Estimation of Mean Using Multi Auxiliary Information in Presence of Non Response

Sunil Kumar[1,a], Housila P. Singh[b]

[a]Department of Statistics, University of Jammu
[b]School of Studies in Statistics, Vikram University

Abstract

For estimating the mean of a finite population, three classes of estimators using multi-auxiliary information with unknown means using two phase sampling in presence of non-response have been proposed with their properties. Asymptotically optimum estimator(AOE) in each class has been identified along with their mean squared error formulae. An empirical study is also given.

Keywords: Study variate, multi-auxiliary variates, mean squared errors, two-phase sampling, non-response.

---

1. Introduction

In survey sampling, it is well established that the use of auxiliary information results in substantial gain in efficiency over the estimators which do not use such information. Out of many ratio, product and regression methods of estimation are good examples in this context. When the correlation between the study variate y and the auxiliary variate x is positive (high) the ratio method of estimation is quite effective. On the other hand if this correlation is negative (high) the product method of estimation envisaged by Robson (1957) and rediscovered by Murthy (1964), can be employed. In large-scale sample surveys, we often collect data on more than one auxiliary character and some of these may be correlated with y. Estimators using information of the known population mean of an auxiliary variable have generalized to the cases when such information is available for more than one auxiliary variables by several authors as Olkin (1958), Raj (1965), Rao and Mudholkar (1967), Singh (1967), Srivastava (1971) and Mohanty and Pattanaik (1982) and Agrawal and Panda (1993) *etc*. But in many situations of practical importance it has been observed that the population means of auxiliary variables are not known. So we use two-phase sampling scheme for estimating the population means of auxiliary variables. Srivastava (1981) suggested a class of estimators for estimating the population mean in two-phase sampling assuming that responses for all variables are available for each unit selected in the sample. But in practice, the problem of non-response often arises in sample surveys. In such situations for single variable survey, the problem of estimating the population mean using sub-sampling scheme has been first considered by Hansen and Hurwitz (1946). Further improvement in the estimation procedure for population mean in presence of non-response using auxiliary variable was considered by Cochran (1977, p.374), Rao (1986, 1987), Sarndal *et al.* (1992, p.583), Khare and Srivastava (1993, 1995, 1997), Okafor (1996), Tabasum and Khan (2004, 2006), Khare and Sinha (2004, 2007), Singh and Kumar (2008a, b; 2009a, b) and Singh *et al.* (2010).

---

[1] Corresponding author: Lecturer, Department of Statistics, University of Jammu, J & K, India,
  E-mail: sunilbhougal06@gmail.com

In this paper we have suggested three classes of estimators for estimating the population mean of the study variate using multi-auxiliary information with unknown population means using two-phase sampling in presence of non-response. The expressions for bias and mean squared errors of the suggested classes of estimators have been derived. The conditions for attaining minimum mean squared errors of the proposed classes have also been investigated. An empirical study is given in support of the present study.

2. Sampling Procedure and Notations

Consider a finite population $U = (U_1, U_2, \ldots, U_N)$ of N units. Let y denote the study character whose population mean $\bar{Y}$ is to be estimated using information on p auxiliary variates $x_1, x_2, \ldots, x_p$. Let $y_j, x_{1j}, x_{2j}, \ldots, x_{pj}$ denote the values of the variates $y, x_1, x_2, \ldots, x_p$ respectively, on the jth unit U_j of the population U, $j = 1, 2, \ldots, N$. When the population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ of the auxiliary variates $x_1, x_2, \ldots, x_p$ respectively are known several multivariate ratio and product estimators of the population mean $\bar{Y}$ have been formulated along with their properties for instance see Olkin (1958), Raj (1965), Rao and Mudholkar (1967), Singh (1967), Srivastava (1971), Mohanty and Pattanaik (1982) and Agarwal and Panda (1993). The population is supposed to be divided in N_1 responding and N_2 non-responding units such that $N_1 + N_2 = N$. However, in certain practical situations population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ of auxiliary variates $x_1, x_2, \ldots, x_p$ respectively are not known a priori in which case the technique of two-phase (or double) sampling can be useful. In two-phase sampling a first phase sample of size n' is drawn from the population by simple random sampling without replacement(SRSWOR) scheme on which only the auxiliary variates are measured in order to furnish the good estimates of $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$. A smaller second phase sample of size $n(< n')$ is selected from n' by simple random sampling without replacement(SRSWOR) and the study variate y is measured on it. Let $(\bar{x}'_1, \bar{x}'_2, \ldots, \bar{x}'_p)$ be the sample mean of the auxiliary variates $x_1, x_2, \ldots, x_p$ respectively based on first phase sample of size n'. Further let $\bar{y}$ and $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)$ be the sample means of the study variate y and auxiliary variates $x_1, x_2, \ldots, x_p$ obtained from the second phase sample of size n when there is no non-response (i.e. complete response) in the second phase sample. In such situations the formulation of two-phase (or double) sampling multivariate ratio and product estimators can be done easily just replacing $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ by $(\bar{x}'_1, \bar{x}'_2, \ldots, \bar{x}'_p)$, see Srivastava (1981). If, however, there is non-response in the second phase sample, take a sub-sample of the non-respondents and re-contact them.

We assume that at the first phase, all the n' units supply information on the auxiliary variates $x_1, x_2, \ldots, x_p$. From the second phase sample of n units, let n_1 units supply information on the study variate y and n_2 units refuse to respond. From the n_2 non-respondents, using Hansen and Hurwitz (1946) procedure we again select a sub sample of size $m = n_2/k$, $(k > 1)$ units using SRSWOR assuming all the m units respond. Here we have $(n_1 + m)$ responding units on the study variate y and consequently the estimator for population mean $\bar{Y}$ using sub sampling scheme envisaged by Hansen and Hurwitz (1946) is defined by

$$\bar{y}^* = \left(\frac{n_1}{n}\right)\bar{y}_{(1)} + \left(\frac{n_2}{n}\right)\bar{y}_{(2)}, \tag{2.1}$$

where $\bar{y}_{(1)}$ and $\bar{y}_{(2)}$ denote the sample means of the study variate y based on n_1 and m units respectively. It is well known that the estimator $\bar{y}^*$ is unbiased estimator of the population mean $\bar{Y}$ and has the variance

$$\mathrm{Var}(\bar{y}^*) = \left(\frac{1-f}{n}\right)S_0^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2, \tag{2.2}$$

where $f = n/N$, $W_i = N_i/N$, $(i = 1, 2)$, S_0^2 and $S_{0(2)}^2$ are the population mean square of the variate y for the entire population and for non-responding group of the population. Similarly for estimating the population mean $\bar{X}_i$ of the auxiliary variate x_i $(i = 1, 2, \ldots, p)$, the unbiased estimator $\bar{x}_i^*$ is given by

$$\bar{x}_i^* = \left(\frac{n_1}{n}\right)\bar{x}_{i(1)} + \left(\frac{n_2}{n}\right)\bar{x}_{i(2)}, \quad i = 1, 2, \ldots, p, \tag{2.3}$$

where $\bar{x}_{i(1)}$ and $\bar{x}_{i(2)}$ are the sample means of the auxiliary variate x_i $(i = 1, 2, \ldots, p)$ based on n_1 and m units respectively. The variance of $\bar{x}_i^*$ is given by

$$\text{Var}(\bar{x}^*) = \left(\frac{1-f}{n}\right)S_i^2 + \frac{W_2(k-1)}{n}S_{i(2)}^2, \quad i = 1, 2, \ldots, p, \tag{2.4}$$

where S_i^2 and $S_{i(2)}^2$ are the population mean square of x_i $(i = 1, 2, \ldots, p)$ for the entire population and non responding group of the population.

Now we define

$$C_0 = \frac{S_0}{\bar{Y}}, \quad C_i = \frac{S_i}{\bar{X}_i}, \quad i = 1, 2, \ldots, p, \quad C_{0(2)} = \frac{S_{0(2)}}{\bar{Y}}, \quad C_{i(2)} = \frac{S_{i(2)}}{\bar{X}_i}, \quad \rho_{0i} = \frac{S_{0i}}{S_0 S_i},$$

$$S_{0i} = \frac{1}{N-1}\sum_{j=1}^{N}\left(y_j - \bar{Y}\right)\left(x_{ij} - \bar{X}_i\right), \quad \rho_{0i(2)} = \frac{S_{0i(2)}}{S_{0(2)}S_{i(2)}}, \quad \bar{Y}_{(2)} = \frac{1}{N_2}\sum_{j=1}^{N_2}y_{ij},$$

$$S_{0i(2)} = \frac{1}{N_2-1}\sum_{j=1}^{N_2}\left(y_j - \bar{Y}_{(2)}\right)\left(x_{ij} - \bar{X}_{i(2)}\right), \quad \bar{X}_{i(2)} = \frac{1}{N_2}\sum_{j=1}^{N_2}x_{ij}, \quad \beta_{0i} = \frac{S_{0i}}{S_i^2}, \quad i = 1, 2, \ldots, p,$$

$$\beta_{0i(2)} = \frac{S_{0i(2)}}{S_{i(2)}^2}, \quad i = 1, 2, \ldots, p, \quad \rho_{il} = \frac{S_{il}}{S_i S_l}, \quad \rho_{il(2)} = \frac{S_{il(2)}}{S_{i(2)}S_{l(2)}}, \quad (i \neq l = 1, 2, \ldots, p), \quad R = \frac{\bar{Y}}{\bar{X}}.$$

Let

$$u_i = \frac{\bar{x}_i^*}{\bar{x}_i'}, \quad i = 1, 2, \ldots, p;$$

$$= \frac{\bar{x}_{i-p}}{\bar{x}_{i-p}'}, \quad i = p + 1, p + 2, \ldots, 2p.$$

Let u denote the column vector of $2p$ elements $u_1, u_2, \ldots, u_{2p}$. Super fix T over a column vector denotes the corresponding row vector.

$$\varepsilon_0 = \frac{\bar{y}^* - \bar{Y}}{\bar{Y}}, \quad \eta_i^* = \frac{\bar{x}_i^* - \bar{X}_i}{\bar{X}_i}, \quad \eta_i' = \frac{\bar{x}_i' - \bar{X}_i}{\bar{X}_i}, \quad \eta_{i-p} = \frac{\bar{x}_{i-p} - \bar{X}_i}{\bar{X}_i}, \quad \eta_{i-p}' = \frac{\bar{x}_{i-p}' - \bar{X}_i}{\bar{X}_i},$$

$$\delta_i = u_i - 1 = \left(\eta_{i-p} - \eta_{i-p}'\right) + \left(\eta_{i-p}'^2 - \eta_{i-p}\eta_{i-p}'\right) + \cdots, \quad i = p + 1, p + 2, \ldots, 2p,$$

$$E(\varepsilon_0) = 0, \quad E(\delta_i) = 0 \quad \forall \ i = 1, 2, \ldots, 2p,$$

$$E(\varepsilon_0 \delta_i) = \left\{\left(\frac{1}{n} - \frac{1}{n'}\right)q_i + \frac{W_2(k-1)}{n}q_{i(2)}\right\} = -a_i, \quad i = 1, 2, \ldots, p,$$

$$= \left(\frac{1}{n} - \frac{1}{n'}\right)q_i = g_i, \quad i = p + 1, p + 2, \ldots, 2p,$$

$$q_i = \rho_{0i}C_0C_i, \quad q_{i(2)} = \rho_{0i(2)}C_{0(2)}C_{i(2)}, \quad i = 1, 2, \ldots, p,$$

$$E\left(\delta_i\delta_l\right) = \left\{\left(\frac{1}{n} - \frac{1}{n'}\right)a_{il} + \frac{W_2(k-1)}{n}a_{il(2)}\right\} = e_{il};$$

$$= \left(\frac{1}{n} - \frac{1}{n'}\right)a_{il}; \quad (i = 1, 2, \ldots, 2p; \ l = p+1, p+2, \ldots, 2p);$$

$a_{il} = \rho_{il}C_iC_l$; $a_{il(2)} = \rho_{il(2)}C_{i(2)}C_{l(2)}$, $(i, l) = 1, 2, \ldots, p$; $\{\rho_{0i}, \rho_{il}, (i, l) = 1, 2, \ldots, p\}$ and $\{\rho_{0i(2)}, \rho_{il(2)}, (i, l) = 1, 2, \ldots, p\}$ are the correlation coefficients between (y, x_i) and (x_i, x_l) respectively for the entire population and for the non-responding group of the population. Putting the above results in matrix notations, we have

$$E\left(\varepsilon_0\delta^T\right) = b^T, \quad E\left(\delta\delta^T\right) = D = \begin{bmatrix} E & F \\ F^T & T \end{bmatrix}, \quad E = F + F_{(2)}, \quad F = \left(\frac{1}{n} - \frac{1}{n'}\right)a = (f_{il})_{p\times p};$$

$$F = \frac{W_2(k-1)}{n}a_{(2)} = (f_{il(2)})_{p\times p}; \quad a = (a_{il})_{p\times p}; \quad a_{(2)} = (a_{il(2)})_{p\times p}; \quad f_{il} = \left(\frac{1}{n} - \frac{1}{n'}\right)a_{il};$$

$$f_{il(2)} = \frac{W_2(k-1)}{n}a_{il(2)}; \quad b^T = \left(Q^T : g^T\right) = \left(Q_1, Q_2, \ldots, Q_p, g_1, g_2, \ldots, g_p\right),$$

$$Q^T = \left\{\left(\frac{1}{n} - \frac{1}{n'}\right)q^T + \frac{W_2(k-1)}{n}q_{(2)}^T\right\}, \quad Q_i = \left\{\left(\frac{1}{n} - \frac{1}{n'}\right)q_i + \frac{W_2(k-1)}{n}q_{i(2)}\right\}, \quad i = 1, 2, \ldots, p,$$

$$q^T = \left(q_1, q_2, \ldots, q_p\right), \quad q_{(2)}^T = \left(q_{1(2)}, q_{2(2)}, \ldots, q_{p(2)}\right), \quad g^T = \left(\frac{1}{n} - \frac{1}{n'}\right)q^T, \quad g_{(2)}^T = \frac{W_2(k-1)}{n}q_{(2)}^T,$$

$$Q^T = \left(g^T + g_{(2)}^T\right), \quad E\left(\varepsilon_0\varphi^T\right) = Q^T, \quad E\left(\varphi\varphi^T\right) = E_{p\times p}, \quad Q = \left(\frac{1}{n} - \frac{1}{n'}\right)q + \frac{W_2(k-1)}{n}q_{(2)},$$

$$e_{11} = \left(\frac{1}{n} - \frac{1}{n'}\right)C_1^2 + \frac{W_2(k-1)}{n}C_{1(2)}^2, \quad e_{22} = \left(\frac{1}{n} - \frac{1}{n'}\right)C_2^2 + \frac{W_2(k-1)}{n}C_{2(2)}^2$$

$$e_{12} = \left(\frac{1}{n} - \frac{1}{n'}\right)a_{12} + \frac{W_2(k-1)}{n}a_{12(2)}, \quad E^* = \left(\frac{1-f}{n}\right)a + \frac{W_2(k-1)}{n}a_{(2)} = \left(e_{il}^*\right)_{p\times p},$$

$$F^* = \left(f_{il}^*\right)_{p\times p}, \quad f_{il}^* = \left(\frac{1-f}{n}\right)a_{il}, \quad e_{il}^* = f_{il}^* + f_{il(2)}^*, \quad (i, l) = 1, 2, \ldots, p,$$

$$Q^* = \left(\frac{1-f}{n}\right)q + \frac{W_2(k-1)}{n}q_{(2)}, \quad Q_i^* = \left(\frac{1-f}{n}\right)q_i + \frac{W_2(k-1)}{n}q_{i(2)}, \quad (i = 1, 2, \ldots, p),$$

$$e_{11}^* = \left(\frac{1-f}{n}\right)C_1^2 + \frac{W_2(k-1)}{n}C_{1(2)}^2, \quad e_{22}^* = \left(\frac{1-f}{n}\right)C_2^2 + \frac{W_2(k-1)}{n}C_{2(2)}^2,$$

$$e_{12}^* = \left(\frac{1-f}{n}\right)\rho_{12}C_1C_2 + \frac{W_2(k-1)}{n}\rho_{12(2)}C_{1(2)}C_{2(2)}, \quad D^* = \begin{bmatrix} E^* & F^* \\ F^{*T} & F^* \end{bmatrix},$$

$$b^{*T} = \left(Q^{*T} : g^{*T}\right) = \left(Q_1^*, Q_2^*, \ldots, Q_p^*, g_1^*, g_2^*, \ldots, g_p^*\right), \quad g^{*T} = \left(\frac{1-f}{n}\right)q^T,$$

$$Q^{*T} = \left(\frac{1-f}{n}\right)q^T + \frac{W_2(k-1)}{n}q_{(2)}^T, \quad d = \rho_{01}\left(\frac{C_0}{C_1}\right), \quad d_{(2)} = \rho_{01(2)}\left(\frac{C_{0(2)}}{C_{1(2)}}\right),$$

$$d_{01(2)} = \frac{\left(\rho_{01(2)} - \rho_{02(2)}\rho_{12(2)}\right)C_{0(2)}}{\left(1 - \rho_{12(2)}^2\right)C_{1(2)}}, \quad d_{02(2)} = \frac{\left(\rho_{02(2)} - \rho_{01(2)}\rho_{12(2)}\right)C_{0(2)}}{\left(1 - \rho_{12(2)}^2\right)C_{2(2)}}$$

$$d_{01} = \frac{(\rho_{01} - \rho_{02}\rho_{12}) C_0}{(1 - \rho_{12}^2) C_1} \quad \text{and} \quad d_{02} = \frac{(\rho_{02} - \rho_{01}\rho_{12}) C_0}{(1 - \rho_{12}^2) C_2}.$$

The matrix D is assumed to be positive definite. The matrices $E = (e_{il})_{p \times p}$, $F = (f_{il})_{p \times p}$ and $F_{(2)} = (f_{il(2)})_{p \times p}$ are $p \times p$ matrices.

In the following sections, utilizing the information an $p(> 1)$ auxiliary variates. We will define different estimators for $\bar{Y}$ in different situations and study their properties.

3. The Suggested Classes of Estimators

3.1. Strategy-I

Population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ of the auxiliary variates $x_1, x_2, \ldots, x_p$ are not known, incomplete information on the study variate y and auxiliary variate x.

In this situation we use $(n_1 + m)$ responding units for y and x from the sample of size n and $\bar{x}'_i$ to estimate $\bar{X}_i$, $(i = 1, 2, \ldots, p)$. Let $v^T = (v_1, v_2, \ldots, v_p)$, $v_i = \bar{x}_i^* / \bar{x}'_i$, $i = 1, 2, \ldots, p$ and e^T denotes the row vector of p unit elements. Whatever be the sample selected, let v^T assume values in a bounded closed convex subset J of the $(p + 1)$ dimensional real space containing the point e^T. Let $A(\bar{y}^*, v^T)$ be a function of $(\bar{y}^*, v^T)$ such that it satisfies the following conditions:

(i) In J, the function $A(\bar{y}^*, v^T)$ is continuous and bounded.

(ii) The first and second order partial derivatives of $A(\bar{y}^*, v^T)$ exist and are continuous and bounded in J.

We propose a class of estimators for the population mean $\bar{Y}$ as

$$\hat{M}_1 = A\left(\bar{y}^*, v^T\right), \tag{3.1}$$

where $A(\bar{y}^*, v^T)$ is the function of $(\bar{y}^*, v^T)$ such that

$$A\left(\bar{Y}, e^T\right) = \bar{Y}, \quad \text{for all } \bar{Y}. \tag{3.2}$$

Since there are only a finite number of samples, the expectations and mean squared error of the class of estimators $\hat{M}_1$ exist under the conditions 1 and 2.

Expanding $A(\bar{y}^*, v^T)$ about the point $v^T = e^T$ by a second order Taylor's series, we obtain

$$\hat{M}_1 = A\left(\bar{Y}, e^T\right) + \left(\bar{y}^* - \bar{Y}\right) \frac{\partial A(\cdot)}{\partial \bar{y}^*}\bigg|_{(\bar{Y}, e^T)} + (v - e)^T A^{(1)}\left(\bar{Y}, e^T\right) + \frac{1}{2} \left\{ \left(\bar{y}^* - \bar{Y}\right)^2 \frac{\partial^2 A(\cdot)}{\partial \bar{y}^{*2}}\bigg|_{(\bar{y}^{**}, v^{*T})} \right.$$

$$\left. + 2 \left(\bar{y}^* - \bar{Y}\right)(v - e)^T \frac{\partial A^{(1)}(\cdot)}{\partial \bar{y}^*}\bigg|_{(\bar{y}^{**}, v^{*T})} + (v - e)^T A^{(2)}\left(\bar{y}^{**}, v^{*T}\right)(v - e) \right\}, \tag{3.3}$$

where $\bar{y}^{**} = \bar{Y} + \theta(\bar{y}^* - \bar{Y})$, $v^{*T} = e^T + \theta(v - e)^T$, $0 < \theta < 1$ and $A^{(1)}(\bar{Y}, e^T)$ denotes the p elements vector of first partial derivatives of $A(\bar{y}^*, v^T)$ with respect to v about the point $v^T = e^T$ and $A^{(2)}(\bar{y}^{**}, v^{*T})$ denotes $2p \times 2p$ matrix of the second partial derivatives of $A(\bar{y}^*, v^T)$ with respect to v about the point $v^T = v^{*T}$. Taking expectation in (3.3) and noting that

$$E(\varepsilon_0) = E\left(\eta_i^*\right) = E\left(\eta_i'\right) = 0, \quad i = 1, 2, \ldots, p$$

and that the expectations of the second degree terms are of order n^{-1}, we obtain

$$E\left(\hat{M}_1\right) = \bar{Y} + o\left(n^{-1}\right).$$

Thus the bias of the estimator $\hat{M}_1$ is of the order n^{-1} and hence its contribution to the mean square error will be of the order n^{-2}.

Noting that

$$A\left(\bar{Y}, e^T\right) = \bar{Y}, \quad \text{for all } \bar{Y} \qquad \Rightarrow \qquad \left.\frac{\partial A(\cdot)}{\partial \bar{y}^*}\right|_{(\bar{Y},e^T)} = 1.$$

Thus from (3.3) we have

$$\hat{M}_1 \approx \bar{Y} + \left(\bar{y}^* - \bar{Y}\right) + (v - e)^T A^{(1)}\left(\bar{Y}, e^T\right) \quad \text{or} \quad \left(\hat{M}_1 - \bar{Y}\right) = \bar{Y}\varepsilon_0 + \phi^T A^{(1)}\left(\bar{Y}, e^T\right), \tag{3.4}$$

where $\phi^T = (v - e)^T$.

Squaring both sides of (3.4) we have

$$\left(\hat{M}_1 - \bar{Y}\right)^2 = \left[\bar{Y}^2 \varepsilon_0^2 + 2\bar{Y}\varepsilon_0 \phi^T A^{(1)}\left(\bar{Y}, e^T\right) + \left\{A^{(1)}\left(\bar{Y}, e^T\right)\right\}^T \phi\phi^T A^{(1)}\left(\bar{Y}, e^T\right)\right]. \tag{3.5}$$

We note that

$$E\left(\varepsilon_0 \phi^T\right) = Q^T \quad \text{and} \quad E\left(\phi\phi^T\right) = E_{p\times p}. \tag{3.6}$$

Taking expectations of both sides in (3.5) and using (3.6) we get the MSE of $\hat{M}_1$ to the first degree of approximation as

$$\text{MSE}\left(\hat{M}_1\right) = \left[\left(\frac{1-f}{n}\right)S_0^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2 + 2\bar{Y}Q^T A^{(1)}\left(\bar{Y}, e^T\right) + \left\{A^{(1)}\left(\bar{Y}, e^T\right)\right\}^T E\left\{A^{(1)}\left(\bar{Y}, e^T\right)\right\}\right]. \tag{3.7}$$

The MSE of $\hat{M}_1$ is minimized for

$$A^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y}E^{-1}Q. \tag{3.8}$$

Thus, the resulting minimum MSE of $\hat{M}_1$ is given by

$$\min \text{MSE}\left(\hat{M}_1\right) = \bar{Y}^2 \left\{\left(\frac{1-f}{n}\right)C_0^2 + \frac{W_2(k-1)}{n}C_{0(2)}^2 - Q^T E^{-1} Q\right\}. \tag{3.9}$$

Now we state the following theorem.

Theorem 1. *To the first degree of approximation*

$$MSE\left(\hat{M}_1\right) \geq \bar{Y}^2 \left\{\left(\frac{1-f}{n}\right)C_0^2 + \frac{W_2(k-1)}{n}C_{0(2)}^2 - Q^T E^{-1} Q\right\}$$

with equality holding if

$$A^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y}E^{-1}Q.$$

The class of estimators (3.1) is very vast, if the parameters in the function $A(\bar{y}^*, v^T)$ are so selected that they satisfy (3.8), the resulting estimator will have MSE given by (3.9). A few examples are:

$$\hat{M}_{1(1)} = \bar{y}^* + \psi^T (v - e), \quad \hat{M}_{1(2)} = \bar{y}^* \exp\left(\psi^T \log v\right), \quad \hat{M}_{1(3)} = \bar{y}^* \left\{1 + \psi^T (v - e)\right\}$$

$$\hat{M}_{1(4)} = \frac{\bar{y}^*}{1 - \psi^T (v - e)}, \quad \hat{M}_{1(5)} = \bar{y}^* \prod_{i=1}^{p} v_i \psi_i, \quad \hat{M}_{1(6)} = \frac{\bar{y}^{*2}}{\bar{y}^* - \psi^T (v - e)},$$

where $\psi^T = (\psi_1, \psi_2, \ldots, \psi_p)$ is vector of p constants. The optimum values of these constants are obtained from the conditions (3.8). Since (3.8) contains p equations, we have taken exactly p unknown constants in defining above estimators of the class.

Remark 1. For the case of a single auxiliary variable x_1, the MSE of $\hat{M}_1$ defined at (3.7) is minimized for

$$A^{(1)}\left(\bar{Y}, 1\right) = -\bar{Y}\left(\frac{Q_1}{e_{11}}\right)$$

and the minimum MSE of $\hat{M}_1$ is given by

$$\min \text{MSE}\left(\hat{M}_1\right)_I = \bar{Y}^2 \left\{\left(\frac{1-f}{n}\right) C_0^2 + \frac{W_2(k-1)}{n} C_{0(2)}^2 - \frac{Q_1^2}{e_{11}}\right\}. \tag{3.10}$$

which equals to the Variance of the optimum estimator

$$\hat{M}_{1\beta} = \bar{y}^* + \beta^* \left(\bar{x}_1' - \bar{x}_1^*\right) \tag{3.11}$$

in the class of estimators

$$\hat{M}_{1d} = \bar{y}^* + d \left(\bar{x}_1' - \bar{x}_1^*\right)$$

where d is suitably chosen constant and

$$\beta^* = \frac{Q_1}{e_{11}}.$$

Remark 2. In case of two auxiliary variables, the MSE of $\hat{M}_1$ defined at (3.7) is minimized for

$$\left[\begin{array}{c} A_1^{(1)}\left(\bar{Y}, 1, 1\right) \\ A_2^{(1)}\left(\bar{Y}, 1, 1\right) \end{array}\right] = \bar{Y}\left[\begin{array}{c} \dfrac{e_{12}Q_2 - e_{22}Q_1}{e_{11}e_{22} - e_{12}^2} \\ \dfrac{e_{12}Q_1 - e_{11}Q_2}{e_{11}e_{22} - e_{12}^2} \end{array}\right]. \tag{3.12}$$

Thus the resulting minimum MSE of $\hat{M}_1$ (with two auxiliary variates) is given by

$$\min \text{MSE}\left(\hat{M}_1\right)_{II} = \bar{Y}^2 \left\{\left(\frac{1-f}{n}\right) C_0^2 + \frac{W_2(k-1)}{n} C_{0(2)}^2 - \frac{Q_1^2 e_{22} + Q_2^2 e_{11} - 2Q_1 Q_2 e_{12}}{e_{11}e_{22} - e_{12}^2}\right\} \tag{3.13}$$

which equals to the variance of the optimum estimator

$$\hat{M}_{10}^* = \bar{y}^* + \beta_{01}^* \left(\bar{x}_1' - \bar{x}_1^*\right) + \beta_{02}^* \left(\bar{x}_2' - \bar{x}_2^*\right)$$

is the class of estimators

$$\hat{M}^*_{1d1} = \bar{y}^* + d_1 \left(\bar{x}'_1 - \bar{x}^*_1 \right) + d_2 \left(\bar{x}'_2 - \bar{x}^*_2 \right),$$

where $\beta^*_{01} = (e_{12}Q_2 - e_{22}Q_1)/(e_{11}e_{22} - e^2_{12})$ and $\beta^*_{02} = (e_{12}Q_1 - e_{11}Q_2)/(e_{11}e_{22} - e^2_{12})$ are the optimum values of d_1 and d_2 respectively.

From (3.10) and (3.13) we have

$$\min \text{MSE} \left(\hat{M}_1 \right)_I - \min \text{MSE} \left(\hat{M}_1 \right)_{II} = \frac{\bar{Y}^2 \left(Q_1 e_{12} + Q_2 e_{11} \right)^2}{e_{11} \left(e_{11}e_{22} - e^2_{12} \right)} \geq 0$$

which establishes that the proposed estimator $\hat{M}_1$ with two auxiliary variables is more efficient than that with one auxiliary variable.

Remark 3. In case $n' = N$ i.e. the population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ of the auxiliary variates $x_1, x_2, \ldots, x_p$ respectively are known, the class of estimators $\hat{M}_1$ defined by (3.1) reduces to:

$$\hat{M}_{11} = A^* \left(\bar{y}^*, v^{*T} \right), \tag{3.14}$$

where $v^{*T} = (v^*_1, v^*_2, \ldots, v^*_p)$, $v^*_i = \bar{x}^*_i/\bar{X}_i$, $i = 1, 2, \ldots, p$ and $A^*(\bar{y}^*, v^{*T})$ is the function of $(\bar{y}^*, v^{*T})$ such that $A^*(\bar{Y}, e^T) = \bar{Y}$ for all $\bar{Y}$.

Putting $n' = N$ in (3.8) and (3.9) we get the optimum values of the derivatives and the minimum MSE of $\hat{M}_{11}$ are respectively given by

$$A^{*(1)} \left(\bar{Y}, e^T \right) = -\bar{Y} E^{*-1} Q^* \tag{3.15}$$

and

$$\min \text{MSE} \left(\hat{M}_{11} \right) = \bar{Y}^2 \left\{ \left(\frac{1-f}{n} \right) C^2_0 + \frac{W_2(k-1)}{n} C^2_{0(2)} - Q^{*T} E^{*-1} Q^* \right\}. \tag{3.16}$$

Remark 4. For the case of a single auxiliary variable x_1 with known population mean $\bar{X}_1$, the expressions (3.15) and (3.16) respectively reduce to:

$$A^{(1)} \left(\bar{Y}, 1 \right) = -\bar{Y} \frac{\left(\frac{1-f}{n} \right) \rho_{01} C_0 C_1 + \frac{W_2(k-1)}{n} \rho_{01(2)} C_{0(2)} C_{1(2)}}{\left(\frac{1-f}{n} \right) C^2_1 + \frac{W_2(k-1)}{n} C^2_{1(2)}} \tag{3.17}$$

and

$$\min \text{MSE} \left(\hat{M}_{11} \right)_I = \bar{Y}^2 \left\{ \left(\frac{1-f}{n} \right) C^2_0 + \frac{W_2(k-1)}{n} C^2_{0(2)} - \frac{Q^{*2}_1}{e^*_{11}} \right\}. \tag{3.18}$$

It can be shown that the variance of the optimum estimator

$$\hat{M}^*_0 = \bar{y}^* + \beta^*_0 \left(\bar{X}_1 - \bar{x}^*_1 \right)$$

in the class of estimators

$$\hat{M}_{1d}^* = \bar{y}^* + d^* \left(\bar{X}_1 - \bar{x}_1^* \right)$$

is same as given by (3.18) *i.e.*, $\text{Var}(\hat{M}_0^*) = \min \text{MSE}(\hat{M}_{11})$, $\beta_0^* = (Q_1^*/e_{11}^*)$ and d^* is a suitably chosen constant.

Remark 5. In case of two auxiliary variables with known population means $\bar{X}_1$ and $\bar{X}_2$, the expressions (3.12) and (3.13) respectively reduce to:

$$\begin{bmatrix} A_1^{*(1)} \left(\bar{Y}, 1, 1 \right) \\ A_2^{*(2)} \left(\bar{Y}, 1, 1 \right) \end{bmatrix} = \bar{Y} \begin{bmatrix} \dfrac{e_{12}^* Q_2^* - e_{22}^* Q_1^*}{e_{11}^* e_{22}^* - e_{12}^{*2}} \\ \dfrac{e_{12}^* Q_1^* - e_{11}^* Q_2^*}{e_{11}^* e_{22}^* - e_{12}^{*2}} \end{bmatrix}, \tag{3.19}$$

$$\min \text{MSE} \left(\hat{M}_{11} \right)_{II} = \bar{Y}^2 \left\{ \left(\frac{1-f}{n} \right) C_0^2 + \frac{W_2(k-1)}{n} C_{0(2)}^2 - \frac{Q_1^{*2} e_{22}^* + Q_2^{*2} e_{11}^* - 2Q_1^* Q_2^* e_{12}^*}{e_{11}^* e_{22}^* - e_{12}^{*2}} \right\}, \tag{3.20}$$

where $A_1^{*(1)}(\bar{Y}, 1, 1) = \partial A^*(\cdot)/\partial v_1^*|_{(\bar{Y},1,1)}$, $A_2^{*(1)}(\bar{Y}, 1, 1) = \partial A^*(\cdot)/\partial v_2^*|_{(\bar{Y},1,1)}$.

It is to be mentioned that the variance of the optimum estimator

$$\hat{M}_{20}^* = \bar{y}^* + \beta_{01}^{**} \left(\bar{X}_1 - \bar{x}_1^* \right) + \beta_{02}^{**} \left(\bar{X}_2 - \bar{x}_2^* \right)$$

in the class of estimators

$$\hat{M}_{2d} = \bar{y}^* + d_1^* \left(\bar{X}_1 - \bar{x}_1^* \right) + d_2^* \left(\bar{X}_2 - \bar{x}_2^* \right)$$

is same as given by (3.20) and $\beta_{01}^{**} = (e_{12}^* Q_2^* - e_{22}^* Q_1^*)/(e_{11}^* e_{22}^* - e_{12}^{*2})$ and $\beta_{02}^{**} = (e_{12}^* Q_1^* - e_{11}^* Q_2^*)/(e_{11}^* e_{22}^* - e_{12}^{*2})$.

From (3.18) and (3.20) we have

$$\min \text{MSE} \left(\hat{M}_{11} \right)_I - \min \text{MSE} \left(\hat{M}_{11} \right)_{II} = \frac{\bar{Y}^2 \left(Q_1^* e_{12}^* + Q_2^* e_{11}^* \right)^2}{e_{11}^* \left(e_{11}^* e_{22}^* - e_{12}^{*2} \right)} \geq 0 \tag{3.21}$$

which shows that the proposed estimator $\hat{M}_{11}$ with two auxiliary variables is more efficient than $\hat{M}_{11}$ with one auxiliary variable.

3.2. Strategy-II

Population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ of the auxiliary variates $x_1, x_2, \ldots, x_p$ are unknown, incomplete information on the study variate y and complete information on the auxiliary variate x.

We consider the situation where information on $(n_1 + m)$ responding units on the study variate y and complete information on the auxiliary variate from the sample of size n are available. Also the population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ are unknown. Let $w^T = (w_1, w_2, \ldots, w_p)$, $w_i = \bar{x}_i/\bar{x}_i'$, $i = 1, 2, \ldots, p$ and e^T denote the row vector of p unit elements. Whatever be the sample chosen, let w^T assume values in a bounded closed convex subset L of the $(p + 1)$ dimensional real space containing the point e^T. We define a class of estimators for $\bar{Y}$ as

$$\hat{M}_2 = B \left(\bar{y}^*, w^T \right), \tag{3.22}$$

where $B(\bar{y}^*, w^T)$ is the function of $(\bar{y}^*, w^T)$ such that

$$B\left(\bar{Y}, e^T\right) = \bar{Y}, \quad \text{for all } \bar{Y}$$

$$B_1\left(\bar{Y}, e^T\right) = \left.\frac{\partial B(\cdot)}{\partial \bar{y}^*}\right|_{(\bar{Y}, e^T)} = 1 \tag{3.23}$$

and it also satisfies certain conditions similar to those given for $\hat{M}_1$ at (3.1).

Since there are only a finite number of samples, the expectations and mean squared error of the class of estimator $\hat{M}_2$ exist. Expanding $B(\bar{y}^*, w^T)$ about the point $(\bar{Y}, e^T)$ in a second order Taylor's series, we have that $E(\hat{M}_2) = \bar{Y} + o(n^{-1})$ and so the bias of $\hat{M}_2$ is of the order n^{-1}. The mean squared error of $\hat{M}_2$ up to the terms of order n^{-1} is

$$\text{MSE}\left(\hat{M}_2\right) = \left[\left(\frac{1-f}{n}\right)S_0^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2 + 2\bar{Y}g^T B^{(1)}\left(\bar{Y}, e^T\right) + \left\{B^{(1)}\left(\bar{Y}, e^T\right)\right\}^T F\left\{B^{(1)}\left(\bar{Y}, e^T\right)\right\}\right] \tag{3.24}$$

$$= \left[\left(\frac{1-f}{n}\right)S_0^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left\{2\bar{Y}q^T B^{(1)}\left(\bar{Y}, e^T\right) + \left\{B^{(1)}\left(\bar{Y}, e^T\right)\right\}^T a\left\{B^{(1)}\left(\bar{Y}, e^T\right)\right\}\right\}\right],$$

where $B^{(1)}(\bar{Y}, e^T)$ denotes the p elements column vector of first partial derivatives of $B(\bar{y}^*, w^T)$ with respect to w about the point $w^T = e^T$. The MSE of $\hat{M}_2$ is minimized for

$$B^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y}a^{-1}q \tag{3.25}$$

and the resulting minimum MSE of $\hat{M}_2$ is given by

$$\min\text{MSE}\left(\hat{M}_2\right) = \left\{\left(\frac{1-f}{n}\right)S_0^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2 - \bar{Y}^2\left(\frac{1}{n} - \frac{1}{n'}\right)q^T a^{-1}q\right\}$$

$$= \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - R_{0.1,2,\ldots,p}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right)S_0^2 R_{0.1,2,\ldots,p}^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2\right\}, \tag{3.26}$$

where $R_{0.1,2,\ldots,p}^2 = (q^T a^{-1}q)/C_0^2$ denotes the square of the multiple correlation coefficient of y on $x_1, x_2, \ldots, x_p$. Thus we state the following theorem

Theorem 2. *To the first degree of approximation,*

$$MSE\left(\hat{M}_2\right) \geq \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - R_{0.123\cdots p}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right)S_0^2 R_{0.123\cdots p}^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2\right\} \tag{3.27}$$

with equality holding if

$$B^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y}a^{-1}q.$$

Remark 6. The class of estimators $\hat{M}_2$ given by (3.22) is very large, if the parameters in the function $B(\bar{y}^*, w^T)$ are so chosen that they satisfy (3.25), the resulting estimator will have MSE given by (3.26). A few examples are

$$\hat{M}_{2(1)} = \bar{y}^* + \mu^T(w - e), \qquad \hat{M}_{2(2)} = \frac{\bar{y}^{*2}}{\bar{y}^* - \mu^T(w - e)},$$

$$\hat{M}_{2(3)} = \bar{y}^* \exp\left(\mu^T \log w\right), \qquad \hat{M}_{2(4)} = \bar{y}^*\left\{1 + \mu^T(w - e)\right\}, \quad etc$$

where $\mu^T = (\mu_1, \mu_2, \ldots, \mu_p)$ is vector of p scalars. The optimum values of these scalars are determined from the condition (3.25). Since (3.25) involves p equations, we have taken exactly p unknown constants in defining above estimators of the class.

Remark 7. For the case of a single auxiliary character x_1, the MSE of $\hat{M}_2$ is minimized for

$$B^{(1)}\left(\bar{Y}, 1\right) = -\bar{Y} K_{01} = -\bar{Y}\left(\frac{\beta}{R}\right), \tag{3.28}$$

where $K_{01} = \rho_{01}(C_0/C_1), \beta = (S_{01}/S_1^2)$ and $R = (\bar{Y}/\bar{X})$.

Thus the minimum MSE of $\hat{M}_2$ is given by

$$\min \mathrm{MSE}\left(\hat{M}_2\right) = \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - \rho_{01}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right)\rho_{01}^2 S_0^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2\right\}, \tag{3.29}$$

which equals to the variance of the optimum estimator

$$\hat{M}_{20}^{(1)} = \bar{y}^* + \beta\left(\bar{x}_1' - \bar{x}_1\right)$$

is the class of estimators

$$\hat{M}_{2d}^{(1)} = \bar{y}^* + d\left(\bar{x}_1' - \bar{x}_1\right),$$

where d is a suitably chosen scalars.

Remark 8. For the case of two auxiliary variates x_1 and x_2, the $\hat{M}_2$ of given by (3.24) is minimized for

$$\left[\begin{array}{c} B_1^{(1)}\left(\bar{Y}, 1, 1\right) \\ B_2^{(1)}\left(\bar{Y}, 1, 1\right) \end{array}\right] = -\bar{Y}\left[\begin{array}{c} d_{01} \\ d_{02} \end{array}\right], \tag{3.30}$$

where

$$B_1^{(1)}\left(\bar{Y}, 1, 1\right) = \left.\frac{\partial B(\cdot)}{\partial w_1}\right|_{(\bar{Y}, 1, 1)}, \qquad B_2^{(1)}\left(\bar{Y}, 1, 1\right) = \left.\frac{\partial B(\cdot)}{\partial w_2}\right|_{(\bar{Y}, 1, 1)}.$$

Thus the resulting minimum MSE of $\hat{M}_2$ is given by

$$\min \mathrm{MSE}\left(\hat{M}_2\right) = \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - R_{0.12}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right)S_0^2 R_{0.12}^2 + \frac{W_2(k-1)}{n}S_{0(2)}^2\right\} \tag{3.31}$$

which equals to the variance of the optimum estimator

$$\hat{M}_{20}^{(2)} = \bar{y}^* + \beta_{01.2}\left(\bar{x}_1' - \bar{x}_1\right) + \beta_{02.1}\left(\bar{x}_2' - \bar{x}_2\right)$$

is the class of estimators

$$\hat{M}_{2d}^{(2)} = \bar{y}^* + d_1\left(\bar{x}_1' - \bar{x}_1\right) + d_2\left(\bar{x}_2' - \bar{x}_2\right),$$

where $R_{0.12}^2 = (\rho_{01}^2 + \rho_{02}^2 - 2\rho_{01}\rho_{02}\rho_{12})/(1 - \rho_{12}^2)$ is the multiple correlation coefficient of y on (x_1, x_2), $\beta_{01.2} = (\beta_{02}\beta_{21} - \beta_{01})/(1 - \beta_{21}\beta_{12})$, $\beta_{02.1} = (\beta_{01}\beta_{12} - \beta_{02})/(1 - \beta_{21}\beta_{12})$ and d_1 and d_2 are suitably

chosen constants, and $(\beta_{01}, \beta_{02}, \beta_{21}, \beta_{12})$ are the entire population regression coefficients of (y on x_1, y on x_2, x_2 on x_1, x_1 on x_2) respectively.

From (3.29) and (3.31) it can be easily shown that the proposed estimator $\hat{M}_2$ with single auxiliary variable is more efficient than $\hat{M}_2$ with two auxiliary variables.

Remark 9. In case $n' = N$ i.e. the population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ of the auxiliary variates $x_1, x_2, \ldots, x_p$ respectively are known, the class of estimators $\hat{M}_2$ given by (3.22) boils down to

$$\hat{M}_{22} = B^* \left(\bar{y}^*, w^{*T} \right), \tag{3.32}$$

where $w^{*T} = (w_1^*, w_2^*, \ldots, w_p^*)$, $w_i^* = \bar{x}_i / \bar{X}_i$, $i = 1, 2, \ldots, p$ and $B^*(\bar{y}^*, w^{*T})$ is the function of $(\bar{y}^*, w^{*T})$ such that $B^*(\bar{Y}, e^T) = \bar{Y}$ for all $\bar{Y}$.

Putting $n' = N$ in (3.25) and (3.26) we get the optimum values of the derivatives and the minimum MSE of $\hat{M}_{22}$ are respectively given by

$$B^{*(1)} \left(\bar{Y}, e^T \right) = -\bar{Y} a^{-1} q \tag{3.33}$$

and

$$\min \text{MSE} \left(\hat{M}_{22} \right) = \bar{Y}^2 \left\{ \left(\frac{1-f}{n} \right) C_0^2 + \frac{W_2(k-1)}{n} C_{0(2)}^2 - \left(\frac{1-f}{n} \right) q^T a^{-1} q \right\} \tag{3.34}$$

$$= \bar{Y}^2 \left\{ \left(\frac{1-f}{n} \right) C_0^2 \left(1 - R_{0.12\cdots p}^2 \right) + \frac{W_2(k-1)}{n} C_{0(2)}^2 \right\},$$

where $B^{*(1)}(\bar{Y}, e^T)$ denotes the p elements column vector of first partial derivatives of $B^*(\bar{y}^*, w^{*T})$ with respect to w^* about the point $w^{*T} = e^T$.

Remark 10. For the case of a single auxiliary variable x, the MSE of $\hat{M}_{22}$ is minimized for

$$B_1^{*(1)} \left(\bar{Y}, 1 \right) = -\beta \bar{X}$$

and thus the resulting minimum MSE is

$$\min \text{MSE} \left(\hat{M}_{22} \right) = \left\{ \left(\frac{1-f}{n} \right) S_0^2 \left(1 - \rho_{01}^2 \right) + \frac{W_2(k-1)}{n} S_{0(2)}^2 \right\} \tag{3.35}$$

which equals to the variance of the optimum estimator

$$\hat{M}_{220}^{(1)} = \bar{y}^* + \beta \left(\bar{X}_1 - \bar{x}_1 \right)$$

in the class of estimators

$$\hat{M}_{22d} = \bar{y}^* + d \left(\bar{X}_1 - \bar{x}_1 \right).$$

Remark 11. In case of two auxiliary variables (x_1, x_2), the MSE of $\hat{M}_{22}$ is minimized for

$$\left[\begin{array}{c} B_1^{*(1)} \left(\bar{Y}, 1, 1 \right) \\ B_2^{*(1)} \left(\bar{Y}, 1, 1 \right) \end{array} \right] = -\bar{Y} \left[\begin{array}{c} d_{01} \\ d_{02} \end{array} \right]. \tag{3.36}$$

Thus the resulting minimum MSE of $\hat{M}_{22}$ is given by

$$\min \text{MSE}\left(\hat{M}_{22}\right) = \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - R_{0.12}^2\right) + \frac{W_2(k-1)}{n}S_{0(2)}^2\right\} \tag{3.37}$$

which equals to the variance of the optimum estimator

$$\hat{M}_{220}^{(2)} = \bar{y}^* + \beta_{01.2}\left(\bar{X}_1 - \bar{x}_1\right) + \beta_{02.1}\left(\bar{X}_2 - \bar{x}_2\right)$$

in the class of estimators

$$\hat{M}_{22d} = \bar{y}^* + d_1\left(\bar{X}_1 - \bar{x}_1\right) + d_2\left(\bar{X}_2 - \bar{x}_2\right).$$

From (3.35) and (3.37) it can be shown that the proposed estimator $\hat{M}_{22}$ with two auxiliary variables is better than $\hat{M}_{22}$ with single auxiliary variable.

3.3. Strategy-III

This strategy is same as that of strategy-II. The difference is made in the formulation of class of estimators. We note that the class of estimators $\hat{M}_2$ at (3.22) utilizes information only on $\bar{x}_i$ and $\bar{x}_i'$ $(i = 1, 2, \ldots, p)$. Here we argue that an unbiased estimator $\bar{x}_i^*$ $(i = 1, 2, \ldots, p)$ can be also computed based on the responding units for which $\bar{y}^*$ is computed. Thus information on $\bar{x}_i^*$ $(i = 1, 2, \ldots, p)$ can be used along with $(\bar{y}^*, \bar{x}_i, \bar{x}_i', i = 1, 2, \ldots, p)$ in formulating the class of estimators. With this background we define a class of estimators for the population mean $\bar{Y}$ as

$$\hat{M}_3 = G\left(\bar{y}^*, u^T\right), \tag{3.38}$$

where $G(\bar{y}^*, u^T)$ is the function of $(\bar{y}^*, u^T)$ such that

$$G\left(\bar{Y}, e^T\right) = \bar{Y}, \quad \text{for all } \bar{Y} \tag{3.39}$$

$$\Rightarrow G_1\left(\bar{Y}, e^T\right) = \left.\frac{\partial G(\cdot)}{\partial \bar{y}^*}\right|_{\left(\bar{Y}, e^T\right)} = 1$$

and it also satisfies the certain conditions similar to those given for $\hat{M}_1$ at (3.1).

To obtain the mean squared error of $\hat{M}_3$, we expand the function $G(\bar{y}^*, u^T)$ about the point $(\bar{Y}, e^T)$ in a second order Taylor's series. We obtain

$$\hat{M}_3 = G\left(\bar{Y}, e^T\right) + \left(\bar{y}^* - \bar{Y}\right)G^{(1)}\left(\bar{Y}, e^T\right) + (u - e)^T G^{(1)}\left(\bar{Y}, e^T\right) + \frac{1}{2}\left\{\left(\bar{y}^* - \bar{Y}\right)^2 \left.\frac{\partial^2 G(\cdot)}{\partial \bar{y}^{*2}}\right|_{\left(\bar{y}^{**}, u^{*T}\right)}\right.$$

$$\left. + 2\left(\bar{y}^* - \bar{Y}\right)(u - e)^T \left.\frac{\partial G^{(1)}(\cdot)}{\partial \bar{y}^*}\right|_{\left(\bar{y}^{**}, u^{*T}\right)} + (u - e)^T G^{(2)}\left(\bar{y}^{**}, u^{*T}\right)(u - e)\right\},$$

where $u^* = e + \theta(u - e)$, $0 < \theta < 1$, $G^{(1)}(\bar{Y}, e^T)$ denotes the $2p$ elements column vector of the first partial derivatives of $G(\bar{y}^*, u^T)$ and $G^{(2)}\left(\bar{Y}, e^T\right)$ denotes the $2p \times 2p$ matrix of second partial derivatives of $G(\bar{y}^*, u^T)$ with respect to u. Substituting for $\bar{y}^*$ and u in terms of ε_0 and δ and using (3.16) we obtain

$$\hat{M}_3 = \bar{Y} + \bar{Y}\varepsilon_0 + \delta^T G^{(1)}\left(\bar{Y}, e^T\right)$$

$$+ \frac{1}{2}\left\{\bar{Y}^2\varepsilon_0^2 \left.\frac{\partial^2 G(\cdot)}{\partial \bar{y}^{*2}}\right|_{\left(\bar{y}^{**}, u^{*T}\right)} + 2\bar{Y}\varepsilon_0\delta^T \left.\frac{\partial G^{(1)}(\cdot)}{\partial \bar{y}^*}\right|_{\left(\bar{y}^{**}, u^{*T}\right)} + \delta^T G^{(2)}\left(\bar{y}^{**}, u^{*T}\right)\delta\right\}. \tag{3.40}$$

Taking expectation in (3.40) and noting that second partial derivatives are bounded, we have

$$E\left(\hat{M}_3\right) = \bar{Y} + o\left(n^{-1}\right)$$

and thus the bias of $\hat{M}_3$ is of order n^{-1}.

From (3.40), we have up to terms of order n^{-1},

$$\min \text{MSE}\left(\hat{M}_3\right) = E\left(\hat{M}_3 - \bar{Y}\right)^2 = E\left\{\bar{Y}\varepsilon_0 + \delta^T G^{(1)}\left(\bar{Y}, e^T\right)\right\}^2 \tag{3.41}$$

$$= \left[\bar{Y}^2\left\{\left(\frac{1-f}{n}\right)C_0^2 + \frac{W_2(k-1)}{n}C_{0(2)}^2\right\} + 2\bar{Y}b^T G^{(1)}\left(\bar{Y}, e^T\right) + \left\{G^{(1)}\left(\bar{Y}, e^T\right)\right\}^T D\left\{G^{(1)}\left(\bar{Y}, e^T\right)\right\}\right]$$

which is minimum when

$$G^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y}D^{-1}b. \tag{3.42}$$

Thus the resulting minimum MSE of $\hat{M}_3$ is given by

$$\min \text{MSE}\left(\hat{M}_3\right) = \bar{Y}^2\left\{\left(\frac{1-f}{n}\right)C_0^2 + \frac{W_2(k-1)}{n}C_{0(2)}^2 - b^T D^{-1}b\right\} \tag{3.43}$$

we note that

$$b^T D^{-1}b = g^T F^{-1}g + g_{(2)}^T F_{(2)}^{-1}g_{(2)} = \left(\frac{1}{n} - \frac{1}{n'}\right)q^T a^{-1}q + \frac{W_2(k-1)}{n}q_{(2)}^T a_{(2)}^{-1}q_{(2)}. \tag{3.44}$$

Using (3.43) in (3.41) we write the minimum MSE of $\hat{M}_3$ as

$$\min \text{MSE}\left(\hat{M}_3\right) = \bar{Y}^2\left\{\left(\frac{1-f}{n}\right)C_0^2 + \frac{W_2(k-1)}{n}C_{0(2)}^2 - \left(\frac{1}{n} - \frac{1}{n'}\right)q^T a^{-1}q - \frac{W_2(k-1)}{n}q_{(2)}^T a_{(2)}^{-1}q_{(2)}\right\}$$

$$= \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - R_{0.12\cdots p}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right)S_0^2 R_{0.12\cdots p}^2 + \frac{W_2(k-1)}{n}\left(1 - R_{(2)0.12\cdots p}^2\right)\right\}, \tag{3.45}$$

where $R_{(2)0.1,2,\ldots,p}^2 = q_{(2)}^T a_{(2)}^{-1}q_{(2)}/C_{0(2)}^2$ is the square of the multiple correlation coefficient of y on $x_1, x_2, \ldots, x_p$ of non-responding group in the population. Now we state the following theorem.

Theorem 3. *To the first degree of approximation,*

$$MSE\left(\hat{M}_3\right) \geq \bar{Y}^2\left\{\left(\frac{1-f}{n}\right)C_0^2 + \frac{W_2(k-1)}{n}C_{0(2)}^2 - b^T D^{-1}b\right\}$$

$$= \left\{\left(\frac{1-f}{n}\right)S_0^2\left(1 - R_{0.12\cdots p}^2\right) + \frac{W_2(k-1)}{n}S_{0(2)}^2\left(1 - R_{(2)0.12\cdots p}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right)S_0^2 R_{0.12\cdots p}^2\right\}$$

with equality holding if

$$G^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y}D^{-1}b.$$

Remark 12. The class of estimators $\hat{M}_3$ given by (3.38) is very large, if the parameters in the function $G(\bar{y}^*, u^T)$ are so chosen that they satisfy (3.42), the resulting estimator will have MSE given by (3.41) or (3.45). A few examples are

$$\hat{M}_{3(1)} = \bar{y}^* + \phi^T (u - e), \quad \hat{M}_{3(2)} = \frac{\bar{y}^{*2}}{\bar{y}^* - \phi^T (u - e)}, \quad \hat{M}_{3(3)} = \bar{y}^* \exp\left(\phi^T \log u\right),$$

$$\hat{M}_{3(4)} = \bar{y}^* \left\{1 + \phi^T (u - e)\right\}, \quad \hat{M}_{3(5)} = \bar{y}^* \exp\left(\phi^T (u - e)\right),$$

where $\phi^T = (\phi_1, \phi_2, \ldots, \phi_{2p})$ is a vector of $2p$ constants. The optimum values of these constants are determined from the conditions (3.42). Since (3.42) contains $2p$ equations, we have taken exactly $2p$ unknown constants in defining estimators of the class.

Remark 13. For the case of a single auxiliary variable x_1, the MSE of $\hat{M}_3$ defined at (3.41) is minimized for

$$G^{(1)}\left(\bar{Y}, 1, 1\right) = \left[\begin{array}{c} G_1^{(1)}\left(\bar{Y}, 1, 1\right) \\ G_2^{(1)}\left(\bar{Y}, 1, 1\right) \end{array} \right] = -\bar{Y} \left[\begin{array}{c} d_{(2)} \\ d - d_{(2)} \end{array} \right], \tag{3.46}$$

$$\left. \begin{array}{c} G_1^{(1)}\left(\bar{Y}, 1, 1\right) = -\bar{Y} d_{(2)} = -\bar{Y}\dfrac{\beta_{01(2)}}{R} = -\beta_{01(2)} \bar{X}_1 \\[2mm] G_2^{(1)}\left(\bar{Y}, 1, 1\right) = -\bar{Y} \left(d - d_{(2)}\right) = -\dfrac{\bar{Y}}{R} \left(\beta_{01} - \beta_{01(2)}\right) \\[2mm] = -\left(\beta_{01} - \beta_{01(2)}\right) \bar{X}_1 \end{array} \right\}, \tag{3.47}$$

where $G_1^{(1)}(\bar{Y}, 1, 1)$ and $G_2^{(1)}\left(\bar{Y}, 1, 1\right)$ denote the first order partial derivatives of $G(\bar{y}^*, \bar{x}_1^*/\bar{x}_1', \bar{x}_1/\bar{x}_1')$ about the point $(\bar{Y}, 1, 1)$.

Thus, the resulting minimum MSE of $\hat{M}_3$ is given by

$$\min \mathrm{MSE}\left(\hat{M}_3\right) = \left\{\left(\frac{1-f}{n}\right) S_0^2 \left(1 - \rho_{01}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right) S_0^2 \rho_{01}^2 + \frac{W_2(k-1)}{n} \left(1 - \rho_{01(2)}^2\right) S_{0(2)}^2\right\} \tag{3.48}$$

which equals to approximate variance of the difference estimator

$$\hat{M}_{30} = \bar{y}^* + \beta_{01(2)} \left(\bar{x}_1 - \bar{x}_1^*\right) + \beta_{01} \left(\bar{x}_1' - \bar{x}_1\right), \tag{3.49}$$

where β_{01} and $\beta_{01(2)}$ are the known regression coefficients of y on x_1 for the entire population and for non-responding group in the population.

It can be also shown to the first degree of approximation that

$$\mathrm{MSE}\left(\hat{M}_{lr1}\right) = \min \mathrm{MSE}\left(\hat{M}_3\right), \tag{3.50}$$

where $\min \mathrm{MSE}(\hat{M}_3)$ is given by (3.48) and

$$\hat{M}_{lr1} = \bar{y}^* + \hat{\beta}_{01(2)} \left(\bar{x}_1 - \bar{x}_1^*\right) + \hat{\beta}_{01} \left(\bar{x}_1' - \bar{x}_1\right),$$

where $\hat{\beta}_{01(2)}$ and $\hat{\beta}_{01}$ are the estimates based on the data available under the given sampling design of the regression coefficients $\beta_{01(2)}$ and β_{01} respectively.

Remark 14. For the case of a two auxiliary variables (x_1, x_2), the MSE of the estimator $\hat{M}_3$ at (3.41) is minimized for

$$G^{(1)}\left(\bar{Y}, e^T\right) = \begin{bmatrix} G_1^{(1)}\left(\bar{Y}, e^T\right) \\ G_2^{(1)}\left(\bar{Y}, e^T\right) \\ G_3^{(1)}\left(\bar{Y}, e^T\right) \\ G_4^{(1)}\left(\bar{Y}, e^T\right) \end{bmatrix} = -\bar{Y} \begin{bmatrix} d_{01(2)} \\ d_{02(2)} \\ d_{01(2)} - d_{01} \\ d_{02(2)} - d_{02} \end{bmatrix}.$$

Thus the resulting minimum MSE of $\hat{M}_3$ (with two auxiliary variables) is given by

$$\min \mathrm{MSE}\left(\hat{M}_3\right) = \left\{ \left(\frac{1-f}{n}\right) S_0^2 \left(1 - R_{0.12}^2\right) + \left(\frac{1}{n'} - \frac{1}{N}\right) S_0^2 R_{0.12}^2 + \frac{W_2(k-1)}{n}\left(1 - R_{(2)0.12}^2\right) S_{0(2)}^2 \right\} \quad (3.51)$$

which equals to variance of the difference estimator

$$\hat{M}_{3d(2)} = \bar{y}^* + \beta_{01.2(2)}\left(\bar{x}_1 - \bar{x}_1^*\right) + \beta_{02.1(2)}\left(\bar{x}_2 - \bar{x}_2^*\right) + \beta_{01.2}\left(\bar{x}_1' - \bar{x}_1\right) + \beta_{02.1}\left(\bar{x}_2' - \bar{x}_2\right),$$

where $\beta_{01.2(2)} = (\beta_{01(2)} - \beta_{02(2)}\beta_{21(2)})/(1 - \beta_{12(2)}\beta_{21(2)})$ and $\beta_{02.1(2)} = (\beta_{02(2)} - \beta_{01(2)}\beta_{12(2)})/(1 - \beta_{12(2)}\beta_{21(2)})$ are known partial regression coefficients of (y on x_1) and (y on x_2) for the non-responding group in the population respectively and $(\beta_{01(2)}, \beta_{02(2)}, \beta_{21(2)}, \beta_{12(2)})$ are the regression coefficients of (y on x_1, y on x_2, x_2 on x_1, x_1 on x_2) for the non- responding units group in the population respectively.

If $(\beta_{01.2}, \beta_{02.1}, \beta_{01.2(2)}, \beta_{02.1(2)})$ are not known, then one can define regression estimator (with two auxiliary variables) for $\bar{Y}$ as

$$\hat{M}_{3lr(2)} = \bar{y}^* + \hat{\beta}_{01.2(2)}\left(\bar{x}_1 - \bar{x}_1^*\right) + \hat{\beta}_{02.1(2)}\left(\bar{x}_2 - \bar{x}_2^*\right) + \hat{\beta}_{01.2}\left(\bar{x}_1' - \bar{x}_1\right) + \hat{\beta}_{02.1}\left(\bar{x}_2' - \bar{x}_2\right),$$

where $(\hat{\beta}_{01.2}, \hat{\beta}_{02.1}, \hat{\beta}_{01.2(2)}, \hat{\beta}_{02.1(2)})$ are the consistent estimates of $(\beta_{01.2}, \beta_{02.1}, \beta_{01.2(2)}, \beta_{02.1(2)})$ respectively based on the data available under the given sampling design.

It can be shown to the first degree of approximation that

$$\mathrm{MSE}\left(\hat{M}_{3lr(2)}\right) = \min \mathrm{MSE}\left(\hat{M}_3\right),$$

where $\min \mathrm{MSE}(\hat{M}_3)$ is given by (3.48).

Remark 15. If $n' = N$ (*i.e.* the population means $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p$ are known) then the class of estimators $\hat{M}_3$ defined by (3.38) reduces to the class of estimators of population mean $\bar{Y}$ as

$$\hat{M}_3^* = H\left(\bar{y}^*, z^T\right),$$

where $z^T = (Z_1, Z_2, \ldots, Z_p, Z_{p+1}, Z_{p+2}, \ldots, Z_{2p})$,

$$Z_i = \frac{\bar{x}_i^*}{\bar{X}_i}, \quad i = 1, 2, \ldots, p$$

$$= \frac{\bar{x}_{i-p}}{\bar{X}_{i-p}}, \quad i = p+1, p+2, \ldots, 2p$$

and $H(\bar{y}^*, z^T)$ is the function of $(\bar{y}^*, z^T)$ such that

$$H\left(\bar{Y}, e^T\right) = \bar{Y}, \quad \text{for all } \bar{Y}$$

$$\Rightarrow \left.\frac{\partial H(\cdot)}{\partial \bar{y}^*}\right|_{(\bar{Y}, e^T)} = H_1\left(\bar{Y}, e^T\right) = 1.$$

Putting $n' = N$ in (3.45) we get the minimum MSE of $\hat{M}_3^*$ as

$$\min \text{MSE}\left(\hat{M}_3^*\right) = \left\{\left(\frac{1-f}{n}\right) S_0^2 \left(1 - R_{0.12\cdots p}^2\right) + \frac{W_2(k-1)}{n}\left(1 - R_{(2)0.12\cdots p}^2\right) S_{0(2)}^2\right\}.$$

Thus we state the following theorem.

Theorem 4. *To the first degree of approximation*

$$MSE\left(\hat{M}_3^*\right) \geq \left\{\left(\frac{1-f}{n}\right) S_0^2 \left(1 - R_{0.12\cdots p}^2\right) + \frac{W_2(k-1)}{n} S_{0(2)}^2 \left(1 - R_{(2)0.12\cdots p}^2\right)\right\}$$

equality holding if

$$H^{(1)}\left(\bar{Y}, e^T\right) = -\bar{Y} D^{*-1} b^{*T},$$

where

$$D^* = \begin{bmatrix} E^* & F^* \\ F^{*T} & F^* \end{bmatrix}, \quad b^{*T} = \left(Q^{*T} : g^{*T}\right) = \left(Q_1^*, Q_2^*, \ldots, Q_p^*, g_1^*, g_2^*, \ldots, g_p^*\right), \quad g^{*T} = \left(\frac{1-f}{n}\right) q^T,$$

$H^{(1)}(\bar{Y}, e^T)$ *denotes the $2p$ elements column vector of the first partial derivatives of $H(\bar{y}^*, z^T)$.*

Remark 16. If $n' = N$ then $\hat{M}_{lr1}$ and $\hat{M}_{lr2}$ reduce to

$$\hat{M}_{lr1}^* = \bar{y}^* + \hat{\beta}_{01(2)}\left(\bar{x}_1 - \bar{x}_1^*\right) + \hat{\beta}_{01}\left(\bar{X}_1 - \bar{x}_1\right)$$

and

$$\hat{M}_{lr2}^* = \bar{y}^* + \hat{\beta}_{01.2(2)}\left(\bar{x}_1 - \bar{x}_1^*\right) + \hat{\beta}_{02.1(2)}\left(\bar{x}_2 - \bar{x}_2^*\right) + \hat{\beta}_{01}\left(\bar{X}_1 - \bar{x}_1\right) + \hat{\beta}_{02}\left(\bar{X}_2 - \bar{x}_2\right).$$

It can be shown to the first degree of approximation that

$$\text{MSE}\left(\hat{M}_{lr1}^*\right) = \left\{\left(\frac{1-f}{n}\right) S_0^2 \left(1 - \rho_{01}^2\right) + \frac{W_2(k-1)}{n}\left(1 - \rho_{01(2)}^2\right) S_{0(2)}^2\right\} \tag{3.52}$$

and

$$\text{MSE}\left(\hat{M}_{lr2}^*\right) = \left\{\left(\frac{1-f}{n}\right) S_0^2 \left(1 - R_{0.12}^2\right) + \frac{W_2(k-1)}{n}\left(1 - R_{(2)0.12}^2\right) S_{0(2)}^2\right\}. \tag{3.53}$$

It is to be mentioned that the regression estimator $\hat{M}_{lr1}^*$ is due to Singh and Kumar (2008a).

From (3.52) and (3.53) we have

$$\text{MSE}\left(\hat{M}_{lr1}^*\right) - \text{MSE}\left(\hat{M}_{lr2}^*\right) = \left\{\left(\frac{1-f}{n}\right)S_0^2\frac{(\rho_{02}-\rho_{01}\rho_{12})^2}{\left(1-\rho_{12}^2\right)} + \frac{W_2(k-1)}{n}\frac{(\rho_{02(2)}-\rho_{01(2)}\rho_{12(2)})^2}{\left(1-\rho_{12(2)}^2\right)}S_{0(2)}^2\right\}$$

which is always positive. Thus the estimator $\hat{M}_{lr2}^*$ is more efficient than Singh and Kumar (2008a) estimator $\hat{M}_{lr1}^*$.

Remark 17. The optimum values of the derivatives $A^{(1)}(\bar{Y}, e^T)$, $B^{(1)}(\bar{Y}, e^T)$ and $G^{(1)}(\bar{Y}, e^T)$ respectively given by (3.8), (3.25) and (3.42) are sometimes in terms of relationship between the parameters and sometimes in the form of the value of constants equated to the function of parameters. When the values of $A^{(1)}(\bar{Y}, e^T)$, $B^{(1)}(\bar{Y}, e^T)$ and $G^{(1)}(\bar{Y}, e^T)$ are of the form of relationship between the parameters then it is rarely applicable in practice. But when the values of $A^{(1)}(\bar{Y}, e^T)$, $B^{(1)}(\bar{Y}, e^T)$ and $G^{(1)}(\bar{Y}, e^T)$ give solution in the form of constants equated to some parametric function then it may be possible to use the optimum values by using the past data regarding parameters or by estimating the parameters contained in the optimum value of constant using the sample data at hand. Reddy (1978) has shown that values of such parameters are stable overtime and region. Das and Tripathi (1978) have illustrated that the guessed values of such parameters are not close enough even though the suggested estimators are better than conventional estimators. On the other hand Singh (1982) and Srivastava and Jhajj (1983) have shown that if the optimum values of constants involved in the estimators are replaced by their consistent estimators based on the sample values, the resulting estimator(s) will have the same mean squared error up to terms of order n^{-1} as that of optimum estimator(s). Thus the proposed classes of estimators $\hat{M}_1$, $\hat{M}_2$ and $\hat{M}_3$ are recommended for the use in case of large sample surveys.

4. Efficiency Comparisons

From (2.2), (3.9), (3.26) and (3.45) we have

$$\text{Var}\left(\bar{y}^*\right) - \min\text{MSE}\left(\hat{M}_1\right) = \bar{Y}^2 Q^T E^{-1} Q \geq 0 \tag{4.1}$$

$$\text{Var}\left(\bar{y}^*\right) - \min\text{MSE}\left(\hat{M}_2\right) = \bar{Y}^2 q^T a^{-1} q \geq 0 \tag{4.2}$$

$$\text{Var}\left(\bar{y}^*\right) - \min\text{MSE}\left(\hat{M}_3\right) = \bar{Y}^2 b^T D^{-1} b \geq 0 \tag{4.3}$$

$$\min\text{MSE}\left(\hat{M}_1\right) - \min\text{MSE}\left(\hat{M}_3\right) = \bar{Y}^2\left(F^T E^{-1} Q - g\right)^T A^{-1}\left(F^T E^{-1} Q^T - g\right) \geq 0 \tag{4.4}$$

$$\min\text{MSE}\left(\hat{M}_2\right) - \min\text{MSE}\left(\hat{M}_3\right) = \bar{Y}^2\frac{W_2(k-1)}{n}q_{(2)}^T a_{(2)}^{-1}q_{(2)} \geq 0, \tag{4.5}$$

where $A = (F - F^T E^{-1} F)$.

Thus all the estimators $\hat{M}_1$, $\hat{M}_2$ and $\hat{M}_3$ are better than the conventional unbiased estimator $\bar{y}^*$, which does not utilize auxiliary information. Further it follows from (4.1)–(4.5) that the estimator $\hat{M}_3$ is the best estimator among all the estimators $\bar{y}^*$, $\hat{M}_1$, $\hat{M}_2$ and $\hat{M}_3$.

5. Empirical Study

To illustrate the performance of various estimators of population mean $\bar{Y}$ relative to usual unbiased estimator $\bar{y}^*$, we consider the same data earlier considered by Khare and Sinha (2007).

Table 1: Percent relative efficiencies(PREs) of $\hat{M}_i$, $i = 1, 2, 3$ with respect to $\bar{y}^*$ for different values of k.

Estimator	Auxiliary variate(s)	Population-I				Population-II			
		1/k				1/k			
		(1/5)	(1/4)	(1/3)	(1/2)	(1/5)	(1/4)	(1/3)	(1/2)
$\bar{y}^*$	-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
$\hat{M}_1$	x_1	123.24	120.01	115.91	110.61	160.92	155.92	149.94	142.53
	x_1, x_2	127.53	125.07	122.27	118.39	186.64	177.36	166.51	153.67
$\hat{M}_2$	x_1	102.34	102.71	103.22	103.97	114.99	117.39	120.73	125.66
	x_1, x_2	107.45	108.71	110.46	113.11	133.11	139.45	148.81	164.01
$\hat{M}_3$	x_1	124.54	121.54	117.66	112.48	193.72	191.25	188.09	183.89
	x_1, x_2	133.02	130.71	127.68	123.54	235.64	228.25	219.18	207.78

The data on physical growth of upper socio-economic group of 95 school going children of Varanasi under ICMR study, Department of Paediatrics, Banaras Hindu University, during 1983–1984 has been taken under study. The first 25% (*i.e.* 24 children) units have been considered as non-responding units. The descriptions of the variates are given below:

Population-I

  y : Height(in cm.) of the children,

  x_1 : Skull circumference(in cm.) of the children,

  x_2 : Chest circumference(in cm.) of the children.

For this population we have

$\bar{Y} = 115.9526$, $\bar{X}_1 = 51.1726$, $\bar{X}_2 = 55.8611$, $C_0 = 0.05146$, $C_1 = 0.03006$, $C_2 = 0.05860$,

$C_{0(2)} = 0.04402$, $C_{1(2)} = 0.02478$, $C_{2(2)} = 0.05402$, $\rho_{01} = 0.3740$, $\rho_{02} = 0.620$

$\rho_{01(2)} = 0.571$, $\rho_{02(2)} = 0.401$, $\rho_{12} = 0.2970$, $\rho_{12(2)} = 0.570$, $N = 95$, $n = 35$, $n' = 45$.

Population-II

  y : Weight(in kg.) of the children,

  x_1 : Chest circumference(in cm.) of the children,

  x_2 : Mid-arm circumference(in cm.) of the children,

For this population we have

$\bar{Y} = 19.4968$, $\bar{X}_1 = 55.8611$, $\bar{X}_2 = 16.7968$, $C_0 = 0.15613$, $C_1 = 0.05860$, $C_2 = 0.08651$,

$C_{0(2)} = 0.12075$, $C_{1(2)} = 0.05402$, $C_{2(2)} = 0.07125$, $\rho_{01} = 0.846$, $\rho_{02} = 0.797$,

$\rho_{01(2)} = 0.729$, $\rho_{02(2)} = 0.757$, $\rho_{12} = 0.725$, $\rho_{12(2)} = 0.641$, $N = 95$, $n = 35$, $n' = 45$.

We have computed the percent relative efficiencies of the proposed estimators(PREs), $\hat{M}_1$, $\hat{M}_2$ and $\hat{M}_3$ with respect to conventional unbiased estimator $\bar{y}^*$ for the two data sets and findings are displayed in Table 1.

Table 1 exhibits that the estimator $\hat{M}_3$ is the best estimator among $\bar{y}^*$, $\hat{M}_1$, $\hat{M}_2$ and $\hat{M}_3$ in both the populations respectively for single as well as double auxiliary variables. It is also observed that for single as well as double auxiliary variables the percent relative efficiencies(PREs) of the estimators $\hat{M}_1$ and $\hat{M}_3$ over the usual unbiased estimator decreases as the value of $\bar{y}^*$ increases while increases for the estimator $\hat{M}_2$ as the value of k increases for both the populations.

References

Agarwal, M. C. and Panda, K. B. (1993). An efficient estimator in post stratification, *Metron*, **L1**, 179–188.

Cochran, W. G. (1977). *Sampling Techniques*, 3rd ed., John Wiley & Sons, New York.

Das, A. K. and Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance, *Sankhya (C)*, **40**, 139–148.

Hansen, M. H. and Hurwitz, W. N. (1946). The problem of non-response in sample surveys, *Journal of the American Statistical Association*, **41**, 517–529.

Khare, B. B. and Sinha, R. R. (2004). Estimation of finite population ratio using two-phase sampling scheme in the presence of non-response, *Aligarh Journal Statistic*, **24**, 43–56.

Khare, B. B. and Sinha, R. R. (2007). Estimation of the ratio of the two population means using multi auxiliary characters in the presence of non-response, *Statistical Techniques in Life Testing, Reliability, Sampling Theory and Quality Control*, 163–171.

Khare, B. B. and Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response, *National Academy of Sciences Letters, India*, **16**, 111–114.

Khare, B. B. and Srivastava, S. (1995). Study of conventional and alternative two phase sampling ratio, product and regression estimators in presence of non-response, *Proceedings of the National Academy of Sciences India*, **65(A)**, II, 195–203.

Khare, B. B. and Srivastava, S. (1997). Transformed ratio type estimators for the population mean in the presence of non-response, *Communications in Statistics - Theory and Methods*, **26**, 1779–1791.

Mohanty, S. and Pattanaik, L. M. (1982). Alternative multivariate ratio estimators using geometric and harmonic means, *Journal of the Indian Society of Agricultural Statistics*, **36**, 110–117.

Murthy, M. N. (1964). Product method of estimation, *Sankhya (A)*, **26**, 69–74.

Okafor, F. C. (1996). On double sampling for Stratification with sub-sampling the non-respondents, *Journal of the Indian Society of Agricultural Statistics*, **48**, 105–113.

Olkin, I. (1958). Multivariate ratio estimation for finite population, *Biometrika*, **45**, 154–165.

Raj, D. (1965). On sampling over two occasions with probability proportional to size, *The Annals of Mathematical Statistics*, **36**, 327–330.

Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the non-respondents, *Survey Methodology*, **12**, 217–230.

Rao, P. S. R. S. (1987). Ratio and regression estimates with sub sampling the non- respondents. Paper presented at a special contributed session of the International Statistical Association Meeting, Sept., 2-16, Tokyo, Japan.

Rao, P. S. R. S. and Mudholkar, G. S. (1967). Generalized multivariate estimator for the mean of finite populations, *Journal of the American Statistical Association*, **62**, 1009–1012.

Reddy, V. N. (1978). A study of use of prior knowledge on certain population parameters in estimation, *Sankhya (C)*, **40**, 29–37.

Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimation, *Journal of the American Statistical Association*, **52**, 511–522.

Sarndal, C. E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*, Springer-Verlag.

Singh, H. P. and Kumar, S. (2008a). A regression approach to the estimation of finite population mean in presence of non-response, *Australian & New Zealand Journal of Statistics*, **50**, 395–408.

Singh, H. P. and Kumar, S. (2008b). Estimation of mean in presence of non-response using two phase

sampling scheme, *Statistical Papers*, DOI 10.1007/s00362-008-040-5.

Singh, H. P. and Kumar, S. (2009a). A general class of estimators of the population mean in survey sampling using auxiliary information with sub sampling the non-respondents, *The Korean Journal of Applied Statistics*, **22**, 387–402.

Singh, H. P. and Kumar, S. (2009b). A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information, *SORT*, **33**, 71–84.

Singh, H. P., Kumar, S. and Kozak, M. (2010). Improved estimation of finite - population mean when subsampling is employed to deal with non- response, *Communications in Statistics - Theory and Methods*, **39**, 791–802.

Singh, M. P. (1967). Multivariate product method of estimation, *Journal of the Indian Society of Agricultural Statistics*, **19**, 1–10.

Singh, R. K. (1982). Generalized estimators of estimation of ratio and product of population parameters, *Journal of Statistical Research*, **16**, 1–2.

Srivastava, S. K. (1971). A generalized estimator for the mean of a finite population using multi auxiliary information, *Journal of the American Statistical Association*, **66**, 404–407.

Srivastava, S. K. (1981). A generalized two-phase sampling estimator, *Journal of the Indian Society of Agricultural Statistics*, **33**, 38–46.

Srivastava, S. K. and Jhajj, H. S. (1983). A class of estimators of the population mean using multi auxiliary information, *Calcutta Statistical Association Bulletin*, **32**, 47–56.

Tabasum, R. and Khan, I. A. (2004). Double sampling for ratio estimation with non- response, *Journal of the Indian Society of Agricultural Statistics*, **58**, 300–306.

Tabasum, R. and Khan, I. A. (2006). Double sampling ratio estimator for the population mean in presence of non-response, *Assam Statistical Review*, **20**, 73–83.