_{t\in U}\mathbf{x}_{jt}$, noting that the population total $\sum_{j\in U_c} M_j\mathbf{z}_j = \mathbf{X}$ (see Lemaître and Dufour (1987) and Stukel and Boyer (1992) for the special case $\tilde{d}_j(s) = \pi_j^{-1}$). It should be noted, however, that the resulting estimator

$$\tilde{Y}_p = \mathbf{X}'\left(\sum_{j\in s_c} M_j\tilde{d}_j(s)\mathbf{z}_j\mathbf{z}'_j\right)^{-1} \times \left(\sum_{j\in s_c} \tilde{d}_j(s) \sum_{t\in U_j} \mathbf{z}_j y_{jt}\right)$$

may not be asymptotically more efficient than the basic estimator $\hat{Y}$, even if y_{jt} and $\mathbf{x}_{jt}$ are positively related.

We now propose an alternative calibration estimator that is asymptotically more efficient than the generalized regression estimator $\hat{Y}_{gr}$ or the basic estimator $\hat{Y}$. Consider the difference estimator $\hat{Y} + (\mathbf{X} - \hat{\mathbf{X}})'\mathbf{C}$ with a fixed p-vector of constants, $\mathbf{C}$, and minimize its variance with respect to $\mathbf{C}$. This leads to the optimum value $\mathbf{C}_{\text{opt}}$ and the resulting optimal estimator $\hat{Y} + (\mathbf{X} - \hat{\mathbf{X}})'\mathbf{C}_{\text{opt}}$, where $\mathbf{C}_{\text{opt}} = \Sigma_{\text{xx}}^{-1}\sigma_{y\text{x}}$ with Σ_{xx} and $\sigma_{y\text{x}}$ respectively denoting the $p \times p$ covariance matrix of $\hat{\mathbf{X}}$ and the p-vector of covariances, $\text{cov}(\hat{Y}, \hat{X}_t)$, $t = 1, \ldots, p$. Replacing Σ_{xx} and $\sigma_{y\text{x}}$ by their unbiased estimators $\hat{\Sigma}_{\text{xx}}$ and $\hat{\sigma}_{y\text{x}}$, we get the estimator

$$\hat{Y}_{\text{opt}} = \hat{Y} + (\mathbf{X} - \hat{\mathbf{X}})'\hat{\Sigma}_{\text{xx}}^{-1}\hat{\sigma}_{y\text{x}}. \qquad (5.7)$$

Fuller and Isaki (1981) and Montanari (1987) have also studied the optimal estimator (5.7) in the context of unistage designs and the Horvitz–Thompson estimator with basic weights $d_i(s) = \pi_i^{-1}$.

The estimator $\hat{Y}_{\text{opt}}$ is also a calibration estimator with respect to $\mathbf{x}$. This follows by letting $y = x_1$ (say) and noting that $\hat{Y} = \hat{X}_1$,

$$\hat{\Sigma}_{\text{xx}}^{-1}\hat{\sigma}_{x_1\text{x}} = (1, 0, \ldots, 0)' \qquad (5.8)$$

and $\hat{Y}_{\text{opt}} = \hat{X}_1 + (X_1 - \hat{X}_1) = X_1$. We used the following matrix results to obtain (5.8). For any nonsingular $p \times p$ matrix $\mathbf{A}$ with elements a_{ij}, $\mathbf{A}^{-1}$ is a $p \times p$ matrix with elements $|\mathbf{A}|^{-1}A_{ij}$, and $|\mathbf{A}|^{-1}\sum_j a_{1j}A_{1j} = 1$, $|\mathbf{A}|^{-1}\sum_j a_{1j}A_{ij} = 0$, $i \neq 1$, where A_{ij} is the cofactor of a_{ij}. If it is considered desirable to use the same weight for estimating totals of both family and individual characteristics, then we simply replace $\mathbf{x}_{jt}$ in (5.7) by $\mathbf{z}_j$ as before, noting that $\mathbf{X}$ and $\hat{\mathbf{X}}$ remain unchanged.

Following Cochran (1977, ch. 7), it is easily verified that, for large samples,

$$V(\hat{Y}_{gr}) - V(\hat{Y}_{\text{opt}}) \doteq (\mathbf{B} - \mathbf{C}_{\text{opt}})'\Sigma_{\text{xx}}(\mathbf{B} - \mathbf{C}_{\text{opt}}) \qquad (5.9)$$

where $\mathbf{B}$ is the vector of population regression coefficients. In the special case of simple random sampling, we have $\mathbf{B} = \mathbf{C}_{\text{opt}}$ and $\hat{Y}_{gr} = \hat{Y}_{\text{opt}}$, but in general the two estimators are not equal even for self-weighting designs.

Another advantage of $\hat{Y}_{\text{opt}}$ is that it leads to valid conditional inferences, noting that its conditional relative bias is asymptotically negligible given $\hat{\mathbf{X}}$. On the other hand, the conditional relative bias of $\hat{Y}_{gr}$ may not be asymptotically negligible, as shown earlier for the ratio estimator under SRS. Liu (1992) developed conditionally valid variance estimators for general stratified multistage design which together with $\hat{Y}_{\text{opt}}$ lead to conditionally valid inferences. Casady and Valliant (1993) also proposed $\hat{Y}_{\text{opt}}$ in the context of one-way post-stratification and studied empirically the conditional and unconditional properties of $\hat{Y}_{\text{opt}}$ and $\hat{Y}_{gr}$ in multistage sampling. They showed that $\hat{Y}_{\text{opt}}$ is the preferred estimator from a conditional point of view.

We now show that $\hat{Y}_{\text{opt}}$ can be expressed in the form $\sum_{i \in s} d_i^*(s)y_i$ for two commonly used sampling designs, stratified simple random sampling and stratified multistage sampling. That is, the same revised weights, $d_i^*(s)$, are used for all characteristics y, as in the case of $\hat{Y}_{gr}$.

5.2. *Stratified simple random sampling*

Suppose the population of size N is partitioned into L strata and a simple random sample, s_h, of size n_h is drawn from the N_h units in stratum h, independently for each $h = 1, \ldots, L$ $(\sum n_h = n)$. The customary unbiased estimator of Y is of the form $\hat{Y} = \sum_h \sum_{i \in s_h} d_{hi}(s)y_{hi}$ with $d_{hi}(s) = N_h/n_h$, where y_{hi} is the y-value of the ith unit in the hth stratum. The elements of $\hat{\Sigma}_{xx}$ and $\hat{\sigma}_{yx}$ may be expressed as

$$\hat{\sigma}_{x_\ell x_m} = \sum_{h=1}^{L}\sum_{i\in s_h} x'_{\ell hi}x'_{mhi} \qquad (5.10)$$
$$\ell, m = 1,\ldots,p$$

$$\hat{\sigma}_{yx_\ell} = \sum_{h=1}^{L}\sum_{i\in s_h} y'_{hi}x'_{\ell hi}, \quad \ell = 1,\ldots,p \qquad (5.11)$$

with

$$x'_{\ell hi} = N_h[(1-f_h)/n_h(n_h-1)]^{1/2} \times (x_{\ell hi} - \bar{x}_{\ell h})$$
$$y'_{hi} = N_h[(1-f_h)/n_h(n_h-1)]^{1/2} \times (y_{hi} - \bar{y}_h)$$

where $f_h = n_h/N_h$, $x_{\ell hi}$ is the value of x_ℓ for the (hi)th unit, and $\bar{y}_h = \sum_{i\in s_h} y_{hi}/n_h$, $\bar{x}_{\ell h} = \sum_{i\in s_h} x_{\ell hi}/n_h$. It follows from (5.10) and (5.11) that the calibration weights associated with $\hat{Y}_{\text{opt}}$ may be written as

$$d^*_{hi}(s) = d_{hi}(s)[1 + N_h(1-f_h)(n_h-1)^{-1} \times (\mathbf{X}-\hat{\mathbf{X}})'\hat{\Sigma}_{xx}^{-1}(\mathbf{x}_{hi} - \bar{\mathbf{x}}_h)] \qquad (5.12)$$

where $\mathbf{x}_{hi} = (x_{1hi},\ldots,x_{phi})'$ and $\bar{\mathbf{x}}_h = \sum_{i\in s_h}\mathbf{x}_{hi}/n_h$. Note that we have only $n-L$ independent observations (y'_{hi}, x'_{hi}) to estimate $\mathbf{C}_{\text{opt}}$ since $\sum_i x'_{\ell hi} = 0$ and $\sum_i y'_{hi} = 0$, whereas $\hat{\mathbf{B}}$ is based on n independent observations.

5.3. *Stratified multistage sampling*

Large-scale surveys often employ stratified multistage designs with large numbers of strata, L, and relatively few primary sampling units (clusters), sampled within each stratum h. We assume that subsampling within sampled clusters is performed to ensure unbiased estimation of cluster totals Y_{hi}. The customary unbiased estimator of Y is of the form $\hat{Y} = \sum_{(hik)\in s} d_{hik}(s)y_{hik}$, where s is the sample of elements and y_{hik} is the y-value associated with the sample element $(hik) \in s$. At the stage of variance estimation, the calculations are greatly simplified by treating the sample as if the sample clusters are sampled with replacement. This approximation leads to overestimation of the variance of $\hat{Y}$. This overestimation can be substantial unless the first-stage sampling fractions are small.

Writing $\hat{Y}$ as $\hat{Y} = \sum_h \bar{r}_h$, with $r_{hi} = \sum_k (n_h d_{hik}(s))y_{hik}$ and $\bar{r}_h = \sum_i r_{hi}/n_h$, the above estimator of variance of $\hat{Y}$ is simply given by

$$v(\hat{Y}) = \sum_h\sum_i (r_{hi} - \bar{r}_h)^2/n_h(n_h-1) \qquad (5.12)$$

where $n_h(\geqslant 2)$ is the number of sample clusters from stratum h. It now follows that $\hat{\Sigma}_{xx}$ and $\hat{\sigma}_{yx}$ may be expressed as

$$\hat{\sigma}_{x_\ell x_m} = \sum_h\sum_i u'_{\ell hi}u'_{mhi} \qquad (5.13)$$
$$\ell, m = 1,\ldots,p$$

$$\hat{\sigma}_{yx_\ell} = \sum_h\sum_i r'_{hi}u'_{\ell hi} \qquad (5.14)$$
$$\ell = 1,\ldots,m$$

with

$$u'_{\ell hi} = [n_h(n_h-1)]^{-1/2}(u_{\ell hi} - \bar{u}_{\ell h})$$
$$r'_{hi} = [n_h(n_h-1)]^{-1/2}(r_{hi} - \bar{r}_h)$$
$$u_{\ell hi} = \sum_k (n_h d_{hik}(s))x_{\ell hik}$$
$$\bar{u}_{\ell h} = \sum_i u_{\ell hi}/n_h.$$

Hence, the calibration weights associated with $\hat{Y}_{\text{opt}}$ may be written as

$$d^*_{hik}(s) = d_{hik}(s)\left[1 + \frac{1}{(n_h-1)} \times (\mathbf{X}-\hat{\mathbf{X}})'\hat{\Sigma}_{xx}^{-1}(\mathbf{u}_{hi} - \bar{\mathbf{u}}_h)\right] \qquad (5.15)$$

where $\mathbf{u}_{hi} = (u_{1hi}, \ldots, u_{phi})'$ and $\bar{\mathbf{u}}_h = \sum_i \mathbf{u}_{hi}/n_h$. Note that we have only $\Sigma n_h - L$ independent observations $(r'_{hi}, \mathbf{u}'_{hi})$ to estimate $\mathbf{C}_{\text{opt}}$ since $\sum_i u'_{\ell hi} = 0$ and $\sum_i r'_{hi} = 0$. The estimator of $\mathbf{C}_{\text{opt}}$, therefore, may not be stable unless $\sum_h n_h - L$ is large relative to the number of auxiliary variables, p. Note that $\sum_h n_h$ is the total number of sample clusters.

6. References

Bankier, M.D. (1992). Two-Step Generalized Least Squares Estimation in the 1991 Canadian Census. Paper presented at the Workshop on Uses of Auxiliary Information in Surveys, October 4–6, 1992, Örebro, Sweden.

Bethlehem, J.G. and Keller, W.J. (1987). Linear Weighting of Sample Survey Data. Journal of Official Statistics, 3, 141–153.

Brewer, K.R.W. (1963). Ratio Estimation and Finite Populations: Some Results Deducible from the Assumption of an Underlying Stochastic Process. Australian Journal of Statistics, 5, 93–105.

Casady, R.J. and Valliant, R. (1993). Conditional Properties of Post-Stratified Estimators under Normal Theory. Survey Methodology, 19, 183–192.

Cochran, W.G. (1977). Sampling Techniques (3rd Edition). New York: John Wiley.

Deville, J. and Särndal, C.E. (1992). Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, 87, 376–382.

Fuller, W.A. and Isaki, C.T. (1981). Survey Design under Superpopulation Models. In Current Topics in Survey Sampling (D. Krewski, R. Platek, J.N.K. Rao, and M.P. Singh eds.), New York: Academic Press, 196–226.

Godambe, V.P. (1955). A Unified Theory of Sampling from Finite Populations. Journal of the Royal Statistical Society, Ser. B, 17, 269–278.

Godambe, V.P. (1989). Estimation of Cumulative Distribution Function of a Survey Population. Technical Report, University of Waterloo.

Hansen, M.H., Madow, W.G., and Tepping, B.J. (1983). An Evaluation of Model-Dependent and Probability-Sampling Inferences in Sample Surveys. Journal of the American Statistical Association, 78, 776–793.

Hidiroglou, M.A., Fuller, W.A. and Hickman, R.D. (1976). SUPER CARP. Statistical Laboratory, Iowa State University, Ames, Iowa.

Huang, E.T. and Fuller, W.A., (1978). Nonnegative Regression Estimation for Sample Survey Data. Proceedings of the Social Statistics Section, American Statistical Association, 300–305.

Kott, P.J. (1990). Estimating the Conditional Variance of a Design Consistent Regression Estimator. Journal of Statistical Planning and Inference, 24, 287–296.

Krieger, A.M. and Pfeffermann, P. (1992). Maximum Likelihood Estimation from Complex Surveys. Unpublished Technical Report, Department of Statistics, University of Pennsylvania.

Lemaître, G. and Dufour, J. (1987). An Integrated Method for Weighting Persons and Families. Survey Methodology, 13, 199–207.

Liu, J. (1992). Inference from Stratified Samples: Application of Edgeworth Expansions. Ph.D. Thesis, Carleton University.

Montanari, G.E. (1987). Post-Sampling Efficient QR-Prediction in Large-Sample Surveys. International Statistical Review, 55, 191–202.

Rao, J.N.K. (1979). On Deriving Mean Square Errors and Their Non-Negative

Unbiased Estimators in Finite Population Sampling. Journal of the Indian Statistical Association, 17, 125–136.

Rao, J.N.K. and Liu, J. (1992). On Estimating Distribution Functions from Sample Survey Data Using Supplementary Information at the Estimation Stage. In Nonparametric Statistics and Related Topics (A.K.Md.E. Saleh, ed.), Amsterdam: Elsevier Science Publishers, 399–407.

Rao, J.N.K., Kovar, J.G., and Mantel, H.J. (1990). On Estimating Distribution Functions and Quantiles from Survey Data Using Auxiliary Information. Biometrika, 77, 365–375.

Robinson, J. (1987). Conditioning Ratio Estimates Under Simple Random Sampling. Journal of the American Statistical Association, 82, 826–831.

Royall, R.M. (1970). On Finite Population Sampling Theory Under Certain Linear Regression Models. Biometrika, 57, 377–387.

Royall, R.M. and Cumberland, W.G. (1981). An Empirical Study of the Ratio Estimator and Estimator of its Variance. Journal of the American Statistical Association, 76, 66–88.

Särndal, C.E. (1980). On π-Inverse Weighting Versus Best Linear Unbiased Weighting in Probability Sampling. Biometrika, 67, 639–650.

Särndal, C.E., Swensson, B., and Wretman, J.H. (1989). The Weighted Regression Technique for Estimating the Variance of the Generalized Regression Estimate. Biometrika, 76, 527–537.

Stukel, D. and Boyer, R. (1992). Calibration Estimation: An Application to the Canadian Labour Force Survey. Technical Report, Social Survey Methods Division, Statistics Canada.

Received January 1993
Revised October 1993