International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 10, 2010, pp 9-17

On A Method of Estimation in the Presence of Non Response Using Auxiliary Information

L. N. Sahoo, G. Mishra
Department of Statistics
Utkal University
Bhubaneswar 751004, India
E-mail: lnsahoostatuu@rediffmail.com

S. R. Nayak
Department of Statistics
Dhenkanal College
Dhekanal, India

[Received November 15, 2006; Revised October 7, 2008; Accepted April 16, 2009]

Abstract

This paper deals with an estimation of a finite population mean in the presence of non response gathering information on two auxiliary variables. It is assumed that full response on the auxiliary variables is available for the intended sample, and the population mean of one auxiliary variable is known whereas that mean of the other is unknown. On adopting the technique of sub sampling of non respondents, as conjectured in Hansen and Hurwitz (1946), a class of estimators is developed. An asymptotic minimum attainable variance for the class is provided and the best estimator is identified. Performance of the non response adjusted mechanism developed here, is also examined analytically and empirically.

Keywords and Phrases: Asymptotic Variance, Auxiliary Variable, Non Response

AMS Classification: 62D05.

1 Introduction

Consider a finite population U of N distinct and identifiable units. Let $\mathcal{Y}$ be the survey variable and y_i its value for the unit i of U, $i = 1, 2, \cdots, N$. In order to estimate the unknown population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$, a sample s of size n is drawn from U

according to simple random sampling without replacement (SRSWOR). Assume that only a subset s_1 of n_1 units of s respond on $\mathcal{Y}$ but the remaining $n_2 = (n - n_1)$ units, constituting a subset $s_2 = (s - s_1)$, do not provide any response. In this case the population U of N units is visualized to be consisting of the response stratum or group U_1 and the non response stratum or group U_2 of sizes N_1 and $N_2 = (N - N_1)$ respectively. It is also convenient to think that s_1 and s_2 *i.e.*, the samples of respondents and non respondents are independent samples from U_1 and U_2 respectively. Let $\bar{Y}_j = \frac{1}{N_j}\sum_{i=1}^{N_j} y_i$ and $\bar{y}_j = \frac{1}{n_j}\sum_{i\in s_j} y_i$ be the means of U_j and s_j respectively such that $W_1\bar{Y}_1 + W_2\bar{Y}_2$ and $\bar{y} = w_1\bar{y}_1 + w_2\bar{y}_2$, where $W_j = \frac{N_j}{N}$ and $w_j = \frac{n_j}{n}$, $j = 1, 2$. Obviously, $E(w_j) = W_j$ and $E(\bar{y}) = \bar{Y}$. However, the sample mean $\bar{y}_1$ is unbiased for $\bar{Y}_1$ but has a bias equal to $W_2\left(\bar{Y}_1 - \bar{Y}_2\right)$ in estimating $\bar{Y}$.

The method of sub sampling of non respondents suggested by Hansen and Hurwitz (1946) is a successful technique for handling non response situation compared to other popular techniques. This technique consists of drawing a sub sample s_m of size $m = n_2/k,\ k \geq 1$, from s_2. It is suitable for surveys in which the initial attempt is made by the mail or telephone calls, perhaps computer aided and the sub sample of persons who do not respond are approached by more expensive method of personal interview. Assuming that response is available from all units of s_m, the authors recommended the estimator

$$\bar{y}^* = w_1\bar{y}_1 + w_2\bar{y}_m,$$

for $\bar{Y}$, where $\bar{y}_m = \frac{1}{m}\sum_{i\in s_m} y_i$. From Cochran (1977, p.329), it may be seen that $E(\bar{y}^*) = \bar{Y}$, and

$$V(\bar{y}^*) = \frac{1-f}{n}S_y^2 + W_2\frac{k-1}{n}S_{y2}^2, \tag{1}$$

where $f = \frac{n}{N}$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2$ and $S_{y2}^2 = \frac{1}{N_2-1}\sum_{i=1}^{N_2}\left(y_i - \bar{Y}_2\right)^2$.

For a situation in which the population mean $\bar{X}\left(= \frac{1}{N}\sum_{i=1}^{N} x_i\right)$ of an auxiliary variable $\mathcal{X}$, taking value x_i for the ith unit of U is available, one can often improve substantially over the Hansen and Hurwitz's (1946) estimator. Motivated by this consideration, Rao (1986, 1990) suggested two simple estimators, defined by

$t_{hr} = \bar{y}^*\frac{\bar{X}}{\bar{x}}$ and $t_{hrg} = \bar{y}^* - b_{yx}^*\left(\bar{x} - \bar{X}\right)$,

where $\bar{X} = W_1\bar{X}_1 + W_2\bar{X}_2$, $\bar{X}_j = \frac{1}{N_j}\sum_{i=1}^{N_j} x_i$ $(j = 1, 2)$, $\bar{x} = w_1\bar{x}_1 + w_2\bar{x}_2$, $\bar{x}_j = \frac{1}{n_j}\sum_{i\in s_j} x_i$, $(j = 1, 2)$, $\bar{x}_m = \frac{1}{m}\sum_{i\in s_m} x_i$, $b_{yx}^* = \frac{\hat{S}_{yx}}{\hat{S}_x^2}$ is an estimator of $\beta_{yx} = \frac{S_{yx}}{S_x^2}$

such that $\hat{S}_{yx}$ and $\hat{S}_x^2$ are unbiased estimators of $S_{yx} = \frac{1}{N-1}\sum\limits_{i=1}^{N}\left(y_i - \bar{Y}\right)\left(x_i - \bar{X}\right)$ and $S_x^2 = \frac{1}{N-1}\sum\limits_{i=1}^{N}\left(x_i - \bar{X}\right)^2$ respectively, given by

$$(n-1)\hat{S}_{yx} = (n_1 - 1)s_{yx1} + k(m-1)s_{yxm} + nw_1w_2(\bar{y}_1 - \bar{y}_m)(\bar{x}_1 - \bar{x}_m) + w_2(k-1)\frac{s_{yxm}}{n},$$
$$(n-1)\hat{S}_x^2 = (n_1 - 1)s_{x1}^2 + k(m-1)s_{xm}^2 + nw_1w_2(\bar{x}_1 - \bar{x}_m)^2 + w_2(k-1)\frac{s_{xm}^2}{n},$$

and $s_{yxj} = \frac{1}{n_j-1}\sum\limits_{i\in s_j}(y_i - \bar{y}_j)(x_i - \bar{x}_j), \quad s_{xj}^2 = \frac{1}{n_j-1}\sum\limits_{i\in s_j}(x_i - \bar{x}_j)^2 \quad (j = 1, 2),$
$s_{yxm} = \frac{1}{m-1}\sum\limits_{i\in s_m}(y_i - \bar{y}_m)(x_i - \bar{x}_m), \quad s_{xm}^2 = \frac{1}{m-1}\sum\limits_{i\in s_m}(x_i - \bar{x}_m)^2.$

Sinha (2001) proposed a class of estimators with the help of multi-auxiliary variables whose population means are known. In case of one auxiliary variable $\mathcal{X}$, this class is defined by

$$t_h = h\left(\bar{y}^*, u\right),$$

where $h\left(\bar{y}^*, u\right)$ is a function of $\bar{y}^*$ and $u = \frac{\bar{x}}{\bar{X}}$ such that $h\left(\bar{Y}, 1\right) = \bar{Y}$. The estimators $\bar{y}^*, t_{hr}, t_{hrg}$ and the product counterpart of t_{hr} defined by $t_{hp} = \bar{y}^*\frac{\bar{x}}{\bar{X}}$ can be easily viewed as particular members of t_h. t_{hrg} in this case is considered as an optimum estimator of class as it possesses the minimum asymptotic variance.

This paper is an attempt to generate a class of estimators for $\bar{Y}$ gathering information on two auxiliary variables $\mathcal{X}$ and $\mathcal{Z}$ by adopting the mechanism of sub sampling of non respondents. It is assumed that there is full response on both auxiliary variables at sample level and population mean $\bar{X}$ (or total X) of $\mathcal{X}$ is known accurately but no information is available on the population mean $\bar{Z}$(or total Z) of $\mathcal{Z}$. To convince readers that situations do arise where such assumptions are met, we may refer to a survey relating to high-income families where data are collected through questionnaire on the variables $\mathcal{Y}$ = savings or consumer durable expenditures, $\mathcal{X}$ = family size, and $\mathcal{Z}$ = family income. Here, all units in s may respond on $\mathcal{X}$ and $\mathcal{Z}$ but some of the units may not respond on $\mathcal{Y}$ because, usually, high-income earners are less inclined to respond on savings and consumer durable expenditures. There is no scope to obtain the value of Z, whereas X *i.e.*, the total population of the study area can be known from the census records. Similarly, if $\mathcal{Y}$, $\mathcal{X}$ and $\mathcal{Z}$ respectively denote percentage of children below 18 engaged in agriculture, number of households and educated parents of a village, response on $\mathcal{Y}$ for some units in the sample may not be available due to adamant refusal of concerned respondents at the first call.

We shall use the following additional notations, for $j = 1, 2$:

$$\bar{Z}_j = \frac{1}{N_j}\sum_{i=1}^{N_j} z_i,\ \bar{Z} = W_1\bar{Z}_1 + W_2\bar{Z}_2,\ \bar{z}_j = \frac{1}{n_j}\sum_{i\in s_j} z_i,\ \bar{z} = w_1\bar{z}_1 + w_2\bar{z}_2,\ \bar{z}_m = \frac{1}{m}\sum_{i\in s_m} z_i,$$

and the quantities $S_{zj}^2, S_{yzj}, S_{xzj}, s_{zj}^2, s_{yzj}, s_{xzj}, s_{zm}^2, s_{yzm}$ and s_{xzm} are defined analogously.

2 The Suggested Class of Estimators

For a given $s_m \subset s_2$, let $d = (\bar{y}_m, \bar{x}_m, \bar{x}_2, \bar{z}_m, \bar{z}_2)$ assume values in 5-dimensional real space R_5 containing the point $D = \left(\bar{Y}_2, \bar{X}_2, \bar{X}_2, \bar{Z}_2, \bar{Z}_2\right)$. Let $\phi_2(d)$ be a function of d such that $\phi_2(D) = \bar{Y}_2$, and it admits the following regularity conditions:
(i) $\phi_2(d)$ is continuous in R_5, and
(ii) The first and second order partial derivatives of $\phi_2(d)$ w.r.t. all arguments exist and are continuous in R_5.

Hence, a class of estimators for $\bar{Y}$ may be defined by
$\hat{\bar{Y}}_s = w_1\bar{y}_1 + w_2\phi_2(d)$.

Further, for a given s, let $\left(\hat{\bar{Y}}_s, \bar{x}\right)$ assume values in 2-dimensional real space R_2 containing the point $\left(\bar{Y}, \bar{X}\right)$. Define $\phi\left(\hat{\bar{Y}}_s, \bar{x}\right)$ as a function of $\hat{\bar{Y}}_s$ and $\bar{x}$, different from ϕ_2, and admitting the said regularity conditions in R_2 with $\phi\left(\bar{Y}, \bar{X}\right) = \bar{Y}$. Then, to estimate $\bar{Y}$ we propose a general class of estimators defined by
$\hat{\bar{Y}} = \phi\left(\hat{\bar{Y}}_s, \bar{x}\right)$.

$\hat{\bar{Y}}$ being a non-linear composite function of many statistics, derivation of expressions for bias and variance seems to be more difficult. However, to circumvent much difficulty, we apply Taylor linearization technique described in many standard text books [*cf.*, Sarndal, Swensson and Wretman (1992, p.172)] in order to obtain an approximate linear form of $\hat{\bar{Y}}$ under the earlier mentioned regularity conditions.

On expanding $\hat{\bar{Y}} = \phi\left(\hat{\bar{Y}}_s, \bar{x}\right)$ around the point $\left(\bar{Y}, \bar{X}\right)$ by the first order Taylor's series and noting that $\phi_0 = 1$, it may be observed that

$$\hat{\bar{Y}} \cong \hat{\bar{Y}}_s + \phi_1\left(\bar{x} - \bar{X}\right) = w_1\bar{y}_1 + w_2\phi_2(d) + \phi_1\left(\bar{x} - \bar{X}\right), \quad (2)$$

where $\phi_0 = \left.\frac{\partial\hat{\bar{Y}}}{\partial\hat{\bar{Y}}_s}\right|_{\substack{\hat{\bar{Y}}_s = \bar{Y} \\ \bar{x} = \bar{X}}}$ and $\phi_1 = \left.\frac{\partial\hat{\bar{Y}}}{\partial\bar{x}}\right|_{\substack{\hat{\bar{Y}}_s = \bar{Y} \\ \bar{x} = \bar{X}}}$.

Similarly, noting that $\phi_{20} = 1$, an expansion of $\phi_2(d)$ about the point D in a first order Taylor's series provides an asymptotic linear form as

$$\phi_2(d) \cong \bar{y}_m + \phi_{21}\left(\bar{x}_m - \bar{X}_2\right) - \phi_{21}\left(\bar{x}_2 - \bar{X}_2\right) + \phi_{22}\left(\bar{z}_m - \bar{Z}_2\right) - \phi_{22}\left(\bar{z}_2 - \bar{Z}_2\right), \quad (3)$$

where $\phi_{20} = \left.\frac{\partial\phi_2(d)}{\partial\bar{y}_m}\right|_{d=D}$, $\phi_{21} = \left.\frac{\partial\phi_2(d)}{\partial\bar{x}_m}\right|_{d=D}$ and $\phi_{22} = \left.\frac{\partial\phi_2(d)}{\partial\bar{z}_m}\right|_{d=D}$.

Hence, from (2) and (3), $\hat{\bar{Y}}$ can be written as

$$\hat{\bar{Y}} = \bar{y} + w_2\left[(\bar{y}_m - \bar{y}_2) + \phi_{21}(\bar{x}_m - \bar{x}_2) + \phi_{22}(\bar{z}_m - \bar{z}_2)\right] + \phi_1\left(\bar{x} - \bar{X}\right). \quad (4)$$

From (4) we see that, to a first order of approximation, $E\left(\hat{\hat{Y}}\right) \cong \bar{Y}$ so that an expression for asymptotic variance (or mean square error) is given by

$$\begin{aligned} V\left(\hat{\hat{Y}}\right) &= \tfrac{1-f}{n}\left(S_y^2+\phi_1^2 S_x^2+2\phi_1 S_{yx}\right) \\ &+ W_2\tfrac{k-1}{n}\left(S_{y2}^2+\phi_{21}^2 S_{x2}^2+\phi_{22}^2 S_{z2}^2+2\phi_{21}S_{yx2}+2\phi_{22}S_{yz2}+2\phi_{21}\phi_{22}S_{xz2}\right). \end{aligned} \tag{5}$$

Minimizing $V\left(\hat{\hat{Y}}\right)$ for ϕ_{21}, ϕ_{22} and ϕ_1, we get

$$\begin{aligned} \phi_{21}^{(opt)} &= -\tfrac{\beta_{yx2}-\beta_{yz2}\beta_{zx2}}{1-\beta_{zx2}\beta_{xz2}} = -\beta_{yx.z2}, \\ \phi_{22}^{(opt)} &= -\tfrac{\beta_{yz2}-\beta_{yx2}\beta_{xz2}}{1-\beta_{zx2}\beta_{xz2}} = -\beta_{yz.x2}, \\ \phi_{1}^{(opt)} &= -\beta_{yx}, \end{aligned}$$

where β_{yx2}, β_{xz2} etc. are the simple regression coefficients and $\beta_{yx.z2}$, $\beta_{yz.x2}$ are the partial regression coefficients in the non response stratum U_2. Use of these optimum values in (5) provides us a minimum asymptotic variance, called the minimum variance bound (MVB) of the class, by the expression

$$\min V\left(\hat{\hat{Y}}\right) = \frac{1-f}{n}S_y^2\left(1-\rho_{yx}^2\right)+W_2\frac{k-1}{n}S_{y2}^2\left(1-\rho_{y.xz2}^2\right), \tag{6}$$

where ρ_{yx} is the simple correlation coefficient between $\mathcal{Y}$ and $\mathcal{X}$ in U, and $\rho_{y.xz2}$ is the multiple correlation coefficient of $\mathcal{Y}$ on $\mathcal{X}$ and $\mathcal{Z}$ in U_2. An estimator attaining this bound *i.e.*, MVB estimator can be defined by

$$\hat{\hat{Y}}_{rg}^{*} = w_1\bar{y}_1 + w_2\left\{\bar{y}_m - \phi_{21}^{(opt)}\left(\bar{x}_m-\bar{x}_2\right) - \phi_{22}^{(opt)}\left(\bar{z}_m-\bar{z}_2\right)\right\} - \phi_1^{(opt)}\left(\bar{x}-\bar{X}\right).$$

But, estimating the unknown parametric functions $\phi_{21}^{(opt)}, \phi_{22}^{(opt)}$ and $\phi_1^{(opt)}$ by
$b_{yx.zm} = -\frac{b_{yxm}-b_{yzm}b_{zxm}}{1-b_{zxm}b_{xzm}}$, $b_{yz.xm} = -\frac{b_{yzm}-b_{yxm}b_{xzm}}{1-b_{zxm}b_{xzm}}$ and b_{yx}^*
respectively, where $b_{yxm} = \frac{s_{yxm}}{s_{xm}^2}, b_{xzm} = \frac{s_{xzm}}{s_{zm}^2}$ etc., we obtain the following regression-type estimator, declaring it as a MVB estimator:

$$\hat{\hat{Y}}_{rg} = w_1\bar{y}_1 + w_2\left\{\bar{y}_m - b_{yx.zm}\left(\bar{x}_m-\bar{x}_2\right) - b_{yz.xm}\left(\bar{z}_m-\bar{z}_2\right)\right\} - b_{yx}^*\left(\bar{x}-\bar{X}\right).$$

This estimator is not only difficult to compute in practice but also biased for $\bar{Y}$. However, to a first order of approximation, $V\left(\hat{\hat{Y}}_{rg}\right) = V\left(\hat{\hat{Y}}_{rg}^*\right)$.

3 Some Particular Cases of $\hat{\hat{Y}}$

When there is no explicit use of auxiliary variables, $\hat{\hat{Y}} = \bar{y}^*$. But, if the emphasis is given on the use of $\mathcal{X}$ or $\mathcal{Z}$ or both, $\hat{\hat{Y}}$ produces a family of estimators because it can be reduced to a series of estimators for various selections of the functions ϕ_2 and ϕ. For instance, the following simple estimators are particular cases of $\hat{\hat{Y}}$ whose design-based large sample properties can be studied in the usual way.

$$\hat{\hat{Y}}_r = \left[w_1\bar{y}_1 + w_2\bar{y}_m \frac{\bar{x}_2}{\bar{x}_m}\frac{\bar{z}_2}{\bar{z}_m}\right]\left(\frac{\bar{X}}{\bar{x}}\right), \quad \hat{\hat{Y}}_p = \left[w_1\bar{y}_1 + w_2\bar{y}_m \frac{\bar{x}_m}{\bar{x}_2}\frac{\bar{z}_m}{\bar{z}_2}\right]\left(\frac{\bar{x}}{\bar{X}}\right),$$

$$\begin{aligned}
\hat{\hat{Y}}_1 &= \left[w_1\bar{y}_1 + w_2\left\{\bar{y}_m + \lambda_1\left(\bar{x}_m - \bar{x}_2\right) + \lambda_2\left(\bar{z}_m - \bar{z}_2\right)\right\}\right]\left(\tfrac{\bar{x}}{\bar{X}}\right)^{\lambda} \\
\hat{\hat{Y}}_2 &= \left[w_1\bar{y}_1 + w_2\bar{y}_m\left(\tfrac{\bar{x}_m}{\bar{x}_2}\right)^{\lambda_1}\left(\tfrac{\bar{z}_m}{\bar{z}_2}\right)^{\lambda_2}\right]\left(\tfrac{\bar{x}}{\bar{X}}\right)^{\lambda} \\
\hat{\hat{Y}}_3 &= \left[w_1\bar{y}_1 + w_2\bar{y}_m\tfrac{\bar{x}_2}{\lambda_1\bar{x}_m + (1-\lambda_1)\bar{x}_2}\tfrac{\bar{z}_2}{\lambda_2\bar{z}_m + (1-\lambda_2)\bar{z}_2}\right]\tfrac{\bar{X}}{\lambda\bar{x} + (1-\lambda)\bar{X}}.
\end{aligned}$$

Suitable selections of ϕ_2 and ϕ on the consideration of different auxiliary quantities can also enable $\hat{\hat{Y}}$ to define some other classes of estimators. Nevertheless, whatever the new class is, its MVB estimator is always a regression-type estimator like $\hat{\hat{Y}}_{rg}$. Let us now consider the following specific cases:

3.1 Suppose that $\bar{X}$ is unknown. Then, $\hat{\hat{Y}} = \hat{\hat{Y}}_s$, a class of separate estimators for $\bar{Y}$ based on $\mathcal{X}$ and $\mathcal{Z}$ whose MVB estimator is given by

$$\hat{\hat{Y}}_{srg} = w_1\bar{y}_1 + w_2\left\{\bar{y}_m - b_{yx.zm}\left(\bar{x}_m - \bar{x}_2\right) - b_{yz.xm}\left(\bar{z}_m - \bar{z}_2\right)\right\}.$$

3.2 When $\bar{X}$ is unknown and no information is available on $\mathcal{X}$, then

$$\hat{\hat{Y}} \to \ell_s = w_1\bar{y}_1 + w_2 g_2\left(\bar{y}_m, \bar{x}_m, \bar{x}_2\right),$$

a class of separate estimators based on $\mathcal{X}$ only whose MVB is given by

$$\ell_{srg} = w_1\bar{y}_1 + w_2\left\{\bar{y}_m - b_{yxm}\left(\bar{x}_m - \bar{x}_2\right)\right\},$$

mentioned earlier in Rao (1990). We also observe that the ratio estimator

$$\ell_{sr} = w_1\bar{y}_1 + w_2\bar{y}_m\frac{\bar{x}_2}{\bar{x}_m},$$

defined by Rao (1986) and the product estimator of the form

$$\ell_{sp} = w_1\bar{y}_1 + w_2\bar{y}_m\frac{\bar{x}_m}{\bar{x}_2},$$

are particular cases of ℓ_s.

3.3 If $\mathcal{X}$ is not involved at the stratum level, then

$$\hat{\hat{Y}} \to \hat{\hat{Y}}_q = q\left(w_1\bar{y}_1 + w_2 q_2(\bar{y}_m, \bar{z}_m, \bar{z}_2), \bar{x}\right),$$

providing a MVB estimator

$$\hat{\hat{Y}}_{qrg} = w_1\bar{y}_1 + w_2\left\{\bar{y}_m - b_{yzm}\left(\bar{z}_m - \bar{z}_2\right)\right\} - b^*_{yx}\left(\bar{x} - \bar{X}\right).$$

3.4 If the estimation procedure is carried out with the involvement of $\mathcal{X}$ only

$$\hat{\hat{Y}} \to \ell_g = g\left(\ell_s, \bar{x}\right).$$

This class of estimators covers Rao's (1986) ratio estimator $\ell_{gr} = \ell_{sr}\frac{\bar{X}}{\bar{x}}$, Rao's (1990) regression estimator $\ell_{grg} = \ell_{srg} - b^*_{yx}\left(\bar{x} - \bar{X}\right)$, and the product estimator $\ell_{gp} = \ell_{sr}\frac{\bar{x}}{\bar{X}}$ as its members. Here ℓ_{grg} is a MVB estimator of ℓ_g.

4 Precision of the Class

From (1) and (5), we see that an estimator of $\hat{\hat{Y}}$ is asymptotically more precise than Hansen and Hurwitz estimator $\bar{y}^*$ if

$$\beta_{yx} < -\frac{\phi_1}{2} \quad and \quad \beta_{yx2} < -\frac{\phi_{21}}{2} - \frac{S^2_{z2}\left(\phi^2_{22} + 2\phi_{22}\beta_{yz2} + 2\phi_{21}\phi_{22}S_{xz2}\right)}{2\phi_{21}S^2_{x2}}. \tag{7}$$

Thus, by taking into account these restrictions, one may apply the suggested technique to hope for an improvement over $\bar{y}^*$ by using the amount of auxiliary information available. But, in the usual practice, this task is not so easy. Because, the conditions mainly depend on the choices of the functions ϕ and ϕ_2, and do not lead to a straight forward conclusion unless nature of these functions are known. A series of other similar complicated sufficient conditions can also be derived if we discuss on the superiority of $\hat{\hat{Y}}$ over other classes considered earlier under the asymptotic variance criterion. In view of this when optimum MVBs can be computed they can be used to compare precision of different classes.

Table 1: Minimum Variance Bounds of Different Classes

Class	Minimum Variance Bound
t_h	$\frac{1-f}{n}S^2_y\left(1-\rho^2_{yx}\right) + W_2\frac{k-1}{n}S^2_{y2}$
ℓ_s	$\frac{1-f}{n}S^2_y + W_2\frac{k-1}{n}S^2_{y2}\left(1-\rho^2_{yx2}\right)$
ℓ_g	$\frac{1-f}{n}S^2_y\left(1-\rho^2_{yx}\right) + W_2\frac{k-1}{n}S^2_{y2}\left(1-\rho^2_{yx2}\right)$
$\hat{\hat{Y}}_q$	$\frac{1-f}{n}S^2_y\left(1-\rho^2_{yx}\right) + W_2\frac{k-1}{n}S^2_{y2}\left(1-\rho^2_{yz2}\right)$
$\hat{\hat{Y}}_s$	$\frac{1-f}{n}S^2_y + W_2\frac{k-1}{n}S^2_{y2}\left(1-\rho^2_{y.xz2}\right)$

Table 1 provides expressions for minimum variance bounds of different comparable classes. Considering these expressions together with (6), the following results are obtained:

(i) Minimum variances of the different classes of estimators are unconditionally less than $V(\bar{y}^*)$.

(ii) Both ℓ_g and $\hat{\hat{Y}}_q$ are superior to t_h

(iii) ℓ_s is inferior to $\hat{\hat{Y}}_s$

(iv) ℓ_g is superior or inferior to $\hat{\hat{Y}}_q$ according as $\rho_{yx2} >$ or $< \rho_{yz2}$

(v) $\hat{\hat{Y}}$ is superior to other classes considered here.

In order to evaluate the relative performance of the suggested method of estimation over other methods in respect of MVB criterion empirically, we consider life data on 3 populations as described below:

Population 1 [Murthy (1967), p.228]. N = 80 factories, $\mathcal{Y}$ = output, $\mathcal{X}$= fixed capital, $\mathcal{Z}$ = number of workers, U_1 = factories having $\mathcal{Z}$ - values less than and equal to 200.

Population 2 [Sarndal, Swensson and Wretman (1992), p.660]. N = 50 cluster countries, $\mathcal{Y}$ = real estate values according to 1984 assessment, $\mathcal{X}$ = 1985 population, $\mathcal{Z}$ = revenues from the 1985 municipal taxation, U_1 = clusters with total number of seats in municipal council in 1982 less than and equal to 250.

Population 3 [Sarndal, Swensson and Wretman (1992), p.662]. N = 124 countries, $\mathcal{Y}$ = 1983 military expenditure, $\mathcal{X}$ = 1980 population, $\mathcal{Z}$ = 1982 gross national product, U_1 = countries with 1980 population less than and equal to 15 millions.

Table 2: Relative Precision of Different Estimators w.r.t. $\bar{y}^*$(in %)

Estimator	k	Population		
		1	2	3
		$n_1 = 10, \quad n_2 = 6$	$n_1 = 6, \quad n_2 = 4$	$n_1 = 17, \quad n_2 = 8$
$\bar{y}^*$	2	100.00	100.00	100.00
	3	100.00	100.00	100.00
	4	100.00	100.00	100.00
t_{hrg}	2	578.42	339.85	117.53
	3	445.50	276.14	115.08
	4	370.38	259.75	114.85
ℓ_{srg}	2	106.88	257.38	101.85
	3	113.73	312.14	101.35
	4	120.55	362.45	101.12
ℓ_{grg}	2	921.55	561.90	118.45
	3	963.85	731.24	116.12
	4	1004.56	828.19	115.93
$\hat{\hat{Y}}_{qrg}$	2	922.55	609.70	118.95
	3	965.91	808.41	119.19
	4	1007.72	944.28	119.48
$\hat{\hat{Y}}_{srg}$	2	106.98	258.59	102.85
	3	113.95	318.35	102.42
	4	120.89	356.29	102.35
$\hat{\hat{Y}}_{rg}$	2	929.31	895.71	121.43
	3	979.90	1067.06	122.37
	4	1029.37	1127.44	124.70

For simplicity, sample sizes in U_1 and U_2 are determined by proportional allocation *i.e.*, $n_j \propto N_j$. Relative precisions (RP) of different MVB estimators with respect to $\bar{y}^*$, for $k = 2, 3$ and 4 are displayed in Table 2. In evaluating the results of the table, it is seen that $\hat{\hat{Y}}_{rg}$ leads to achieving the highest gain in precision amongst all indicating that other comparable methods are inferior to the proposed method of estimation.

Findings of this empirical study, which of course has a limited scope, may not fit to other situations. However, it clearly shows that there are practical situations where the use of more auxiliary information at sample level may provide estimates with greater precision for the non response affected situations through the suggested method, even if this may require more computational efforts.

Acknowledgement

The authors are grateful to the referees for providing some useful comments on an earlier draft of the paper.

References

[1] Cochran, W. G. (1977). *Sampling Techniques* (Third Edition). John Wiley & Sons, New York.

[2] Hansen, M. H. and Hurwitz, W. N. (1946). The problem of non response in sample surveys. *Jour. Amer. Stat. Assoc.*, **41**, 517-529.

[3] Murthy, M. N. (1967). *Sampling Theory and Methods.* Statistical Publishing Society, Calcutta.

[4] Rao, P. S. R. S. (1986). Ratio estimation with sub sampling of non respondents. *Survey Methodology*, **12**, 217-230.

[5] Rao, P. S. R. S. (1990). Regression estimators with sub sampling of non respondents. *Data Quality Control: Theory and Pragmatics.* Gunar E. Liepins and V. R. R. Uppuluri (Eds), Marcel Dekker, New York, 191-208.

[6] Sarndal, C. E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling.* Springr – Verlag.

[7] Sinha, R. R. (2001). Some problems on the estimation of population parameters using multi-auxiliary characters in presence of non response. Ph. D. Thesis submitted to Banaras Hindu University, Varanasi, India.