Some Estimation Procedures in Presence of Non-Response in Two-Phase Sampling

A. Bandyopadhyay[1*] and G. N. Singh[2]

[1]Department of Mathematics, Asansol Engineering College, Asansol–713305, India.

[2]Department of Applied Mathematics, Indian School of Mines, Dhanbad-826004, India.

Corresponding Author: arnabbandyopadhyay4@gmail.com

Received: 31, March 2014 / Revised: 01, December 2014 / Accepted: 25, February 2015

ABSTRACT

The present investigation deals with the problem of estimation of population mean in presence of non-response under two-phase (double) sampling. Following the technique of sub-sampling of non-responding group adopted by Hansen and Hurwitz (1946) and using information on two auxiliary variables, four general classes of estimators have been suggested for four different situations of non-response and their properties are studied. It is shown that several estimators can be generated from our proposed classes of estimators. Comparisons of the proposed strategies with some contemporary estimators of population mean in presence of non-response are carried out. The results obtained are illustrated numerically through empirical studies which present the effectiveness of the suggested classes of estimators.

Keywords: Two-phase sampling, study variable, auxiliary variable, non response, bias, mean square error.

1. Introduction

In surveys covering human populations in most cases, information is not obtained from all the units in the survey at the first attempt even after some call-backs. An estimate obtained from such incomplete data may be misleading especially when the respondents differ from the non-respondents. In order to reduce the effect of non-response in such situations, (Hansen and Hurwitz, 1946) gave a technique of sub-sampling of the non-responding group. It is well known fact that in sample surveys precision in estimating the population mean may be increased by using information on single or multiple auxiliary variables. Following (Hansen and Hurwitz, 1946) technique, several authors including (Cochran, 1977; Rao, 1986;

Khare and Srivastava, 1993, 1995, 1997; Okafor and Lee, 2000; Tabasum and Khan, 2004, 2006; Singh and Kumar, 2010 a) have contributed towards the improvement of the estimation procedure of population mean in presence of non-response using information on auxiliary variable. (Olkin, 1958; Mohanty, 1967; Srivastava, 1971; Singh and Kumar, 2010 b; Khare and Sinha, 2012) and others have made the extension of the ratio method of estimation to the case where multiple auxiliary variables are used to increase the precision of estimates. In many situations, information on the auxiliary variable may be readily available on all the units of the population; for example, tonnage (or seat capacity) of each vehicle or ship is known in survey sampling of transportation and number of beds in different hospitals may be known in hospital surveys. When such information is lacking, it is sometimes, relatively cheap to take a large preliminary sample in which auxiliary variable alone is measured. This technique is known as double sampling or two-phase sampling. Two-phase sampling happens to be a powerful and cost effective (economical) technique for obtaining the reliable estimate in first-phase (preliminary) sample for the unknown population parameters of the auxiliary variables. For example, (Tabasum and Khan, 2004) have mentioned that the procedure of double sampling can be applied in a household survey where the household size is used as an auxiliary variable for the estimation of family expenditure. Information can be obtained completely on the family size, while there may be non-response on the household expenditure.

Motivated with the above arguments and following the technique of sub-sampling of the non-responding group, we have proposed four general classes of estimators for four different situations of non-response in two-phase sampling using information on two auxiliary variables and studied their properties. It is shown that several estimators can be generated from our proposed classes of estimators. The superiorities of the proposed classes of estimators over some existing estimators have been established through theoretical and empirical comparisons.

2. Proposed Classes of Estimators

Consider a finite population $U = (U_1, U_2, \ldots, U_N)$ of N units, y, x and z are the variables under study, first auxiliary variable and second auxiliary variable respectively with population means $\bar{Y}, \bar{X}$ and $\bar{Z}$. Let y_k, x_k and z_k be the values of y, x and z for the k-th (k = 1, 2, …, N) unit in the population. If the information on an auxiliary variable x whose population mean is known and highly correlated to y is readily available for all the units of the population, it is well known that regression and ratio type estimators of population mean $\bar{Y}$ could be used for good performance. However, in certain practical situations when population mean $\bar{X}$ is not known, a priori in such case the technique of two-phase sampling is useful. Thus, to estimate the population mean $\bar{X}$, a first phase sample of size n' is drawn from the entire population U by simple random sampling without replacement (SRSWOR) and a second phase sample of size n (i.e., $n' > n$) is selected from the first phase by SRSWOR and the variable y under investigation is measured on it. If there is non-response in the second phase sample one may form an estimator by utilizing the information only from the respondents or take a sub-sample of the non-respondents and re-contact them. We assume that at the first phase sample of size n', all the units supplied information on the auxiliary variables x and z and at the second phase sample of size n, let n_1 units supply information on y and n_2 refuse to respond. Following (Hansen and Hurwitz, 1946) technique of sub-sampling the non-responding group, a sub-sample of size m units $\left(m = n_2/k,\ k > 1\right)$ is selected at random (without replacement) from the n_2 non-respondent units and is enumerated by direct interview. It is assumed that response is obtained for all the m units and the whole population (i. e., U) is supposed to be consisting of two non-overlapping strata of N_1 and N_2 units. Stratum of N_1 responding units $\left(\text{denoted by } U_1\right)$ would respond on the first call at the second phase and the stratum of $N_2\left(N_2 = N - N_1\right)$ non-responding units $\left(\text{denoted by } U_2\right)$ would not respond on the first call at the second phase but will respond on the second call. Further, we assume that the strata sizes of N_1 and N_2 are not known well in advance, see (Tripathi and Khare, 1997). The stratum weights of responding and non-responding groups are given by $W_1 = \frac{N_1}{N}$ and $W_2 = \frac{N_2}{N}$ and their estimates are considered as $\hat{W}_1 = \frac{n_1}{n}$ and $\hat{W}_2 = \frac{n_2}{n}$ respectively. Let first and second phase sample be denoted by u' and u respectively and $u_1 = u \cap U_1$ and $u_2 = u \cap U_2$. The sub-sample of u_2 will be denoted by u_{2m}.

If non-response occurs on the study variable y as well as on the auxiliary variable x in the second phase sample, the conventional two-phase ratio, product and regression estimators for population mean $\bar{Y}$ are considered as

$$t_1 = \frac{\bar{y}^*}{\bar{x}^*}\bar{x}', \tag{1}$$

$$t_2 = \frac{\bar{y}^*}{\bar{x}'}\bar{x}^* \tag{2}$$

and

$$t_3 = \bar{y}^* + b_1^* \left(\bar{x}' - \bar{x}^* \right) \tag{3}$$

where $\bar{x}^*$ and $\bar{y}^*$ are the Hansen–Hurwitz estimators for population means $\bar{X}$ and $\bar{Y}$ respectively and are defined by $\bar{x}^* = \dfrac{n_1 \bar{x}_1 + n_2 \bar{x}_{2m}}{n}$ and $\bar{y}^* = \dfrac{n_1 \bar{y}_1 + n_2 \bar{y}_{2m}}{n}$ with $\left(\bar{x}_1, \bar{y}_1\right)$ and $\left(\bar{x}_{2m}, \bar{y}_{2m}\right)$ are the sample means of (x, y) variables based on the samples of n_1 and m units respectively, $\bar{x}'$ denotes the sample mean of the variable x based on first phase sample of size n', $b_1^* = \dfrac{S_{yx}^*}{S_x^{*2}}$ is the estimate of the population (i. e., U) regression coefficient β_{yx} of y on x,

$$S_{yx}^* = \frac{1}{(n-1)} \left(\sum_{u_1} x_i y_i + m \sum_{u_{2m}} x_i y_i - n \bar{x}^* \bar{y}^* \right),$$

$$S_x^{*2} = \frac{1}{(n-1)} \left(\sum_{u_1} x_i^2 + m \sum_{u_{2m}} x_i^2 - n \bar{x}^{*2} \right)$$ {for instance see Okafor and Lee, 2000)}.

But, when non-response situation is observed only on the study variable y, while the complete information on the auxiliary variable x is available in the second phase sample, the conventional double sampling ratio, product and regression estimators are suggested as

$$t_4 = \frac{\bar{y}^*}{\bar{x}} \bar{x}', \tag{4}$$

$$t_5 = \frac{\bar{y}^*}{\bar{x}'} \bar{x}, \tag{5}$$

and

$$t_6 = \bar{y}^* + b_1' \left(\bar{x}' - \bar{x} \right) \tag{6}$$

where $\bar{x}$ are the sample mean of the variable x based on the second phase sample of size n, $b_1' = \dfrac{S_{yx}^*}{S_x^2}$ is the estimate of the population regression coefficient β_{yx} and $S_x^2 = \left(N-1\right)^{-1} \sum_{i=1}^{N} (x_i - \bar{X})^2$.

It is to be noted that the estimators $t_i \left(i = 1, 2, \ldots, 4\right)$ were first proposed and studied by Khare and Srivastava (1993). The estimator t_1 was revisited by

(Okafor and Lee, 2000; Tabasum and Khan, 2004). Further the estimator t_4 was reconsidered by (Tabasum and Khan, 2006). The estimator t_6 were first envisaged by (Khare and Srivastava, 1995) and the estimator t_3 was revisited by (Okafor and Lee, 2000).

Motivated by the above suggestions and following the two-phase sampling structure defined above with the assumption that the population mean of the auxiliary variable x be unknown, we have proposed following four general classes of estimators of population mean $\bar{Y}$ of the study variable y applicable for four different situations of non-response.

Situation I

In this case, we assume that the non- response conditions occur on the study variable y as well as on the auxiliary variable x in the second phase sample of size n and also the population mean $\bar{Z}$ of the second auxiliary variable z be known. Accordingly, we have suggested the general class of estimators of population mean $\bar{Y}$ in two-phase sampling as

$$T_1 = f\left(\bar{y}^*, \bar{x}^*, h_1\left(\bar{x}', \bar{z}'\right)\right) \tag{7}$$

where $\bar{z}'$ is the sample mean of the variable z based on the first phase sample of size n', $h_1\left(\bar{x}', \bar{z}'\right)$ be a class of estimators of $\bar{X}$ using information on $\bar{x}'$ and $\bar{z}'$, such that

$$h_1\left(\bar{X}, \bar{Z}\right) = \bar{X}. \tag{8}$$

We treat the composite function $f\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ as one-to-one function of $\bar{y}^*, \bar{x}^*, \bar{x}'$ and $\bar{z}'$ denoted by $T_1 = F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ such that

$$F\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right) = \bar{Y} \Rightarrow \left.\frac{\partial F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)}{\partial \bar{y}^*}\right|_{\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)} = 1 \tag{9}$$

with $\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)$ and $F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ satisfies the following regularity conditions:

1. Whatever the chosen samples, $\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ assume values in a closed convex subspace, R^4 of the four dimensional real space containing the point $\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)$.

2. The function $F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ is continuous and bounded in R^4.

3. The first, second and third order partial derivatives of $F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ exist and are continuous and bounded in R^4.

It can be observed from equation (7) that the class of estimators T_1 is very wide in the sense for any parametric function, $f\left(\bar{y}^*, \bar{x}^*, h_1\left(\bar{x}', \bar{z}'\right)\right)$ satisfying above regularity conditions with $F\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right) = \bar{Y}$ can generate an estimators of $\bar{Y}$. For examples, the following ratio, product, regression and exponential type estimators are the member of the class T_1.

$$t_{1i} = \bar{y}^* \frac{\bar{x}_i'}{\bar{x}^*},\ t_{2i} = \bar{y}^* \frac{\bar{x}^*}{\bar{x}_i'},\ t_{3i} = \bar{y}^* + b_1^*\left(\bar{x}_i' - \bar{x}^*\right),\ t_{4i} = \bar{y}^* \exp\left(\frac{\bar{x}_i' - \bar{x}^*}{\bar{x}_i' + \bar{x}^*}\right);\ \left(i = 1, 2, \ldots, 4\right)$$

where $\bar{x}_1' = \bar{x}' \frac{\bar{Z}}{\bar{z}'},\ \bar{x}_2' = \bar{x}' \frac{\bar{z}'}{\bar{Z}},\ \bar{x}_3' = \bar{x}' + b_{xz}(n')\left(\bar{Z} - \bar{z}'\right),\ \bar{x}_4' = \bar{x}' \exp\left(\frac{\bar{Z} - \bar{z}'}{\bar{Z} + \bar{z}'}\right)$ and

$b_{xz}(n')$ is the estimate of the population (entire population i. e. U) regression coefficient β_{xz} of x on z based on the first phase sample of size n'.

Situation II

In this situation, we assume that the non- response occurs on the study variable y as well as on the auxiliary variables x and z in the second phase sample of size n and the population mean of the auxiliary variable z be unknown. Considering these aspects, we have formed the general class of estimators of $\bar{Y}$ in two-phase sampling as

$$T_2 = g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right) \tag{10}$$

where $g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)$ is a function of $\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}'$ and $\bar{z}'$ such that

$$g\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right) = \bar{Y} \Rightarrow \left.\frac{\partial g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)}{\partial \bar{y}^*}\right|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} = 1 \tag{11}$$

where $\bar{z}^*$ is the Hansen–Hurwitz estimator for population mean $\bar{Z}$ and is defined by $\bar{z}^* = \frac{n_1\bar{z}_1 + n_2\bar{z}_{2m}}{n}$ with $\bar{z}_1$ and $\bar{z}_{2m}$ are the sample means of the variable z based on the samples of n_1 and m units respectively, $\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)$ and $g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)$ satisfies regularity conditions similar to those given for $\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)$ and $F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ in equation (9).

It can be observed from equation (10) that for any parametric function, $g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)$ satisfying above regularity conditions with $g\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)=\bar{Y}$ for all $\bar{Y}$, can generate an estimator of $\bar{Y}$. For examples, we present below few ratio, product, regression and exponential type estimators as the members of the class of estimators T_2 .

$$t'_{1i}=\bar{y}^*\frac{\bar{x}'}{\bar{x}_i^*},\ t'_{2i}=\bar{y}^*\frac{\bar{x}_i^*}{\bar{x}'},\ t'_{3i}=\ \bar{y}^*+b_1^*\left(\bar{x}'-\bar{x}_i^*\right),\ t'_{4i}=\ \bar{y}^*\exp\left(\frac{\bar{x}'-\bar{x}_i^*}{\bar{x}'+\bar{x}_i^*}\right);\left(i=1, 2, \ldots, 4\right),$$

$$t''_{1j}=\bar{y}_j^*\frac{\bar{x}'}{\bar{x}^*},\ t''_{2j}=\bar{y}_j^*\frac{\bar{x}'}{\bar{x}^*},\ t''_{3j}=\ \bar{y}_j^*+b_1^*\left(\bar{x}'-\bar{x}^*\right),\ t''_{4j}=\ \bar{y}_j^*\exp\left(\frac{\bar{x}'-\bar{x}^*}{\bar{x}'+\bar{x}^*}\right);\left(j=1, 2, \ldots, 4\right)$$

where $\bar{x}_1^*=\bar{x}^*\frac{\bar{z}'}{\bar{z}^*}$, $\bar{x}_2^*=\bar{x}^*\frac{\bar{z}^*}{\bar{z}'}$, $\bar{x}_3^*=\bar{x}^*+b_2^*\left(\bar{z}'-\bar{z}^*\right)$, $\bar{x}_4^*=\ \bar{x}^*\exp\left(\frac{\bar{z}'-\bar{z}^*}{\bar{z}'+\bar{z}^*}\right)$,

$\bar{y}_1^*=\bar{y}^*\frac{\bar{z}'}{\bar{z}^*}$, $\bar{y}_2^*=\bar{y}^*\frac{\bar{z}^*}{\bar{z}'}$, $\bar{y}_3^*=\bar{y}^*+b_3^*\left(\bar{z}'-\bar{z}^*\right)$, $\bar{y}_4^*=\ \bar{y}^*\exp\left(\frac{\bar{z}'-\bar{z}^*}{\bar{z}'+\bar{z}^*}\right)$, $b_2^*=\frac{S_{xz}^*}{S_z^{*2}}$ is the estimate of the population regression coefficient β_{xz} , $b_3^*=\frac{S_{yz}^*}{S_z^{*2}}$ is the estimate of the population (i. e., U) regression coefficient β_{yz} of y on z,

$$S_{xz}^*=\frac{1}{(n\text{-}1)}\left(\sum_{u_1}x_iz_i+m\sum_{u_{2m}}x_iz_i\ -n\bar{x}^*\bar{z}^*\right),$$

$$S_{yz}^*=\frac{1}{(n\text{-}1)}\left(\sum_{u_1}y_iz_i+m\sum_{u_{2m}}y_iz_i\ -n\bar{y}^*\bar{z}^*\right),\ S_z^{*2}=\frac{1}{(n\ \text{-}1)}\left(\sum_{u_1}z_i^2+m\sum_{u_{2m}}z_i^2\ -n\bar{z}^{*2}\right).$$

Situation III

In this case, we assume that the non-response situation occurs only on the study variable y while the complete information on the auxiliary variable x is available in second phase sample of size n and also the population mean $\bar{Z}$ of the second auxiliary variable z be known. Considering this situation, we have proposed the general class of estimators of population mean $\bar{Y}$ in two-phase sampling as

$$T_3=\varphi\left(\bar{y}^*, \bar{x}, h_2\left(\bar{x}', \bar{z}'\right)\right) \tag{12}$$

where $h_2\left(\bar{x}', \bar{z}'\right)$ be a class of estimators of $\bar{X}$ using information on $\bar{x}'$ and $\bar{z}'$, such that

$$h_2\left(\bar{X}, \bar{Z}\right)=\bar{X}. \tag{13}$$

We treat the composite function $\varphi\left(\bar{y}^{*}, \bar{x}, \bar{x}', \bar{z}'\right)$ as one-to-one function of $\bar{y}^{*}, \bar{x}, \bar{x}'$ and $\bar{z}'$ denoted by $T_3 = G\left(\bar{y}^{*}, \bar{x}, \bar{x}', \bar{z}'\right)$ such that

$$G\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right) = \bar{Y} \Rightarrow \left.\frac{\partial G\left(\bar{y}^{*}, \bar{x}, \bar{x}', \bar{z}'\right)}{\partial \bar{y}^{*}}\right|_{\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)} = 1 \tag{14}$$

with $\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)$ and $G\left(\bar{y}^{*}, \bar{x}, \bar{x}', \bar{z}'\right)$ satisfies the similar regularity conditions given for $\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)$ and $F\left(\bar{y}^{*}, \bar{x}^{*}, \bar{x}', \bar{z}'\right)$ in equation (9).

Situation IV

In this case, we assume that in the second phase sample non- response situation is found on the study variable y and the auxiliary variable z with unknown $\bar{Z}$ while the complete information on the auxiliary variable x is available. Considering this situation, we have suggested the general class of estimators of population mean $\bar{Y}$ in two-phase sampling as

$$T_4 = \psi\left(\bar{y}^{*}, \bar{x}, \bar{z}^{*}, \bar{x}', \bar{z}'\right) \tag{15}$$

where $\psi\left(\bar{y}^{*}, \bar{x}, \bar{z}^{*}, \bar{x}', \bar{z}'\right)$ is a function of $\bar{y}^{*}, \bar{x}, \bar{z}^{*}, \bar{x}'$ and $\bar{z}'$ such that

$$\psi\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right) = \bar{Y} \Rightarrow \left.\frac{\partial \psi\left(\bar{y}^{*}, \bar{x}, \bar{z}^{*}, \bar{x}', \bar{z}'\right)}{\partial \bar{y}^{*}}\right|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} = 1 \tag{16}$$

and $\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)$ and $\psi\left(\bar{y}^{*}, \bar{x}, \bar{z}^{*}, \bar{x}', \bar{z}'\right)$ satisfies the similar regularity conditions as presented for the class of estimators T_1 above.

Proceeding as above, it can also be found that the classes of estimators T_3 and T_4 are also very wide and the following estimators can be identified as their member.

Estimators belonging to the class T_3 :

$$t_{1i}^{*} = \bar{y}^{*}\frac{\bar{x}_i'}{\bar{x}},\ t_{2i}^{*} = \bar{y}^{*}\frac{\bar{x}}{\bar{x}_i'},\ t_{2i}^{*} = \bar{y}^{*} + b_1'\left(\bar{x}_i' - \bar{x}\right),\ t_{4i}^{*} = \bar{y}^{*}\exp\left(\frac{\bar{x}_i' - \bar{x}}{\bar{x}_i' + \bar{x}}\right); \left(i = 1, 2, \ldots, 4\right)$$

Estimators belonging to the class T_4 :

$$t_{1i}^{**} = \bar{y}_i^{*}\frac{\bar{x}'}{\bar{x}},\ t_{2i}^{**} = \bar{y}_i^{*}\frac{\bar{x}'}{\bar{x}},\ t_{3i}^{**} = \bar{y}_i^{*} + b_1'\left(\bar{x}' - \bar{x}\right),\ t_{4i}^{**} = \bar{y}_i^{*}\exp\left(\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\right);\ \left(i = 1, 2, \ldots, 4\right)$$

$$t_{1j}^{***} = \bar{y}^* \frac{\bar{x}'}{\bar{x}_j},\ t_{2j}^{***} = \bar{y}^* \frac{\bar{x}_j}{\bar{x}'},\ t_{3j}^{***} = \bar{y}^* + b_1'\left(\bar{x}' - \bar{x}_j\right),\ t_{4j}^{***} = \bar{y}^* \exp\left(\frac{\bar{x}' - \bar{x}_j}{\bar{x}' + \bar{x}_j}\right);\ \left(j = 1, 2, 3\right)$$

where $\bar{x}_1 = \bar{x}\frac{\bar{z}'}{\bar{z}^*},\ \bar{x}_2 = \bar{x}\frac{\bar{z}^*}{\bar{z}'}$ and $\bar{x}_3 = \bar{x}\exp\left(\frac{\bar{z}' - \bar{z}^*}{\bar{z}' + \bar{z}^*}\right)$.

3. Bias and Mean Square Errors of the Proposed Classes of Estimators

$T_i \left(i = 1, 2, \ldots, 4\right)$

The bias and mean square errors (M. S. E.s) of the proposed classes of estimators $T_i \left(i = 1, 2, \ldots, 4\right)$ to the first order of approximations are derived under large sample approximations using the following transformations:

$$\bar{y}^* = \bar{Y}\left(1 + e_0\right),\ \bar{x}^* = \bar{X}\left(1 + e_1\right),\ \bar{z}^* = \bar{Z}\left(1 + e_2\right),\ \bar{x}' = \bar{X}\left(1 + e_1'\right),\ \bar{z}' = \bar{Z}\left(1 + e_2'\right),\ \bar{x} = \bar{X}\left(1 + e_3\right).$$

Such that $\left|e_i\right|$ and $\left|e_j'\right|$ are < 1 (i = 0, 1, . . ., 3; j= 1, 2). Further, we have the following expectations:

$$\left.\begin{aligned}
&E\left(e_0^2\right) = f_1 C_y^2 + W_2 \frac{(k-1)}{n} C_{y(2)}^2,\ E\left(e_1^2\right) = f_1 C_x^2 + W_2 \frac{(k-1)}{n} C_{x(2)}^2,\ E\left(e_1'^2\right) = f_2 C_x^2,\ E\left(e_2'^2\right) = f_2 C_z^2,\\
&E\left(e_2^2\right) = f_1 C_z^2 + W_2 \frac{(k-1)}{n} C_{z(2)}^2,\ E\left(e_3^2\right) = f_1 C_x^2,\ E\left(e_0 e_1\right) = f_1 \rho_{yx} C_y C_x + W_2 \frac{(k-1)}{n} \rho_{yx(2)} C_{y(2)} C_{x(2)},\\
&E\left(e_0 e_2\right) = f_1 \rho_{yz} C_y C_z + W_2 \frac{(k-1)}{n} \rho_{yz(2)} C_{y(2)} C_{z(2)},\ E\left(e_1 e_2\right) = f_1 \rho_{xz} C_x C_z + W_2 \frac{(k-1)}{n} \rho_{xz(2)} C_{x(2)} C_{z(2)},\\
&E\left(e_0 e_1'\right) = f_2 \rho_{yx} C_y C_x,\ E\left(e_0 e_3\right) = f_1 \rho_{yx} C_y C_x,\ E\left(e_0 e_2'\right) = f_2 \rho_{yz} C_y C_z,\ E\left(e_2 e_3\right) = f_1 \rho_{xz} C_x C_z\\
&E\left(e_1' e_3\right) = E\left(e_1 e_1'\right) = f_2 C_x^2,\ E\left(e_1 e_2'\right) = E\left(e_1' e_2\right) = E\left(e_1' e_2'\right) = E\left(e_2' e_3\right) = f_2 \rho_{xz} C_x C_z,\ E\left(e_0\right) = 0\\
&E\left(e_1\right) = E\left(e_2\right) = E\left(e_3\right) = E\left(e_1'\right) = E\left(e_2'\right) = 0,\ f_1 = \left(\frac{1}{n} - \frac{1}{N}\right),\ f_2 = \left(\frac{1}{n'} - \frac{1}{N}\right),\ f_3 = \left(\frac{1}{n} - \frac{1}{n'}\right)
\end{aligned}\right\} \quad (17)$$

where

$\rho_{yx}, \rho_{yz}, \rho_{xz}$: correlation coefficients between the variables shown in suffice based on the whole population (i. e. U),

C_x, C_y, C_z : coefficient of variations of the variables x, y and z respectively based on the whole population,

$\rho_{yx(2)}, \rho_{yz(2)}, \rho_{xz(2)}$: correlation coefficients between the variables shown in suffice in the non-response group of the population (i. e. U_2),

$\bar{X}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} x_i$, $\bar{Y}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} y_i$, $\bar{Z}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} z_i$: population means of the variables x, y and z respectively in the non-response group of the population,

$S^2_{x(2)} = (N_2 - 1)^{-1}\sum_{i=1}^{N_2}(x_i - \bar{X}_2)^2$, $S^2_{y(2)} = (N_2 - 1)^{-1}\sum_{i=1}^{N_2}(y_i - \bar{Y}_2)^2$, $S^2_{z(2)} = (N_2 - 1)^{-1}\sum_{i=1}^{N_2}(z_i - \bar{Z}_2)^2$: population mean squares of the variables x, y and z respectively in the non-response group of the population and $C_{x(2)} = \frac{S_{x(2)}}{\bar{X}}$, $C_{y(2)} = \frac{S_{y(2)}}{\bar{Y}}$, $C_{z(2)} = \frac{S_{z(2)}}{\bar{Z}}$.

Now, to express T_1 in terms of e's, we expand $F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')$ about the point $(\bar{Y}, \bar{X}, \bar{X}, \bar{Z})$ in a third order Taylor's series and we have

$$F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}') = F(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}) + d_1(\bar{y}^* - \bar{Y}) + d_2(\bar{x}^* - \bar{X}) + d_3(\bar{x}' - \bar{X}) + d_4(\bar{z}' - \bar{Z}) \tag{18}$$

$$+ \frac{1}{2}\Big\{d_{11}(\bar{y}^* - \bar{Y})^2 + d_{22}(\bar{x}^* - \bar{X})^2 + d_{33}(\bar{x}' - \bar{X})^2 + d_{44}(\bar{z}' - \bar{Z})^2$$

$$+ 2d_{12}(\bar{y}^* - \bar{Y})(\bar{x}^* - \bar{X}) + 2d_{13}(\bar{y}^* - \bar{Y})(\bar{x}' - \bar{X}) + 2d_{14}(\bar{y}^* - \bar{Y})(\bar{z}' - \bar{Z})$$

$$+ 2d_{23}(\bar{x}^* - \bar{X})(\bar{x}' - \bar{X}) + 2d_{24}(\bar{x}^* - \bar{X})(\bar{z}' - \bar{Z}) + 2d_{34}(\bar{x}' - \bar{X})(\bar{z}' - \bar{Z})\Big\}$$

$$+ \frac{1}{6}\left\{(\bar{y}^* - \bar{Y})\frac{\partial}{\partial \bar{y}^*} + (\bar{x}^* - \bar{X})\frac{\partial}{\partial \bar{x}^*} + (\bar{x}' - \bar{X})\frac{\partial}{\partial \bar{x}'} + (\bar{z}' - \bar{Z})\frac{\partial}{\partial \bar{z}'}\right\}^3 F(\bar{y}^*_\theta, \bar{x}^*_\theta, \bar{x}'_\theta, \bar{z}'_\theta)$$

where

$$d_1 = \frac{\partial}{\partial \bar{y}^*}F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')\Big|_{(\bar{Y}, \bar{X}, \bar{X}, \bar{Z})}, \quad d_2 = \frac{\partial}{\partial \bar{x}^*}F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')\Big|_{(\bar{Y}, \bar{X}, \bar{X}, \bar{Z})},$$

$$d_3 = \frac{\partial}{\partial \bar{x}'}F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')\Big|_{(\bar{Y}, \bar{X}, \bar{X}, \bar{Z})}, \quad d_4 = \frac{\partial}{\partial \bar{z}'}F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')\Big|_{(\bar{Y}, \bar{X}, \bar{X}, \bar{Z})}$$

$(d_{11}, d_{22}, d_{33}, d_{44}, d_{12}, d_{13}, d_{14}, d_{23}, d_{24}, d_{34})$ are the second order partial derivatives of $F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')$ at the point $(\bar{Y}, \bar{X}, \bar{X}, \bar{Z})$ and $\bar{y}^*_\theta = \bar{Y} + \theta(\bar{y}^* - \bar{Y})$, $\bar{x}^*_\theta = \bar{X} + \theta(\bar{x}^* - \bar{X})$,

$\bar{x}'_\theta = \bar{X} + \theta(\bar{x}' - \bar{X})$, $\bar{z}'_\theta = \bar{Z} + \theta(\bar{z}' - \bar{Z})$ for $(0 < \theta < 1)$.

In the light of the conditions mentioned for $F(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}')$ in equation (9), it is noted that

$$F\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)=\bar{Y} \Rightarrow d_1 = 1 \text{ and } d_{11} = \frac{\partial^2}{\partial \bar{y}^{*2}} F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)\Bigg|_{\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)} = 0. \quad (19)$$

Since the population mean of the auxiliary variable x is unknown, we have to impose the constraint as

$$d_2 = - d_3 . \quad (20)$$

Thus, expressing $F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right)$ in terms of e's and neglecting the terms of e's having power greater than two we get

$$T_1 = F\left(\bar{y}^*, \bar{x}^*, \bar{x}', \bar{z}'\right) = \bar{Y}\left(1+ e_0\right) + d_2\bar{X}\left(e_1 - e_1'\right) + d_4\bar{Z}e_2' + \frac{1}{2}\left\{\begin{array}{l} d_{22}\bar{X}^2e_1^2 + d_{33}\bar{X}^2e_1'^2 + d_{44}\bar{Z}^2e_2'^2 + 2d_{12}\bar{Y}\bar{X}e_0e_1 + 2d_{13}\bar{Y}\bar{X}e_0e_1' \\ +2d_{14}\bar{Y}\bar{Z}e_0e_2' + 2d_{23}\bar{X}^2e_1e_1' + 2d_{24}\bar{X}\bar{Z}e_1e_2' + 2d_{34}\bar{X}\bar{Z}e_1'e_2' \end{array}\right\} \quad (21)$$

Similarly, expressing $T_i\left(i = 2, 3, 4\right)$ in terms of e's we have

$$T_2 = g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right) = \bar{Y}\left(1+ e_0\right) + c_2\bar{X}\left(e_1 - e_1'\right) + c_3\bar{Z}\left(e_2 - e_2'\right) \quad (22)$$

$$+\frac{1}{2}\left\{\begin{array}{l} c_{22}\bar{X}^2e_1^2 + c_{33}\bar{Z}^2e_2^2 + c_{44}\bar{X}^2e_1'^2 + c_{55}\bar{Z}^2e_2'^2 + 2c_{12}\bar{Y}\bar{X}e_0e_1 + 2c_{13}\bar{Y}\bar{Z}e_0e_2 \\ +2c_{14}\bar{Y}\bar{X}e_0e_1' + 2c_{15}\bar{Y}\bar{Z}e_0e_2' + 2c_{23}\bar{X}\bar{Z}e_0e_2 + 2c_{24}\bar{X}^2e_1e_1' \\ +2c_{25}\bar{X}\bar{Z}e_1e_2' + 2c_{34}\bar{X}\bar{Z}e_1'e_2 + 2c_{35}\bar{Z}^2e_2e_2' + 2c_{45}\bar{X}\bar{Z}e_1'e_2' \end{array}\right\}$$

$$T_3 = G\left(\bar{y}^*, \bar{x}, \bar{x}', \bar{z}'\right) = \bar{Y}\left(1+e_0\right) + p_2\bar{X}\left(e_3 - e_1'\right) + p_4\bar{Z}e_2' + \frac{1}{2}\left\{\begin{array}{l} p_{22}\bar{X}^2e_3^2 + p_{33}\bar{X}^2e_1'^2 + p_{44}\bar{Z}^2e_2'^2 + 2p_{12}\bar{Y}\bar{X}e_0e_3 + 2p_{13}\bar{Y}\bar{X}e_0e_1' \\ +2p_{14}\bar{Y}\bar{Z}e_0e_2' + 2p_{23}\bar{X}^2e_1'e_3 + 2p_{24}\bar{X}\bar{Z}e_2'e_3 + 2p_{34}\bar{X}\bar{Z}e_1'e_2' \end{array}\right\} \quad (23)$$

and

$$T_4 = \psi\left(\bar{y}^*, \bar{x}, \bar{z}^*, \bar{x}', \bar{z}'\right) = \bar{Y}\left(1+e_0\right) + q_2\bar{X}\left(e_3 - e_1'\right) + q_3\bar{Z}\left(e_2 - e_2'\right) + \frac{1}{2}\left\{\begin{array}{l} q_{22}\bar{X}^2e_1^2 + q_{33}\bar{Z}^2e_2^2 + q_{44}\bar{X}^2e_1'^2 + q_{55}\bar{Z}^2e_2'^2 + 2q_{12}\bar{Y}\bar{X}e_0e_3 \\ +2q_{14}\bar{Y}\bar{X}e_0e_1' + 2q_{15}\bar{Y}\bar{Z}e_0e_2' + 2q_{23}\bar{X}\bar{Z}e_2e_3 + 2q_{24}\bar{X}^2e_1'e_3 \\ +2q_{34}\bar{X}\bar{Z}e_1'e_2 + 2q_{35}\bar{Z}^2e_2e_2' + 2q_{45}\bar{X}\bar{Z}e_1'e_2' + 2q_{25}\bar{X}\bar{Z}e_2'e_3 + 2q_{13}\bar{Y}\bar{Z}e_0e_2 \end{array}\right\} \quad (24)$$

where

$$c_2 = \frac{\partial}{\partial \bar{x}^*} g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} = -\frac{\partial}{\partial \bar{x}'} g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} \left\{\text{as } \bar{X} \text{ is unknown}\right\},$$

$$c_3 = \frac{\partial}{\partial \bar{z}^*} g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} = -\frac{\partial}{\partial \bar{z}'} g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} \left\{\text{as } \bar{Z} \text{ is unknown}\right\},$$

$$p_2 = \frac{\partial}{\partial \bar{x}} G\left(\bar{y}^*, \bar{x}, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)} = -\frac{\partial}{\partial \bar{x}'} G\left(\bar{y}^*, \bar{x}, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)} \left\{\text{as } \bar{X} \text{ is unknown}\right\},$$

$$q_2 = \frac{\partial}{\partial \bar{x}} \psi\left(\bar{y}^*, \bar{x}, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} = -\frac{\partial}{\partial \bar{x}'} \psi\left(\bar{y}^*, \bar{x}, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} \left\{\text{as } \bar{X} \text{ is unknown}\right\}$$

$$q_3 = \frac{\partial}{\partial \bar{z}^*} \psi\left(\bar{y}^*, \bar{x}, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} = -\frac{\partial}{\partial \bar{z}'} \psi\left(\bar{y}^*, \bar{x}, \bar{z}^*, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)} \left\{\text{as } \bar{Z} \text{ is unknown}\right\}$$

$$p_4 = \frac{\partial}{\partial \bar{z}'} G\left(\bar{y}^*, \bar{x}, \bar{x}', \bar{z}'\right)\Big|_{\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)},$$

$\left(c_{22}, c_{33}, c_{44}, c_{55}, c_{12}, c_{13}, c_{14}, c_{15}, c_{23}, c_{24}, c_{25}, c_{34}, c_{35}, c_{45}\right)$ are the second order partial derivatives of $g\left(\bar{y}^*, \bar{x}^*, \bar{z}^*, \bar{x}', \bar{z}'\right)$ at the point $\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)$, $\left(p_{22}, p_{33}, p_{44}, p_{12}, p_{13}, p_{14}, p_{23}, p_{24}, p_{34}\right)$ are the second order partial derivatives of $G\left(\bar{y}^*, \bar{x}, \bar{x}', \bar{z}'\right)$ at the point $\left(\bar{Y}, \bar{X}, \bar{X}, \bar{Z}\right)$ and ($q_{22}, q_{33}, q_{44}, q_{55}, q_{12}, q_{13}, q_{14}, q_{15}, q_{23}, q_{24}, q_{25}, q_{34}, q_{35}, q_{45}$) are the second order partial derivatives of $\psi\left(\bar{y}^*, \bar{x}, \bar{z}^*, \bar{x}', \bar{z}'\right)$ at the point $\left(\bar{Y}, \bar{X}, \bar{Z}, \bar{X}, \bar{Z}\right)$.

Taking expectations on both sides of the equations (21)-(24) and using the results from equation (17), we obtain the expressions for bias B(.) and mean square errors M(.) of the classes of estimators $T_i (i = 1, 2, \ldots, 4)$ to the first order of approximations as

$$B\left(T_1\right) = \frac{1}{2}\begin{bmatrix} d_{22}\bar{X}^2 A' + d_{33}\bar{X}^2 f_2 C_x^2 + d_{44}\bar{Z}^2 f_2 C_z^2 + 2d_{12}\bar{Y}\bar{X}C' + 2d_{13}f_2\bar{Y}\bar{X}\rho_{yx}C_yC_x \\ 2d_{14}f_2\bar{Y}\bar{Z}\rho_{yz}C_yC_z + 2d_{23}\bar{X}^2 f_2 C_x^2 + 2d_{24}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2d_{34}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z \end{bmatrix}, \quad (25)$$

$$B(T_2)=\frac{1}{2}\begin{bmatrix} c_{22}\bar{X}^2A' + c_{33}\bar{Z}^2B' + c_{44}f_2\bar{X}^2C_x^2 + c_{55}f_2\bar{Z}^2C_z^2 + 2c_{12}\bar{Y}\bar{X}C' + 2c_{13}\bar{Y}\bar{Z}D' \\ +2c_{14}f_2\bar{Y}\bar{X}\rho_{yx}C_yC_x + 2c_{15}f_2\bar{Y}\bar{Z}\rho_{yz}C_yC_z + 2c_{23}\bar{X}\bar{Z}E' + 2c_{24}f_2\bar{X}^2C_x^2 \\ +2c_{25}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2c_{34}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2c_{35}f_2\bar{Z}^2C_z^2 \\ 2c_{45}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z \end{bmatrix}, \quad (26)$$

$$B(T_3)=\frac{1}{2}\begin{bmatrix} p_{22}f_1\bar{X}^2C_x^2 + p_{33}f_2\bar{X}^2C_x^2 + p_{44}f_2\bar{Z}^2C_z^2 + 2p_{12}f_1\bar{Y}\bar{X}\rho_{yx}C_yC_x + 2p_{13}f_2\bar{Y}\bar{X}\rho_{yx}C_yC_x \\ 2p_{14}f_2\bar{Y}\bar{Z}\rho_{yz}C_yC_z + 2p_{23}f_2\bar{X}^2C_x^2 + 2p_{24}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2p_{34}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z \end{bmatrix}, (27)$$

$$B(T_4)=\frac{1}{2}\begin{bmatrix} q_{22}\bar{X}^2f_1C_x^2 + q_{33}\bar{Z}^2B' + q_{44}f_2\bar{X}^2C_x^2 + q_{55}f_2\bar{Z}^2C_z^2 + 2q_{12}f_1\bar{Y}\bar{X}\rho_{yx}C_yC_x + 2q_{13}\bar{Y}\bar{Z}D' \\ +2q_{14}f_2\bar{Y}\bar{X}\rho_{yx}C_yC_x + 2q_{15}f_2\bar{Y}\bar{Z}\rho_{yz}C_yC_z + 2q_{23}f_1\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2q_{24}f_2\bar{X}^2C_x^2 \\ +2q_{25}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2q_{34}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z + 2q_{35}f_2\bar{Z}^2C_z^2 + 2q_{45}f_2\bar{X}\bar{Z}\rho_{xz}C_xC_z \end{bmatrix}, (28)$$

$$M(T_1)=\bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} + d_2^2\bar{X}^2A + d_4^2\bar{Z}^2f_2C_z^2 + 2d_2\bar{Y}\bar{X}C + 2d_4f_2\bar{Y}\bar{Z}\rho_{yz}C_yC_z, \quad (29)$$

$$M(T_2)=\bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} + c_2^2\bar{X}^2A + c_3^2\bar{Z}^2B + 2c_2\bar{Y}\bar{X}C + 2c_3\bar{Y}\bar{Z}D + 2c_2c_3\bar{X}\bar{Z}E, \quad (30)$$

$$M(T_3)=\bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} + p_2^2f_3\bar{X}^2C_x^2 + p_4^2f_2\bar{Z}^2C_z^2 + 2p_2f_3\bar{Y}\bar{X}\rho_{yx}C_yC_x + 2p_4f_2\bar{Y}\bar{Z}\rho_{yz}C_yC_z \quad (31)$$

and

$$M(T_4)=\bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} + q_2^2f_3\bar{X}^2C_x^2 + q_3^2\bar{Z}^2B + 2q_2f_3\bar{Y}\bar{X}\rho_{yx}C_yC_x + 2q_3\bar{Y}\bar{Z}D$$
$$+2q_2q_3\bar{X}\bar{Z}f_3\rho_{xz}C_xC_z \quad (32)$$

where

$A = f_3C_x^2 + W_2\frac{(k\text{-}1)}{n}C_{x(2)}^2$, $B = f_3C_z^2 + W_2\frac{(k\text{-}1)}{n}C_{z(2)}^2$, $C = f_3\rho_{yx}C_yC_x + W_2\frac{(k\text{-}1)}{n}\rho_{yx(2)}C_{y(2)}C_{x(2)}$,

$D = f_3\rho_{yz}C_yC_z + W_2\frac{(k\text{-}1)}{n}\rho_{yz(2)}C_{y(2)}C_{z(2)}$, $E = f_3\rho_{xz}C_xC_z + W_2\frac{(k\text{-}1)}{n}\rho_{xz(2)}C_{x(2)}C_{z(2)}$,

$A' = f_1C_x^2 + W_2\frac{(k\text{-}1)}{n}C_{x(2)}^2$, $B' = f_1C_z^2 + W_2\frac{(k\text{-}1)}{n}C_{z(2)}^2$, $C' = f_1\rho_{yx}C_yC_x + W_2\frac{(k\text{-}1)}{n}\rho_{yx(2)}C_{y(2)}C_{x(2)}$,

$D' = f_1\rho_{yz}C_yC_z + W_2\frac{(k\text{-}1)}{n}\rho_{yz(2)}C_{y(2)}C_{z(2)}$, $E' = f_1\rho_{xz}C_xC_z + W_2\frac{(k\text{-}1)}{n}\rho_{xz(2)}C_{x(2)}C_{z(2)}$.

Remark 3.1

The bias and mean square errors of the various estimators (indicated in section 2) belonging to the classes of estimators $T_i\left(i = 1, 2, ..., 4\right)$ can be easily obtained by substituting the suitable values of the derivatives in equations (25)-(32) as suggested by (Singh *et al.*, 2007; Singh and Vishwakarma, 2007).

4. Minimum M.S.E of the Proposed Classes of Estimators $T_i\left(i = 1, 2, ..., 4\right)$

It is obvious from the equations (29)-(32) and remark 3.1 that the mean square errors of the proposed classes of estimators $T_i\left(i = 1, 2, ..., 4\right)$ depend on the different values of the derivatives $d_2, d_4, c_2, c_3, p_2, p_4, q_2$ and q_3. Therefore, we desire to minimize the mean square errors of the classes of estimators T_i. The optimality conditions, that are, the conditions under which our proposed classes of estimators have minimum M. S. E.s are obtained as

$$\left.\begin{aligned} & d_2 = -\frac{\bar{Y}C}{\bar{X}A},\ d_4 = -\rho_{yz}\frac{\bar{Y}C_y}{\bar{Z}C_z},\ c_2 = \frac{\bar{Y}\left(ED\text{-}BC\right)}{\bar{X}\left(AB - E^2\right)},\ c_3 = \frac{\bar{Y}\left(EC\text{-}AD\right)}{\bar{Z}\left(AB- E^2\right)},\ p_2 = -\rho_{yx}\frac{\bar{Y}C_y}{\bar{X}C_x}, \\ & p_4 = -\rho_{yz}\frac{\bar{Y}C_y}{\bar{Z}C_z},\ q_2 = \frac{\bar{Y}\left(f_3\rho_{xz}C_xC_zD - f_3B\rho_{yx}C_yC_x\right)}{\bar{X}\left\{f_3BC_x^2 - \left(f_3\rho_{xz}C_xC_z\right)^2\right\}}\ \text{and}\ q_3 = \frac{\bar{Y}\left(f_3^2\rho_{xz}\rho_{yx}C_x^2C_yC_z - f_3C_x^2D\right)}{\bar{Z}\left\{f_3BC_x^2 - \left(f_3\rho_{xz}C_xC_z\right)^2\right\}}. \end{aligned}\right\} \quad (33)$$

Substituting these optimum values of the derivatives in equations (29)-(32), we have minimum M. S. E.s of the classes of estimators $T_i\left(i = 1, 2, ..., 4\right)$ as

$$\text{Min. } M\left(T_1\right) = \bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} - f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 - \bar{Y}^2\frac{C^2}{A}, \quad (34)$$

$$\text{Min. } M\left(T_2\right) = \bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} - \bar{Y}^2\frac{AD^2 + BC^2 - 2EDC}{AB - E^2}, \quad (35)$$

$$\text{Min. } M\left(T_3\right) = \bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} - f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 - f_3\left(\rho_{yx}\bar{Y}C_y\right)^2 \quad (36)$$

and

$$\text{Min. } M(T_4) = \bar{Y}^2\left\{f_1C_y^2 + W_2\frac{(k\text{-}1)}{n}C_{y(2)}^2\right\} - \bar{Y}^2\frac{f_3C_x^2D^2 + B\left(f_3\rho_{yx}C_yC_x\right)^2 - 2Df_3^2\rho_{yx}\rho_{xz}C_x^2C_yC_z}{f_3BC_x^2 - \left(f_3\rho_{xz}C_xC_z\right)^2}. \quad (37)$$

Remark 4.1

It is to be noted from the optimality conditions in equation (33) that optimum values of derivatives of the proposed classes of estimators $T_i\left(i = 1, 2, \ldots, 4\right)$ depend on unknown population parameters such as $C_y, C_x, C_z, \rho_{yx}, \rho_{xz}, \rho_{yz}, \rho_{yx(2)}, \rho_{xz(2)}, \rho_{yz(2)}, C_{y(2)}, C_{x(2)}$ and $C_{z(2)}$. Thus, to use such estimators one has to use guessed or estimated values of these parameters. Guessed values of these population parameters can be obtained either from past data or experience gathered over time; see (Murthy, 1967; Reddy, 1978; Tracy *et al.,* 1996). If such guessed values are not available then it is advisable to use sample data to estimate these parameters as suggested by (Gupta and Shabbir, 2008). In case, non-response situations occur in the sample data, it is advised to utilize the technique of sub-sampling of the non-responding group to estimate these parameters as suggested in this paper. After substitutions of the above population parameters with their respective estimated values, it could be observed that the mean square errors of the resulting estimators are same (up to first order of approximations) to those derived.

5. Efficiency Comparisons of the Proposed Classes of Estimators $T_i\left(i = 1, 2, \ldots, 4\right)$

It is important to investigate the situations under which our proposed classes of estimators $T_i\left(i = 1, 2, \ldots, 4\right)$ are preferable over the existing estimators such as (Hansen and Hurwitz, 1946) sample mean estimator $\bar{y}^*$ and $t_i (i = 1, 2, \ldots, 6)$. The variance V(.)/ M. S. E.s of the estimators t_i to the first order of approximations are obtained as

$$V\left(\bar{y}^*\right) = \bar{Y}^2 \left\{ f_1 C_y^2 + W_2 \frac{(k-1)}{n} C_{y(2)}^2 \right\} \tag{38}$$

$$M(t_1) = f_2 \bar{Y}^2 C_y^2 + f_3 \bar{Y}^2 C_r^2 + W_2 \frac{(k-1)}{n} \bar{Y}^2 C_{r(2)}^2 \tag{39}$$

$$M(t_2) = f_2 \bar{Y}^2 C_y^2 + f_3 \bar{Y}^2 C_p^2 + W_2 \frac{(k-1)}{n} \bar{Y}^2 C_{p(2)}^2 \tag{40}$$

$$M(t_3) = f_2 \bar{Y}^2 C_y^2 + f_3 \bar{Y}^2 C_y^2 \left(1 - \rho_{yx}^2\right) + W_2 \frac{(k-1)}{n} \left\{ S_{y(2)}^2 + \beta_{yx} S_{x(2)}^2 \left(\beta_{yx} - 2\beta_{yx(2)}\right) \right\} \tag{41}$$

$$M(t_4) = f_2 \bar{Y}^2 C_y^2 + f_3 \bar{Y}^2 C_r^2 + W_2 \frac{(k-1)}{n} \bar{Y}^2 C_{y(2)}^2 \tag{42}$$

$$M(t_5) = f_2\bar{Y}^2C_y^2 + f_3\bar{Y}^2C_p^2 + W_2\frac{(k-1)}{n}\bar{Y}^2C_{y(2)}^2 \quad (43)$$

and

$$M(t_6) = f_2\bar{Y}^2C_y^2 + f_3\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right) + W_2\frac{(k-1)}{n}\bar{Y}^2C_{y(2)}^2. \quad (44)$$

where

$C_r^2 = C_y^2 + C_x^2 - 2\rho_{yx}C_yC_x$, $C_p^2 = C_y^2 + C_x^2 + 2\rho_{yx}C_yC_x$, $C_{r(2)}^2 = C_{y(2)}^2 + C_{x(2)}^2 - 2\rho_{yx(2)}C_{y(2)}C_{x(2)}$,

$C_{p(2)}^2 = C_{y(2)}^2 + C_{x(2)}^2 + 2\rho_{yx(2)}C_{y(2)}C_{x(2)}$, $\beta_{yx(2)} = \frac{S_{yx(2)}}{S_{x(2)}^2}$ and $S_{yx(2)} = (N_2-1)^{-1}\sum_{i=1}^{N_2}\left(y_i - \bar{Y}_2\right)\left(x_i - \bar{X}_2\right)$.

5.1. Efficiency Comparisons of the Estimators T_1 and T_2

In this section, we compare the efficiencies of our proposed classes of estimators T_1 and T_2 under their respective optimality conditions with existing estimators when non- response situations is observed on the study variable y as well as on the auxiliary variable x in the second phase sample of size n.

(a) Efficiency Comparisons of T_1

It could be concluded from equations (34) and (38)-(41) that

(i). T_1 is more efficient than $\bar{y}^*$, provided $V\left(\bar{y}^*\right) - \text{Min. } M\left(T_1\right) > 0.$

$$\Rightarrow f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 + \bar{Y}^2\frac{C^2}{A} > 0. \quad (45)$$

It may be observed from the equation (45) that the class of estimators T_1 is always preferable over $\bar{y}^*$ as $\left(f_1, f_2, f_3 > 0\right)$ and $A = f_3C_x^2 + W_2\frac{(k-1)}{n}C_{x(2)}^2 > 0$ always.

Similarly,

(ii) T_1 is more precise than t_1 when $M(t_1) - \text{Min. } M\left(T_1\right) > 0$

$$\Rightarrow f_3\bar{Y}^2\left(C_x^2 - 2\rho_{yx}C_yC_x\right) + W_2\frac{(k-1)}{n}\bar{Y}^2\left(C_{x(2)}^2 - 2\rho_{yx(2)}C_{y(2)}C_{x(2)}\right)$$
$$+ f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 + \bar{Y}^2\frac{C^2}{A} > 0. \quad (46)$$

which is possible if

$$C_x^2 \geq 2\rho_{yx}C_yC_x \text{ and } C_{x(2)}^2 \geq 2\rho_{yx(2)}C_{y(2)}C_{x(2)}$$
$$\Rightarrow \rho_{yx} \leq \frac{C_x}{2C_y} \text{ and } \rho_{yx(2)} \leq \frac{C_{x(2)}}{2C_{y(2)}}. \quad (47)$$

(iii) T_1 is more efficient than t_2, if $M(t_2)$ - Min. $M(T_1) > 0$

$$\Rightarrow f_3\bar{Y}^2\left(C_x^2+2\rho_{yx}C_yC_x\right)+W_2\frac{(k-1)}{n}\bar{Y}^2\left(C_{x(2)}^2+2\rho_{yx(2)}C_{y(2)}C_{x(2)}\right)$$
$$+f_2\left(\bar{Y}\rho_{yz}C_y\right)^2+\bar{Y}^2\frac{C^2}{A}>0. \quad (48)$$

It may be noted from equation (48) that T_1 is always more precise than t_2 when

$C_x^2 + 2\rho_{yx}C_yC_x \geq 0$ and $C_{x(2)}^2+2\rho_{yx(2)}C_{y(2)}C_{x(2)} \geq 0$ which is possible provided

$$\rho_{yx} \geq -\frac{C_x}{2C_y} \text{ and } \rho_{yx(2)} \geq -\frac{C_{x(2)}}{2C_{y(2)}}. \quad (49)$$

(iv) T_1 is preferable over t_3 when $M(t_3)$ - Min. $M(T_1) > 0$

$$\Rightarrow W_2\frac{(k-1)}{n}\beta_{yx}S_{x(2)}^2\left(\beta_{yx}-2\beta_{yx(2)}\right)+\bar{Y}^2C_y^2\left(f_2\rho_{yz}^2-f_3\rho_{yx}^2\right)+\bar{Y}^2\frac{C^2}{A}>0 \quad (50)$$

which is always possible provided

$$\beta_{yx} > 2\beta_{yx(2)} \text{ and } f_2\rho_{yz}^2 > f_3\rho_{yx}^2. \quad (51)$$

(b) Efficiency Comparisons of T_2

It can be observed from equations that (35), and (38) - (41) that

(i). T_2 is more efficient than $\bar{y}^*$, provided $V\left(\bar{y}^*\right)$- Min. $M\left(T_2\right) > 0$

$$\Rightarrow \bar{Y}^2\frac{AD^2+BC^2-2EDC}{AB-E^2} \geq 0 \quad (52)$$

which occurs always since $\left(-1 \leq \rho_{xz} \leq 1, -1 \leq \rho_{xz(2)} \leq 1\right)$

$\Rightarrow AD^2 + BC^2 - 2EDC$

$$\Rightarrow f_3\left(C_x^2D^2+C_z^2C^2-2DC\rho_{xz}C_xC_z\right)+W_2\frac{(k-1)}{n}\left(C_{x(2)}^2D^2+C_{z(2)}^2C^2-2DC\rho_{xz(2)}C_{x(2)}C_{z(2)}\right) \geq 0 \quad (53)$$

and

$$\Rightarrow AB - E^2$$

$$\Rightarrow f_3^2 C_x^2 C_z^2 \left(1 - \rho_{xz}^2\right) + \left\{W_2 \frac{(k-1)}{n}\right\}^2 C_{x(2)}^2 C_{z(2)}^2 \left(1 - \rho_{xz(2)}^2\right)$$

$$+ f_3 W_2 \frac{(k-1)}{n}\left(C_x^2 C_z^2 + C_{x(2)}^2 C_{z(2)}^2 - 2\rho_{xz}\rho_{xz(2)} C_x C_z C_{x(2)} C_{z(2)}\right) \geq 0 \qquad (54)$$

Similarly,

(ii) T_2 is more precise than t_1 when $M(t_1) - \text{Min. } M(T_2) > 0$

$$\Rightarrow f_3 \bar{Y}^2 \left(C_x^2 - 2\rho_{yx} C_y C_x\right) + W_2 \frac{(k-1)}{n} \bar{Y}^2 \left(C_{x(2)}^2 - 2\rho_{yx(2)} C_{y(2)} C_{x(2)}\right)$$

$$+ \bar{Y}^2 \frac{AD^2 + BC^2 - 2EDC}{AB - E^2} > 0. \qquad (55)$$

which is possible when

$$C_x^2 \geq 2\rho_{yx} C_y C_x \text{ and } C_{x(2)}^2 \geq 2\rho_{yx(2)} C_{y(2)} C_{x(2)}$$

$$\Rightarrow \rho_{yx} \leq \frac{C_x}{2C_y} \text{ and } \rho_{yx(2)} \leq \frac{C_{x(2)}}{2C_{y(2)}}. \qquad (56)$$

(iii) T_2 is more efficient than t_2 if $M(t_2) - \text{Min. } M(T_2) > 0$

$$\Rightarrow f_3 \bar{Y}^2 \left(C_x^2 + 2\rho_{yx} C_y C_x\right) + W_2 \frac{(k-1)}{n} \bar{Y}^2 \left(C_{x(2)}^2 + 2\rho_{yx(2)} C_{y(2)} C_{x(2)}\right) + \bar{Y}^2 \frac{AD^2 + BC^2 - 2EDC}{AB - E^2} > 0. \qquad (57)$$

which is always possible when

$$C_x^2 + 2\rho_{yx} C_y C_x \geq 0 \text{ and } C_{x(2)}^2 + 2\rho_{yx(2)} C_{y(2)} C_{x(2)} \geq 0.$$

$$\Rightarrow \rho_{yx} \geq -\frac{C_x}{2C_y} \text{ and } \rho_{yx(2)} \geq -\frac{C_{x(2)}}{2C_{y(2)}}. \qquad (58)$$

(iv) T_2 is preferable over t_3 when

$$M(t_3) - \text{Min. } M(T_2) > 0$$

$$\Rightarrow W_2 \frac{(k-1)}{n} \beta_{yx} S_{x(2)}^2 \left(\beta_{yx} - 2\beta_{yx(2)}\right) - \bar{Y}^2 f_3 \rho_{yx}^2 C_y^2 + \bar{Y}^2 \frac{AD^2 + BC^2 - 2EDC}{AB - E^2} > 0 \qquad (59)$$

which is possible provided

$$\beta_{yx} > 2\beta_{yx(2)} \text{ and } \frac{AD^2 + BC^2 - 2EDC}{AB - E^2} > f_3 \rho_{yx}^2 C_y^2. \qquad (60)$$

5.2 Efficiency Comparisons of the Estimators T_3 and T_4

Now, we compare the efficiencies of the classes of estimators T_3 and T_4 under their respective optimality conditions with the estimators $\bar{y}^*$, t_4, t_5 and t_6 when there is non-response only on the study variable y but the complete information on the auxiliary variable x is available from the second phase sample of size n.

(a) Efficiency Comparisons of T_3

(i) It may be observed from equations (36) and (38) that the class of estimators T_3 is preferable over $\bar{y}^*$ when $V(\bar{y}^*) - \text{Min. } M(T_3) > 0$

on simplification which turns out to be

$$f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 + f_3\left(\rho_{yx}\bar{Y}C_y\right)^2 > 0. \tag{61}$$

It may be noted that class of estimators T_3 is always more efficient than $\bar{y}^*$.

(ii) From equations (36) and (42), it may be detected that the class of estimators T_3 is more precise than t_4, provided $M(t_4) - \text{Min. } M(T_3) > 0$

$$\Rightarrow f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 + f_3\left(\rho_{yx}\bar{Y}C_y\right)^2 + f_3\bar{Y}^2\left(C_x^2 - 2\rho_{yx}C_yC_x\right) > 0. \tag{62}$$

which is possible when

$$\rho_{yx} \le \frac{C_x}{2C_y} \tag{63}$$

(iii) It may be seen from the equations (36) and (43) that the class of estimators T_3 is more efficient than t_5, provided $M(t_5) - \text{Min. } M(T_3) > 0$

$$\Rightarrow f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 + f_3\left(\rho_{yx}\bar{Y}C_y\right)^2 + f_3\bar{Y}^2\left(C_x^2 + 2\rho_{yx}C_yC_x\right) > 0. \tag{64}$$

which is possible if

$$\rho_{yx} \ge -\frac{C_x}{2C_y} \tag{65}$$

(iv) A comparison of efficiencies of the class of estimators T_3 with the estimator t_6 from equations (36) and (44) revels the fact that T_3 is superior than t_6 when

$$M(t_6) - \text{Min. } M(T_3) > 0.$$

$$\Rightarrow f_2\left(\bar{Y}\rho_{yz}C_y\right)^2 > 0 \tag{66}$$

which claims that the class of estimators T_3 is always more efficient than the estimator t_6.

(b) Efficiency Comparisons of T_4

Comparisons of efficiencies of the class of estimators T_4 with the estimators $\bar{y}^*$, indicate the fact that the class of estimators T_4 is preferable over $\bar{y}^*$, provided $V(\bar{y}^*) - \text{Min. } M(T_4) > 0$

$$\Rightarrow \bar{Y}^2 \frac{PD^2 + BQ^2 - 2(f_3\rho_{xz}C_xC_z)DQ}{PB - (f_3\rho_{xz}C_xC_z)^2} \geq 0. \tag{67}$$

Thus, it may be established from the above comparisons that the class of estimators T_4 is always more efficient than the estimator $\bar{y}^*$ since $(-1 \leq \rho_{xz} \leq 1)$ indicates that

$$PD^2 + BQ^2 - 2(f_3\rho_{xz}C_xC_z)DQ = f_3\left(C_x^2D^2 + C_z^2Q^2 - 2DQ\rho_{xz}C_xC_z\right) + W_2\frac{(k-1)}{n}C_{z(2)}^2Q^2 \geq 0 \tag{68}$$

and

$$PB - (f_3\rho_{xz}C_xC_z)^2 = f_3^2C_x^2C_z^2\left(1 - \rho_{xz}^2\right) + f_3W_2\frac{(k-1)}{n}C_x^2C_{z(2)}^2 \geq 0. \tag{69}$$

where $P = f_3C_x^2$ and $Q = f_3\rho_{yx}C_yC_x$.

Similarly, from equations (37) and (42) - (44), it may be noted that:

(ii) T_4 is more precise than t_4 when $M(t_4) - \text{Min. } M(T_4) > 0$

$$\Rightarrow f_3\bar{Y}^2\left(C_x^2 - 2\rho_{yx}C_yC_x\right) + \bar{Y}^2\frac{PD^2 + BQ^2 - 2(f_3\rho_{xz}C_xC_z)DQ}{PB - (f_3\rho_{xz}C_xC_z)^2} > 0. \tag{70}$$

which is possible when

$$\rho_{yx} \leq \frac{C_x}{2C_y} \tag{71}$$

(iii) T_4 is more preferable over t_5 when $M(t_5) - \text{Min. } M(T_4) > 0$

$$\Rightarrow f_3\bar{Y}^2\left(C_x^2 + 2\rho_{yx}C_yC_x\right) + \bar{Y}^2\frac{PD^2 + BQ^2 - 2(f_3\rho_{xz}C_xC_z)DQ}{PB - (f_3\rho_{xz}C_xC_z)^2} > 0. \tag{72}$$

which gives the dominance of situation the class of estimators T_4 over the estimator t_5 when

$$\rho_{yx} \geq -\frac{C_x}{2C_y} \tag{73}$$

(iv) T_4 is preferable over t_6 when $M(t_6) - \text{Min. } M(T_4) > 0$

$$\Rightarrow \frac{PD^2 + BQ^2 - 2\left(f_3\rho_{xz}C_xC_z\right)DQ}{PB - \left(f_3\rho_{xz}C_xC_z\right)^2} > f_3\rho_{yx}^2C_y^2 \quad (74)$$

It may be noted from equations (60) and (74) that the dominance of the class of estimators T_2 over t_3 and T_4 over t_6 are difficult to establish theoretically. However, their performances are examined below through empirical studies carried over different population which establish their superiority over the traditional ones.

6. Numerical Illustration

We have chosen three natural population data sets to illustrate the efficacious performances of our proposed classes of estimators. The source of the populations, the nature of the variables y, x, z and the values of the various parameters are given as follows.

Population I- Source: (Khare and Sinha, 2007)

The present data belongs to the physical growth of upper socio-economic group of 95 school going children of Varanasi under an ICMR study, Department of Pediatrics, B. H. U., during 1983-84. The first 25% (i.e. 24 children) units have been considered as non-response group of the population. The weight (in kg) of the children is taken as study variable (y) while the skull circumference (in cm) of the children and the chest circumference (in cm) of the children are taken as auxiliary variable x and z respectively. It is to be noted that this population was also considered by several authors including (Singh and Kumar, 2010 b).

Population II- Source: District Census Handbook, 1981, Orissa, published by Govt. of India

The 109 Village / Town / Ward wise population of urban area under Police-station – Baria, Tahasil–Champua, Orissa, India has been taken under study. The last 25% villages (i. e. 27 villages) have been considered as non-response group of the population. The number of literate persons in the village is considered as study variable (y) while the number of main workers in the village and the number of non -workers in the village are considered as auxiliary variable x and z respectively. This population was also considered as numerical evidence in the works of several authors including (Khare and Sinha, 2012).

Table1: Parametric values of the above populations.

Population	W_2	$\bar{Y}$	$\bar{X}$	$\bar{Z}$	C_y	C_x	C_z	$C_{y(2)}$
I **Size:** N= 95	0.25	19.49	51.17	55.86	0.15	0.03	0.05	0.12
	$C_{x(2)}$	$C_{z(2)}$	ρ_{yx}	ρ_{yz}	ρ_{xz}	$\rho_{yx(2)}$	$\rho_{yz(2)}$	$\rho_{xz(2)}$
	0.02	0.05	0.32	0.84	0.29	0.47	0.72	0.57
II **Size:** N= 109	W_2	$\bar{Y}$	$\bar{X}$	$\bar{Z}$	C_y	C_x	C_z	$C_{y(2)}$
	0.25	145.30	165.26	259.08	0.76	0.68	0.76	0.68
	$C_{x(2)}$	$C_{z(2)}$	ρ_{yx}	ρ_{yz}	ρ_{xz}	$\rho_{yx(2)}$	$\rho_{yz(2)}$	$\rho_{xz(2)}$
	0.57	0.54	0.81	0.90	0.81	0.78	0.87	0.74
III **Size:** N= 96	W_2	$\bar{Y}$	$\bar{X}$	$\bar{Z}$	C_y	C_x	C_z	$C_{y(2)}$
	0.25	137.92	144.87	185.21	1.32	0.81	1.05	2.08
	$C_{x(2)}$	$C_{z(2)}$	ρ_{yx}	ρ_{yz}	ρ_{xz}	$\rho_{yx(2)}$	$\rho_{yz(2)}$	$\rho_{xz(2)}$
	0.94	1.48	0.77	0.78	0.81	0.72	0.78	0.72

Population III- Source: District Census Handbook, 1981, West Bengal, published by Govt. of India

Ninety-six village wise population of rural area under Police-station–Singur, District – Hooghly, West Bengal has been taken under the study. The 25% villages (i.e. 24 villages) whose area is greater than 160 hectares have been considered as non-response group of the population. The number of agricultural labours in the village is taken as study variable (y) while the area (in hectares) of the village and the number of cultivators in the village are taken as auxiliary variables x and z respectively. It is to be noted that this population was also considered by (Khare and Sinha, 2009).

The values of various parameters obtained from the above populations are shown in Table 1.

To have a tangible idea about the performance of the proposed classes of estimators T_i (i= 1, 2, . . ., 4), we have computed the percent relative efficiencies (PREs) of the estimators T_i (under their respective optimality conditions) and the other existing estimators considered in this present work with respect to $\bar{y}^*$. The findings are displayed in Table 2 and Table 3 where we have designated the percent relative efficiency (PRE) of an estimator T with respect to sample mean estimator $\bar{y}^*$ as

$$PRE = \frac{V(\bar{y}^*)}{M(T)} \times 100 \,. \tag{75}$$

Table 2: PREs of the different estimators with respect to $\bar{y}^*$ when there is non-response situation on the study variable y as well as on the auxiliary variable x in second phase sample.

Population I											
n	K	t_1	t_2	t_3	T_1	T_2	t_1	t_2	t_3	T_1	T_2
		$n' = 65$					$n' = 75$				
30	2	109.3	87.0	109.7	130.4	222.8	110.1	86.1	110.7	123.1	252.6
	3	110.5	86.0	110.0	130.0	219.1	111.2	85.3	110.8	123.9	242.8
	4	111.5	85.3	110.3	129.8	217.2	112.1	84.6	111.0	124.5	237.0
35	2	108.9	87.5	109.1	135.2	207.3	109.9	86.4	110.3	125.7	240.1
	3	110.3	86.4	109.6	134.0	206.5	111.1	85.4	110.6	126.1	232.8
	4	111.3	85.5	109.9	133.3	206.6	112.0	84.7	110.8	126.4	228.7
40	2	108.5	88.1	108.5	141.2	192.2	109.7	86.6	110.0	128.7	227.4
	3	110.0	86.7	109.1	138.9	194.1	111.0	85.6	110.3	128.7	222.6
	4	111.1	85.8	109.5	137.3	196.1	112.0	84.8	110.5	128.7	220.2
Population II											
n	K	t_1	t_2	t_3	T_1	T_2	t_1	t_2	t_3	T_1	T_2
		$n' = 50$					$n' = 60$				
25	2	187.8	38.8	192.3	331.9	245.8	210.1	36.1	216.0	318.0	298.2
	3	197.3	37.7	204.7	317.4	268.6	217.4	35.6	226.7	306.8	319.2
	4	204.7	36.9	214.5	308.0	288.0	222.9	35.2	234.8	299.4	336.6
30	2	172.2	41.5	176.0	343.9	213.9	196.3	37.7	201.5	325.4	264.9
	3	183.4	39.7	189.9	325.8	237.9	205.5	36.8	213.9	312.0	288.4
	4	192.1	38.5	200.9	314.3	258.4	212.3	36.2	223.4	303.4	307.9
35	2	157.8	44.8	161.0	358.3	187.2	183.1	39.6	187.6	334.0	235.9
	3	170.4	42.1	176.1	335.5	211.8	193.9	38.2	201.5	317.9	261.1
	4	180.2	40.3	188.1	321.4	233.0	202.1	37.3	212.2	307.7	282.2
Population III											
n	K	t_1	t_2	t_3	T_1	T_2	t_1	t_2	t_3	T_1	T_2
		$n' = 50$					$n' = 60$				
30	2	163.4	55.1	180.6	227.8	206.6	176.0	52.2	196.9	226.6	230.7
	3	168.9	54.7	192.9	221.0	229.4	177.8	52.7	205.1	220.1	248.8
	4	171.9	54.5	199.8	218.0	243.5	178.8	53.0	209.7	217.2	259.6
35	2	155.9	57.4	171.6	227.8	194.2	170.0	53.6	189.7	226.3	220.4
	3	163.6	56.2	186.2	221.1	219.8	173.7	53.7	200.1	219.9	241.3
	4	167.8	55.6	194.6	218.0	235.7	175.7	53.7	205.8	217.1	253.7
40	2	148.5	60.1	162.8	228.0	182.4	164.0	55.2	182.4	226.1	210.2
	3	158.4	57.8	179.7	221.3	210.6	169.6	54.7	195.0	219.7	234.0
	4	163.75	56.78	189.44	218.26	228.15	172.57	54.54	201.91	217.02	248.0

Table 3: PREs of the different estimators with respect to $\bar{y}^*$ when the non-response situation is observed only on the study variable y in second phase sample while complete response is available on the auxiliary variable x.

Population I											
N	K	t_4	t_5	t_6	T_3	T_4	t_4	t_5	t_6	T_3	T_4
		$n' = 65$					$n' = 75$				
30	2	106.1	90.4	107.4	124.1	224.1	106.8	89.4	108.3	117.6	254.4
	3	105.1	91.7	106.2	119.7	220.7	105.7	90.9	107.0	114.5	244.8
	4	104.4	92.7	105.3	116.7	218.8	104.9	92.0	106.0	112.3	238.9
35	2	105.5	91.2	106.7	128.1	208.5	106.4	89.9	107.9	119.5	241.8
	3	104.6	92.5	105.6	122.5	207.8	105.3	91.4	106.5	115.9	234.7
	4	103.9	93.4	104.8	118.8	207.9	104.6	92.5	105.6	113.4	230.4
40	2	104.9	92.0	106.0	132.9	193.2	106.0	90.5	107.4	121.8	228.9
	3	104.0	93.3	104.9	125.8	195.2	104.9	92.0	106.0	117.4	224.3
	4	103.4	94.2	104.2	121.3	197.1	104.2	93.0	105.1	114.5	221.7
Population II											
N	K	t_4	t_5	t_6	T_3	T_4	t_4	t_5	t_6	T_3	T_4
		$n' = 50$					$n' = 60$				
25	2	151.3	46.4	152.0	232.5	236.8	165.5	42.6	166.4	225.6	284.5
	3	139.0	51.1	139.5	189.3	254.0	148.7	47.2	149.3	185.4	298.0
	4	131.5	55.0	131.8	167.3	268.9	138.7	51.2	139.2	164.7	309.8
30	2	139.8	50.7	140.3	234.8	207.1	155.2	45.2	156.0	226.0	253.9
	3	130.5	55.6	130.8	189.1	226.6	141.2	50.1	141.7	184.3	271.0
	4	124.7	59.7	124.9	166.5	243.2	132.9	54.2	133.3	163.4	285.5
35	2	128.9	56.6	129.3	237.3	182.1	145.3	48.4	145.9	226.4	227.1
	3	122.3	61.6	122.5	188.9	203.0	134.0	53.6	134.4	183.2	246.9
	4	118.2	65.6	118.3	165.7	220.8	127.2	57.8	127.5	161.9	263.4
Population III											
N	K	t_4	t_5	t_6	T_3	T_4	t_4	t_5	t_6	T_3	T_4
		$n' = 50$					$n' = 60$				
30	2	121.2	71.1	122.3	146.7	194.0	128.0	66.3	129.6	146.4	214.2
	3	113.4	78.4	114.1	127.5	211.3	117.4	74.4	118.3	127.4	226.9
	4	109.8	82.7	110.3	119.5	221.7	112.6	79.3	113.2	119.4	234.4
35	2	115.8	75.9	116.6	144.4	183.3	123.4	69.4	124.7	144.0	205.4
	3	110.0	82.5	110.5	125.9	203.3	114.5	77.2	115.3	125.7	220.8
	4	107.3	86.2	107.7	118.3	215.3	110.5	81.8	111.0	118.1	229.7
40	2	110.5	81.9	111.0	142.0	173.0	118.8	73.1	119.8	141.6	196.6
	3	106.6	87.3	107.0	124.2	195.6	111.6	80.5	112.2	124.0	214.6
	4	104.9	90.22	105.14	117.07	209.13	108.46	84.70	108.87	116.92	225.01

7. Conclusions

The following conclusions can be read-out from the present study.

(a) It is observed from the efficiency comparisons of our suggested classes of estimators $T_i \left(i = 1, 2, \ldots, 4\right)$ in section 5 that under their respective optimality conditions they are always preferable over the (Hansen and Hurwitz, 1946) sample mean estimator $\overline{y}^*$. The dominance conditions of the proposed classes of estimators over the existing estimators $t_i\left(i = 1, 2, \ldots, 6\right)$ are also shown in section 5.

(b) From Table 2 and Table 3, it may be observed that under the similar realistic situations, proposed classes of estimators $T_i\left(i = 1, 2, \ldots, 4\right)$ yield impressive gains in efficiencies over the existing estimators $t_i(i = 1, 2, \ldots, 6)$ for all different choices of the sub-sampling fraction $\left(\frac{1}{k}\right)$ and sample sizes $(n,\ n')$.

Therefore, it is clear that our proposed classes of estimators are more justifiable in compare with the previous work of similar nature and may be recommended for their practical applications.

Acknowledgements

Authors are thankful to the reviewers for their valuable suggestions.

References

1. Cochran, W.G. (1977): Sampling Techniques, (3rd ed.). *John Wiley and Sons, NewYork.*
2. Gupta, S. and Shabbir, J. (2008): On improvement in estimating the population mean in simple random sampling. *Journal of Applied Statistics, 35 (5)*: 559–566. DOI: 10.1080/02664760701835839
3. Hansen, M. H. and Hurwitz, W.N. (1946): The problem of non-response in sample surveys. *Journal of American Statistical Association, 41*: 517–529. DOI: 10.1080/01621459.1946.10501894
4. Khare, B. B. and Srivastava, S. R. (1993): Estimation of population mean using auxiliary character in presence of non-response. *National Academy of Science Letters. India 16*: 111–114.
5. Khare, B. B. and Srivastava, S. R. (1995): Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non- response. *Proceedings of National Academy of Science India, 65(A) (II)*: 195–203.
6. Khare, B. B. and Srivastava, S. R. (1997): Transformed ratio type estimators for the population mean in the presence of non-response. *Communication in Statistics- Theory Methods, 26*: 1779–1791. DOI: 10.1080/03610929708832012

7. 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. In: Statistical Techniques in Life Testing, Reliability, *Sampling Theory and Quality Control, 1*: 63–171.

8. Khare, B. B. and Sinha, R. R. (2009): On class of estimators for population mean using multi-auxiliary characters in the presence of non-response. *Statistics in Transition –new series, 10(1)*: 3-14.

9. Khare, B. B. and Sinha, R. R.(2012): Improved classes of estimators for ratio of two means with double sampling the non respondents. *Statistica, 49(3)*: 75-83.

10. Khare, B. B. and Sinha, R. R.(2012): Improved classes of estimators for ratio of two means with double sampling the non respondents. *Statistica, 49(3)*: 75-83.

11. Mohanty, S. (1967): Combination of regression and ratio estimates. *Journal of Indian Statistical Association, 5* :1–14.

12. Murthy, M. N. (1967): Sampling Theory and Methods. *Statistical Publishing Society, Calcutta, India.*

13. Olkin, I. (1958): Multivariate ratio estimation for finite population. *Biometrika, 45:* 154–165. DOI: 10.1093/biomet/45.1-2.154.

14. Okafor, F. C. and Lee, H. (2000): Double sampling for ratio and regression estimation with subsampling the non-respondent. *Survey Methodology, 26(2)*: 183–188.

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

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

17. Srivastava, S. K. (1971): A generalized estimator for the mean of a finite population using multi auxiliary information. *Journal of American Statistical Association, 66*: 404–407. DOI: 10.1080/01621459.1971.10482277

18. Singh, H. P. Chandra, P., Joarder, A. H. and Singh, S. (2007): Family of estimators of mean, ratio and product of a finite population using random nonresponse. *Test*, 16: 565– 597. DOI: 10.1007/s11749-006-0020-z

19. Singh, H. P. and Vishwakarma, G. K. (2007): A general class of estimators in successive sampling. *Metron, LXV (2):* 201–227.

20. Singh, G. N. and Priyanka, K. (2007): Effect of non-response on current

occasion in search of good rotation patterns on successive occasions. *Statistics in Transition- new series, 8 (2)*: 273 - 292.

21. Singh, H. P. and Kumar, K. (2010 a): Estimation of mean in presence of non-response using two phase sampling scheme. *Statistical Papers, 51*: 559–582. DOI: 10.1007/s00362-008-0140-5
22. Singh, H. P. and Kumar, K. (2010 b): Improved estimation of population mean under two-phase sampling with sub sampling the non-respondents. *Journal of Statistical Planning and Inference,* 140: 2536–2550. DOI: 10.1016/j.jspi.2010.03.023
23. Tracy, D. S., Singh, H. P. and Singh, R. (1996): An alternative to the ratio-cum-product estimator in sample surveys. *Journal of Statistical Planning and Inference, 53*: 375- 387. DOI: 10.1016/0378-3758(95)00136-0
24. Tripathi, T. P. and Khare, B. B. (1997): Estimation of mean vector in presence of non-response. *Communications in Statistics - Theory and Methods, 26 (9)*: 2255–2269. DOI: 10.1080/03610929708832045
25. 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.
26. 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.

Appendix:

Proof/ explanations of the expectations on e's presented in equation (17)

It may be noted from sections (2) - (3) that $\bar{y}^* = \dfrac{n_1\bar{y}_1 + n_2\bar{y}_{2m}}{n}$ and it is written in form of transformation as $\bar{y}^* = \bar{Y}(1+e_0)$ when $E(e_0) = 0$ and $|e_0| < 1$.

Thus, taking expectation we have

$$E\left(e_0^2\right) = \frac{E\left(\bar{y}^* - \bar{Y}\right)^2}{\bar{Y}^2} = \frac{V\left\{E\left(\bar{y}^* | n_1, n_2\right)\right\} + E\left\{V\left(\bar{y}^* | n_1, n_2\right)\right\}}{\bar{Y}^2}$$

$$= \frac{V\left(\bar{y}\right) + E\left\{\frac{n_2^2}{n^2}\left(\frac{1}{m} - \frac{1}{n_2}\right)S_{y(2)}^2\right\}}{\bar{Y}^2} = \frac{f_1 S_y^2 + E\left\{\frac{n_2}{n^2}\left(\frac{n_2}{m} - 1\right)s_{y(2)}^2\right\}}{\bar{Y}^2}$$

$$= \frac{f_1 S_y^2 + \frac{(k-1)}{n} E\left(\frac{n_2}{n}\right) E\left(s_{y(2)}^2\right)}{\bar{Y}^2} = f_1 C_y^2 + W_2 \frac{(k-1)}{n} C_{y(2)}^2.$$

where

$\bar{y}$ is the sample mean of the study variable y based on the second phase sample of size n

and $s_{y(2)}^2$ is the sample mean square of the study variable y based in the sample of size n_2. It may be noted $s_{y(2)}^2$ is unbiased estimator of $S_{y(2)}^2$.

The remaining proofs of the equation (17), may be obtained easily by following the proof of $E\left(e_0^2\right)$ shown above and the ways adopted by (Singh and Priyanka, 2007; Singh and Kumar, 2010 a, b).