Estimation of Population Variance in Two-Phase Sampling in Presence of Random Non - Response

A. Bandyopadhyay
Department of Mathematics, Asansol Engineering College
Asansol–713305, India
arnabbandyopadhyay4@gmail.com

G. N. Singh
Department of Applied Mathematics
Indian School of Mines, Dhanbad-826004, India
gnsingh_ism@yahoo.com

Abstract

The present investigation deals with the problem of estimation of population variance in presence of random non-response in two-phase (double) sampling. Using information on two auxiliary variables, two general classes of estimators have been suggested in two different situations of random non-response and studied their properties under two different set up of two-phase sampling. It is shown that several estimators may be generated from our proposed classes of estimators. Proposed classes of estimators are empirically compared with some contemporary estimators of population variance under the similar realistic situations and their performances have been demonstrated through numerical illustration and graphical interpretation which are followed by suitable recommendations.

Mathematics Subject Classification: 62D05

Keywords: Variance estimation, Two-phase sampling, Random non-response, Study variable, Auxiliary variable, Bias, Mean square error.

1. Introduction

The problem of estimation of population variance arises in many practical situations. For example, a physician needs a full understanding of variations in the degree of human blood pressure, body temperature and pulse rate for adequate prescription. An agriculturist needs an adequate understanding of the variations in climatic factors especially from place to place (or time to time) to be able to plan on when, how and where to plant his crop. The variance estimation technique using auxiliary variable was first considered by Das and Tripathi (1978). Further this was extended by Srivastava and Jhaji (1980), Isaki (1983), Singh (1983), Upadhyay and Singh (1983), Tripathi et al. (1988), Singh and Joarder (1998) and Ahamed et al. (2003) among others. In many situations, information on an auxiliary variable may be readily available on all unit 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.

However in some practical situations, it is common experience in sample surveys that data cannot always be collected from all the units selected in the sample. For example, the selected families may not be at home at the first attempt and some of them may refuse to cooperate with the interviewer even if contacted. As many respondents do not reply, available sample of returns is incomplete. The resulting incompleteness is called non-response and is sometimes so large that can completely vitiate the results. Statisticians have long known that failure to account for the stochastic nature of incompleteness can damage the actual conclusion. An obvious problem, that one needs to justify, arises when ignoring the incomplete mechanism. Rubin (1976) advocated three concepts: missing at random (MAR), observed at random (OAR), and parameter distribution (PD). Rubin defined: “The data are MAR if the probability of the observed missingness pattern, given the observed and unobserved data, does not depend on the value of the unobserved data”. Singh and Joarder (1998) studied the properties of ratio type estimator of population variance suggested by Isaki (1983) under two different situations of random non-response (MAR) advocated by Tracy and Osahan (1994) when (i) random non-response on both the study and auxiliary variables and (ii) only on the study variable. Singh *et al.* (2012) revisited the family of estimators of population variance suggested by Srivastava and Jhajj (1980) under the above situations of random non-responses.

It is worth to be mentioned that all the above recent works of estimation of population variance in presence of random non-response are discussed on the assumption that either population mean or both population mean and variance of the auxiliary variable are known and even if they are unknown, it is assumed that no non-response situations occur on the auxiliary variable in the sampled unit. This may not often be the case. In such situations, it is more generously advisable to draw 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. Motivated with these arguments and using information on two auxiliary variables, we have proposed two general classes of estimators of population variance in two-phase sampling applicable for two different realistic situations of random non-response and studied their properties under two different set up of two-phase sampling. It is shown that several estimators may be generated as member of the proposed classes of estimators. The superiorities of the proposed classes of estimators over some contemporary estimators of population variance under the similar realistic conditions have been established through numerical illustration and graphical interpretation. Suitable recommendations have been put forward to the survey statistician.

2. Formulation of Estimators

2.1. Two-Phase Sampling Structure

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. We wish to estimate the population variance

$S_y^2\left(S_y^2=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})^2\right)$ of the study variable y in the presence of the auxiliary variables x and z, when the population variance $S_x^2\left(S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{X})^2\right)$ of x is unknown but the information on z is available for all the units of population. To estimate S_y^2, a first phase sample S' of size n is drawn by simple random sampling without replacement scheme (SRSWOR) from the entire population U and observed for the auxiliary variables x and z to estimate S_x^2. Again a second-phase sample S of size m (m < n) is drawn according to the following cases by SRSWOR scheme to observe the characteristic y and x.

Case I: Second phase sample is drawn as a subsample of the first phase sample $(\text{i. e. } S\subset S')$.

Case II: Second phase sample S is drawn independently of the first phase sample S'.

Hence onwards, we use the following notations:

$\bar{y}_m, \bar{x}_m, \bar{z}_n$: Sample means of the respective variables based on the sample sizes shown in suffices.

$S_z^2=(N-1)^{-1}\sum_{i=1}^{N}(z_i-\bar{Z})^2$: Population variance of the auxiliary variable z.

$s_{x_m}^2=(m-1)^{-1}\sum_{i=1}^{m}(x_i-\bar{x}_m)^2$: Sample variance of the auxiliary variable x based on sample of size m.

$s_{y_m}^2, s_{x_n}^2$ and $s_{z_n}^2$: Sample variances of the respective variables based on sample sizes shown in suffices.

We assume that no non-response situations occur at the first phase sample S' while random non-response situations occur either on both the variables y and x or on the variable y alone in the second phase sample S. We have considered that the occurrences of random non-response situation follow the discrete probability distribution as presented below.

2.2. Non-Response Probability Model

If random non-response situations occur at the second phase sample S of size m and r $\{r = 0, 1, 2, . . ., (m - 2)\}$ denotes the number of sampling units on which information could not be collected due to random non-response, then the observations of the respective variables on which random non-response occur can be taken from the remaining $(m-r)$ units of the second phase sample. It is assumed that r is less than $(m-1)$, that is, $0\leq r\leq(m-2)$. We also assume that if p denotes the probability of a non-response among the $(m - 2)$ possible values of non-response, then r has the following discrete distribution

$$P(r) = \frac{(m - r)}{mq + 2p} \, {}^{m-2}C_r \, p^r q^{m-2-r}, \; r = 0, 1, 2, \ldots, (m - 2) \tag{1}$$

where $q = 1 - p$ and ${}^{m-2}C_r$ denote the total number of ways of obtaining r non-responses out of the $(m - 2)$ total possible non-responses, for instance, see Singh and Joarder (1998).

It is to be noted, the probability model, defined in equation (1), is free from actual data values; hence, can be considered as a model suitable for MAR situation.

We have defined following variables based on the responding part of the sample as $\bar{x}_m^* = \frac{1}{m - r}\sum_{i=1}^{m-r} x_i$, $\bar{y}_m^* = \frac{1}{m - r}\sum_{i=1}^{m-r} y_i$: Sample means of the respective variables based on the responding part of the second phase sample S.

$s_{x_m}^{*2} = (m - r - 1)^{-1} \sum_{i=1}^{m-r} (x_i - \bar{x}_m^*)^2$: Sample variance of the variable x based on the responding part of the second phase sample S.

$s_{y_m}^{*2}$: Sample variance of the study variable y based on the responding part of the second phase sample S.

2.3. Proposed Estimation Strategies

Utilizing information on an auxiliary variable x with unknown S_x^2 and following the work of Isaki (1983), one may propose the ratio type estimator of population variance S_y^2 in two-phase sampling as

$$t_R = s_{y_m}^2 \frac{s_{x_n}^2}{s_{x_m}^2} \tag{2}$$

Similarly, if $\bar{X}$ and S_x^2 both are unknown, then following the work of Srivastava and Jhajj (1980), one may define a general class of estimators of population variance S_y^2 in two-phase sampling set up as

$$t_g = g\left(s_{y_m}^2, u, v\right) \tag{3}$$

where $u = \frac{\bar{x}_m}{\bar{x}_n}$, $v = \frac{s_{x_m}^2}{s_{x_n}^2}$ and $g\left(s_{y_m}^2, u, v\right)$ is a parametric function that satisfies similar regularity conditions as given in Srivastava and Jhajj (1980) and is such that $g\left(S_y^2, 1, 1\right) = S_y^2$.

Motivated with the work of Singh and Joarder (1998) and Singh *et al.* (2012), one may propose the estimators t_R and t_g for two different situations of a random non-response at the second phase sample S as presented below.

(i) If random non-response occurs on both the variables y and x, the estimators t_R and t_g may be considered as

$$t_R^* = s^{*2}_{y_m} \frac{s^2_{x_n}}{s^{*2}_{x_m}} \tag{4}$$

and $$t_g^* = g\left(s^{*2}_{y_m}, u^*, v^*\right) \tag{5}$$

where $u^* = \frac{\bar{x}^*_m}{\bar{x}_n}$ and $v^* = \frac{s^{*2}_{x_m}}{s^2_{x_n}}$.

(ii) If random non-response occurs only on the study variable y, then the estimators t_R and t_g may be consider as

$$t_R^{**} = s^{*2}_{y_m} \frac{s^2_{x_n}}{s^2_{x_m}} \tag{6}$$

and $$t_g^{**} = g\left(s^{*2}_{y_m}, u, v\right) \tag{7}$$

Motivated with the above suggestions and assuming that the population variance S_x^2 of the auxiliary variable x is unknown, we have proposed two general classes of estimators of population variance S_y^2 in two-phase sampling set up applicable for two different situations of random non-response and presented below.

Situation I: In this situation, we assume that random non-response conditions occur on both the study variable y and the auxiliary variable x at the second phase sample S and also the population variance S_z^2 of the auxiliary variable z is known. Accordingly, we have suggested the general class of estimators of population variance S_y^2 in two-phase sampling set up as

$$T_1 = f\left(s^{*2}_{y_m}, s^{*2}_{x_m}, h_1\left(s^2_{x_n}, s^2_{z_n}\right)\right) \tag{8}$$

where $h_1\left(s^2_{x_n}, s^2_{z_n}\right)$ is a class of estimators of S_x^2 using information on $s^2_{x_n}$ and $s^2_{z_n}$, such that

$$h_1\left(S_x^2, S_z^2\right) = S_x^2. \tag{9}$$

We consider the composite function $f\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ as one-to-one function of $s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}$ and $s_{z_n}^{2}$ denoted by $T_1 = F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ such that

$$F\left(S_y^2, S_x^2, S_x^2, S_z^2\right) = S_y^2 \Rightarrow \left.\frac{\partial F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)}{\partial s_{y_m}^{*2}}\right|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)} = 1 \tag{10}$$

with $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$ and $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ satisfy the following regularity conditions:

1. Whatever be the chosen samples, $\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ assume values in a closed convex subspace, R^4 of the four dimensional real space containing the point $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$.

2. The function $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ is continuous and bounded in R^4.

3. The first, second and third order partial derivatives of $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ exist and are continuous and bounded in R^4.

It can be observed from equation (8) that the class of estimators T_1 is very wide in the sense for any parametric function, $f\left(s_{y_m}^{*2}, s_{x_m}^{*2}, h_1\left(s_{x_n}^{2}, s_{z_n}^{2}\right)\right)$ satisfying above regularity conditions with $F\left(S_y^2, S_x^2, S_x^2, S_z^2\right) = S_y^2$ may generate an estimators of S_y^2. For examples, the following ratio, product, regression and exponential type estimators of S_y^2 are the members of the class T_1.

$$t_{1i} = s_{y_m}^{*2}\frac{s_{x_{in}}^{2}}{s_{x_m}^{*2}},\ t_{2i} = s_{y_m}^{*2}\frac{s_{x_m}^{*2}}{s_{x_{in}}^{2}},\ t_{3i} = s_{y_m}^{*2} + b_1\left(s_{x_{in}}^{2} - s_{x_m}^{*2}\right),\ t_{4i} = s_{y_m}^{*2}\exp\left(\frac{s_{x_{in}}^{2} - s_{x_m}^{*2}}{s_{x_{in}}^{2} + s_{x_m}^{*2}}\right);\ \left(i = 1, 2, \ldots, 4\right)$$

where $s_{1x_n}^{2} = s_{x_n}^{2}\frac{S_z^2}{s_{z_n}^{2}},\ s_{2x_n}^{2} = s_{x_n}^{2}\frac{s_{z_n}^{2}}{S_z^2},\ s_{3x_n}^{2} = s_{x_n}^{2} + b_2\left(S_z^2 - s_{z_n}^{2}\right),\ s_{4x_n}^{2} = s_{x_n}^{2}\exp\left(\frac{S_z^2 - s_{z_n}^{2}}{S_z^2 + s_{z_n}^{2}}\right)$ and b_1, b_2 are the real scalars.

Situation II: In this case, we assume that random non-response situation occurs only on the study variable y while the complete information on the auxiliary variable x is available at the second phase sample S and also the population variance S_z^2 is known. Considering this aspect, we have proposed the general class of estimators of population variance S_y^2 in two-phase sampling set up as

$$T_2 = g\left(s_{y_m}^{*2}, s_{x_m}^{2}, h_1\left(s_{x_n}^{2}, s_{z_n}^{2}\right)\right). \tag{11}$$

We consider the composite function $g\left(s_{y_m}^{*2}, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2\right)$ as one-to-one function of $s_{y_m}^{*2}, s_{x_m}^2$, $s_{x_n}^2$ and $s_{z_n}^2$ denoted by $T_2 = G\left(s_{y_m}^{*2}, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2\right)$ such that

$$G\left(S_y^2, S_x^2, S_x^2, S_z^2\right) = S_y^2 \;\Rightarrow\; \left.\frac{\partial G\left(s_{y_m}^{*2}, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2\right)}{\partial s_{y_m}^{*2}}\right|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)} = 1 \qquad (12)$$

with $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$ and $G\left(s_{y_m}^{*2}, s_{x_m}^2, s_{x_n}^2, s_{z_n}^2\right)$ satisfy the similar regularity conditions as given for $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$ and $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^2, s_{z_n}^2\right)$ in equation (10).

Proceeding as above, it may be found that the class of estimators T_2 is also very wide and we present below some estimators of S_y^2 which are members of the class T_2.

$$t'_{1i} = s_{y_m}^{*2}\frac{s_{x_{in}}^2}{s_{x_m}^2},\; t'_{2i} = s_{y_m}^{*2}\frac{s_{x_m}^2}{s_{x_{in}}^2},\; t'_{3i} = s_{y_m}^{*2} + b_3\left(s_{x_{in}}^2 - s_{x_m}^2\right),\; t'_{4i} = s_{y_m}^{*2}\exp\left(\frac{s_{x_{in}}^2 - s_{x_m}^2}{s_{x_{in}}^2 + s_{x_m}^2}\right);\; (i = 1, 2, \ldots, 4)$$

where b_3 is a real scalar.

3. Bias and Mean Square Errors of the Proposed Classes of Estimators T_1 and T_2

The bias and mean square errors (M. S. E.s) of our proposed classes of estimators T_1 and T_2 are derived up to first order of approximations under large sample assumptions and using the following transformations:

$$s_{y_m}^{*2} = S_y^2\left(1+e_0\right), s_{x_m}^{*2} = S_x^2\left(1+e_1\right), s_{z_n}^2 = S_z^2\left(1+e_2\right), s_{x_n}^2 = S_x^2\left(1+e_3\right),\; s_{x_m}^2 = S_x^2\left(1+e_4\right).$$

Such that $\left|e_i\right| < 1 \;\forall\, (i = 0, 1, \ldots, 4)$.

We have derived the bias and mean square errors of the proposed classes of estimators T_1 and T_2 separately for the cases I and II of the two-phase sampling structure defined in section 2.1 and present them below.

3. 1 Bias and Mean Square Errors of the Proposed Classes of Estimators under Case I

In this section, we have considered that the second phase sample S of size m is drawn as a subsample from of the first phase sample S' of size n and we have the following results.

$$\left.\begin{aligned}
&E\left(e_0^2\right) = f^*C_0^2, E\left(e_1^2\right) = f^*C_1^2, E\left(e_2^2\right) = f_2C_2^2, E\left(e_3^2\right) = f_2C_1^2, E\left(e_4^2\right) = f_1C_1^2, E\left(e_0e_1\right) = f^*\rho_{01}C_0C_1,\\
&E\left(e_0e_2\right) = f_2\rho_{02}C_0C_2, E\left(e_0e_3\right) = f_2\rho_{01}C_0C_1, E\left(e_0e_4\right) = f_1\rho_{01}C_0C_1, E\left(e_1e_2\right) = f_2\rho_{12}C_1C_2,\\
&E\left(e_1e_3\right) = E\left(e_3e_4\right) = f_2C_1^2,\; E\left(e_1e_4\right) = f_1C_1^2,\; E\left(e_2e_3\right) = E\left(e_2e_4\right) = f_2\rho_{12}C_1C_2,
\end{aligned}\right\} \quad (13)$$

where

$f^{*}=\left(\frac{1}{mq+2p}-\frac{1}{N}\right),\ f_1=\left(\frac{1}{m}-\frac{1}{N}\right),\ f_2=\left(\frac{1}{n}-\frac{1}{N}\right),\ f_3=\left(\frac{1}{m}-\frac{1}{n}\right),\ f'=\left(\frac{1}{mq+2p}-\frac{1}{n}\right),$

$\mu_{abc}=\frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{Y})^a(x_i-\bar{X})^b(z_i-\bar{Z})^c$; (a, b, c) being non negative integers,

$\lambda_{abc}=\mu_{abc}\Big/\left\{\mu_{200}^{a/2}\ \mu_{020}^{b/2}\ \mu_{002}^{c/2}\right\},\ C_0=\sqrt{(\lambda_{400}-1)},\ C_1=\sqrt{(\lambda_{040}-1)},\ C_2=\sqrt{(\lambda_{004}-1)},$

$\rho_{01}=(\lambda_{220}-1)\Big/\sqrt{(\lambda_{400}-1)(\lambda_{040}-1)},\ \rho_{02}=(\lambda_{202}-1)\Big/\sqrt{(\lambda_{400}-1)(\lambda_{004}-1)},$

$\rho_{12}=(\lambda_{022}-1)\Big/\sqrt{(\lambda_{040}-1)(\lambda_{004}-1)}.$

From the above expectations, it is to be noted that:

(a) If p = 0 (there is no non-response), the above expected values of the sample statistics on which random non-responses occur coincide with the usual results.

(b) ρ_{01} is the correlation between $\left(y-\bar{Y}\right)^2$ and $\left(x-\bar{X}\right)^2$. Similarly ρ_{12} is the correlation between $\left(x-\bar{X}\right)^2$ and $\left(z-\bar{Z}\right)^2$ and ρ_{02} is the correlation between $\left(y-\bar{Y}\right)^2$ and $\left(z-\bar{Z}\right)^2$; see for instance Upadhyaya and Singh (2006).

Now, to express the class of estimators T_1 in terms of e's, we expand $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ about the point $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$ in a third order of Taylor's series expansions and we have

$$\begin{aligned}
F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right) &= F\left(S_y^2, S_x^2, S_x^2, S_z^2\right)+d_1\left(s_{y_m}^{*2}-S_y^2\right)+d_2\left(s_{x_m}^{*2}-S_x^2\right)+d_3\left(s_{x_n}^{2}-S_x^2\right)+d_4\left(s_{z_n}^{2}-S_z^2\right) \qquad (14)\\
&+\frac{1}{2}\Big\{d_{11}\left(s_{y_m}^{*2}-S_y^2\right)^2+d_{22}\left(s_{x_m}^{*2}-S_x^2\right)^2+d_{33}\left(s_{x_n}^{2}-S_x^2\right)^2+d_{44}\left(s_{z_n}^{2}-S_z^2\right)^2\\
&+2d_{12}\left(s_{y_m}^{*2}-S_y^2\right)\left(s_{x_m}^{*2}-S_x^2\right)+2d_{13}\left(s_{y_m}^{*2}-S_y^2\right)\left(s_{x_n}^{2}-S_x^2\right)+2d_{14}\left(s_{y_m}^{*2}-S_y^2\right)\left(s_{z_n}^{2}-S_z^2\right)\\
&+2d_{23}\left(s_{x_m}^{*2}-S_x^2\right)\left(s_{x_n}^{2}-S_x^2\right)+2d_{24}\left(s_{x_m}^{*2}-S_x^2\right)\left(s_{z_n}^{2}-S_z^2\right)+2d_{34}\left(s_{x_n}^{2}-S_x^2\right)\left(s_{z_n}^{2}-S_z^2\right)\Big\}\\
&+\frac{1}{6}\left\{\left(s_{y_m}^{*2}-S_y^2\right)\frac{\partial}{\partial s_{y_m}^{*2}}+\left(s_{x_m}^{*2}-S_x^2\right)\frac{\partial}{\partial s_{x_m}^{*2}}+\left(s_{x_n}^{2}-S_x^2\right)\frac{\partial}{\partial s_{x_n}^{2}}+\left(s_{z_n}^{2}-S_z^2\right)\frac{\partial}{\partial s_{z_n}^{2}}\right\}^3 F\left(s_{y_m}'^{*2}, s_{x_m}'^{*2}, s_{x_n}'^{2}, s_{z_n}'^{2}\right)
\end{aligned}$$

where

$$d_1=\frac{\partial}{\partial s_{y_m}^{*2}}F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)},\quad d_2=\frac{\partial}{\partial s_{x_m}^{*2}}F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)},$$

$$d_3=\frac{\partial}{\partial s_{x_n}^{2}}F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)},\quad d_4=\frac{\partial}{\partial s_{z_n}^{2}}F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)}$$

and $\left(d_{11}, d_{22}, d_{33}, d_{44}, d_{12}, d_{13}, d_{14}, d_{23}, d_{24}, d_{34}\right)$ are the second order partial derivatives of $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ at the point $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$ and $s_{y_m}^{\prime *2} = S_y^2 + \theta\left(s_{y_m}^{*2} - S_y^2\right)$, $s_{x_m}^{\prime *2} = S_x^2 + \theta\left(s_{x_m}^{*2} - S_x^2\right)$, $s_{x_n}^{\prime 2} = S_x^2 + \theta\left(s_{x_n}^{2} - S_x^2\right)$, $s_{z_m}^{\prime 2} = S_z^2 + \theta\left(s_{z_m}^{2} - S_z^2\right)$ for $\left(0 < \theta < 1\right)$.

In the light of the conditions mentioned for $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ in equation (10), it is noted that

$$F\left(S_y^2, S_x^2, S_x^2, S_z^2\right) = S_y^2 \;\Rightarrow d_1 = 1 \text{ and } d_{11} = \frac{\partial^2}{\partial\left(s_{y_m}^{*2}\right)^2} F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)} = 0. \quad (15)$$

Since the population variance S_x^2 of the auxiliary variable x is unknown, therefore, we have to impose the constraint as

$$d_2 = - d_3 . \quad (16)$$

Thus, expressing $F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ in terms of e's and neglecting the terms of e's having power greater than two we get

$$\begin{aligned} T_1 = F\left(s_{y_m}^{*2}, s_{x_m}^{*2}, s_{x_n}^{2}, s_{z_n}^{2}\right) &= S_y^2\left(1+e_0\right) + d_2 S_x^2\left(e_1 - e_3\right) + d_4 S_z^2 e_2 + \frac{1}{2}\Big\{S_x^4\left(d_{22}e_1^2 + d_{33}e_3^2 + 2d_{23}e_1e_3\right) + d_{44}S_z^4e_2^2 \\ &+ 2S_y^2S_x^2\left(d_{12}e_0e_1 + d_{13}e_0e_3\right) + 2d_{14}S_y^2S_z^2e_0e_2 + 2S_x^2S_z^2\left(d_{24}e_1e_2 + d_{34}e_2e_3\right)\Big\} \end{aligned} \quad (17)$$

Similarly, expressing T_2 in terms of e's we have

$$\begin{aligned} T_2 = G\left(s_{y_m}^{*2}, s_{x_m}^{2}, s_{x_n}^{2}, s_{z_n}^{2}\right) &= S_y^2\left(1+e_0\right) + c_2 S_x^2\left(e_4 - e_3\right) + c_4 S_z^2 e_2 + \frac{1}{2}\Big\{S_x^4\left(c_{22}e_4^2 + c_{33}e_3^2 + 2c_{23}e_3e_4\right) + c_{44}S_z^4e_2^2 \\ &+ 2S_y^2S_x^2\left(c_{12}e_0e_4 + c_{13}e_0e_3\right) + 2c_{14}S_y^2S_z^2e_0e_2 + 2S_x^2S_z^2\left(c_{24}e_2e_4 + c_{34}e_2e_3\right)\Big\} \end{aligned} \quad (18)$$

where

$$c_2 = \frac{\partial}{\partial s_{x_m}^2} G\left(s_{y_m}^{*2}, s_{x_m}^{2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)} = -\frac{\partial}{\partial s_{x_n}^2} G\left(s_{y_m}^{*2}, s_{x_m}^{2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)} \left\{\text{as } S_x^2 \text{ is unknown}\right\},$$

$$c_4 = \frac{\partial}{\partial s_{z_n}^2} G\left(s_{y_m}^{*2}, s_{x_m}^{2}, s_{x_n}^{2}, s_{z_n}^{2}\right)\Bigg|_{\left(S_y^2, S_x^2, S_x^2, S_z^2\right)}$$

and $\left(c_{22}, c_{33}, c_{44}, c_{12}, c_{13}, c_{14}, c_{23}, c_{24}, c_{34}\right)$ are the second order partial derivatives of $G\left(s_{y_m}^{*2}, s_{x_m}^{2}, s_{x_n}^{2}, s_{z_n}^{2}\right)$ at the point $\left(S_y^2, S_x^2, S_x^2, S_z^2\right)$.

Taking expectations on both sides of the equations (17), (18) and using the results in equation (13), we obtain the expressions for bias B(.) and mean square errors M(.) of the proposed classes of estimators T_1 and T_2 to the first order of approximations as

$$B\left(T_1\right) = E\left(T_1 - \bar{Y}\right) = \frac{1}{2}\left[\begin{aligned} &S_x^4C_1^2\left(d_{22}f^* + d_{33}f_2 + 2d_{23}f_2\right) + 2S_y^2S_x^2\left(d_{12}f^* + d_{13}f_2\right)\rho_{01}C_0C_1 \\ &+ d_{44}f_2S_z^4C_2^2 + 2d_{14}f_2S_y^2S_z^2\rho_{02}C_0C_2 + 2f_2\left(d_{24} + d_{34}\right)S_x^2S_z^2\rho_{12}C_1C_2 \end{aligned}\right], \quad (19)$$

$$B(T_2)=E(T_2-\bar{Y})=\frac{1}{2}\left[\begin{array}{l} S_x^4C_1^2(c_{22}f_1+c_{33}f_2+2c_{23}f_2)+2S_y^2S_x^2(c_{12}f_1+c_{13}f_2)\rho_{01}C_0C_1 \\ +c_{44}f_2S_z^4C_2^2+2c_{14}f_2S_y^2S_z^2\rho_{02}C_0C_2+2f_2(c_{24}+c_{34})S_x^2S_z^2\rho_{12}C_1C_2 \end{array}\right], \quad (20)$$

$$M(T_1)=E(T_1-\bar{Y})^2=f^*S_y^4C_0^2+d_2^2S_x^4f'C_1^2+d_4^2f_2S_z^4C_2^2+2d_2f'S_x^2S_y^2\rho_{01}C_0C_1+2d_4f_2S_y^2S_z^2\rho_{02}C_0C_2 \quad (21)$$

$$M(T_2)=E(T_2-\bar{Y})^2=f^*S_y^4C_0^2+c_2^2f_3S_x^4C_1^2+c_4^2f_2S_z^4C_2^2+2c_2f_3S_x^2S_y^2\rho_{01}C_0C_1+2c_4f_2S_y^2S_z^2\rho_{02}C_0C_2 \quad (22)$$

3. 2 Bias and Mean Square Errors of the Proposed Classes of Estimators under Case II

If the second phase sample S is drawn independently of the first phase sample S', then we have the following results.

$$\left.\begin{array}{l} E(e_0^2)=f^*C_0^2,\ E(e_1^2)=f^*C_1^2,\ E(e_2^2)=f_2C_2^2,\ E(e_3^2)=f_2C_1^2,\ E(e_4^2)=f_1C_1^2, \\ E(e_0e_1)=f^*\rho_{01}C_0C_1,\ E(e_0e_4)=f_1\rho_{01}C_0C_1,\ E(e_1e_4)=f_1C_1^2,\ E(e_2e_3)=f_2\rho_{12}C_1C_2, \\ E(e_0e_2)=E(e_0e_3)=E(e_1e_3)=E(e_3e_4)=E(e_2e_4)=E(e_1e_2)=0 \end{array}\right\} \quad (23)$$

Proceeding as section 3.1 and using the results in equation (23), we have derived the expressions for bias B(.) and mean square errors M(.) of the proposed classes of estimators T_1 and T_2 to the first order of approximations as

$$B(T_1)=E(T_1-\bar{Y})=\frac{1}{2}\left[S_x^4C_1^2(d_{22}f^*+d_{33}f_2)+d_{44}f_2S_z^4C_2^2+2d_{12}f^*S_x^2S_y^2\rho_{01}C_0C_1+2d_{34}f_2S_x^2S_z^2\rho_{12}C_1C_2\right], \quad (24)$$

$$B(T_2)=E(T_2-\bar{Y})=\frac{1}{2}\left[S_x^4C_1^2(c_{22}f_1+c_{33}f_2)+c_{44}f_2S_z^4C_2^2+2c_{12}f_1S_x^2S_y^2\rho_{01}C_0C_1+2c_{34}f_2S_x^2S_z^2\rho_{12}C_1C_2\right], \quad (25)$$

$$M(T_1)=E(T_1-\bar{Y})^2=f^*S_y^4C_0^2+d_2^2S_x^4(f^*+f_2)C_1^2+d_4^2f_2S_z^4C_2^2+2d_2f^*S_x^2S_y^2\rho_{01}C_0C_1-2d_2d_4f_2S_x^2S_z^2\rho_{12}C_1C_2, \quad (26)$$

and

$$M(T_2)=E(T_2-\bar{Y})^2=f^*S_y^4C_0^2+c_2^2S_x^4(f_1+f_2)C_1^2+c_4^2f_2S_z^4C_2^2+2c_2f_1S_x^2S_y^2\rho_{01}C_0C_1-2c_2c_4f_2S_x^2S_z^2\rho_{12}C_1C_2. \quad (27)$$

Remark 3.1.

The bias and mean square errors of the various estimators (indicated in section 2.3) belonging to the classes of estimators T_1 and T_2 can be easily obtained by substituting the suitable values of the derivatives in equations (19)-(22) and (24)-(27) as suggested by Singh *et al.* (2007) and Singh and Vishwakarma (2007).

4. Minimum M. S. E.s of the Proposed Classes of Estimators T_1 and T_2

It is obvious from the equations (21), (22), (26), (27) and remark 3.1 that the mean square errors of the proposed classes of estimators $T_i\ (i = 1, 2)$ depend on the different values of the derivatives d_2, d_4, c_2 and c_4. Therefore, we desire to minimize the mean square errors of the proposed classes of estimators T_i separately for two different cases of two-phase sampling set up considered in this work and shown below:

Case I

When second phase sample S is drawn as a sub sample of the first phase sample S', the optimality conditions under which proposed classes of estimators $T_i\left(i = 1, 2\right)$ have minimum M. S. Es are obtained as

$$\left. d_2 = c_2 = -\rho_{01}\frac{S_y^2 C_0}{S_x^2 C_1},\ d_4 = c_4 = -\rho_{02}\frac{S_y^2 C_0}{S_z^2 C_2} \right\} \tag{28}$$

Substituting these optimum values of the derivatives in equations (21) and (22), we have the minimum M. S. E.s of the classes of estimators $T_i\left(i = 1, 2\right)$ as

$$\text{Min. } M\left(T_1\right) = \left(f^* - f'\rho_{01}^2 - f_2\rho_{02}^2\right)C_0^2 S_y^4 \tag{29}$$

and $$\text{Min. } M\left(T_2\right) = \left(f^* - f_3\rho_{01}^2 - f_2\rho_{02}^2\right)C_0^2 S_y^4 . \tag{30}$$

Case II

When second phase sample S is selected independently of the first phase sample S', the optimality conditions which minimize the mean square errors of the proposed classes of estimators T_1 and T_2 are obtained as

$$\left. d_2 = -\frac{f^*\rho_{01}S_y^2C_0}{\left(f^*+f_2\right)S_x^2C_1},\ d_4 = -\frac{f^*\rho_{01}\rho_{12}S_y^2C_0}{\left(f^*+f_2\right)S_z^2C_2},\ c_2 = -\frac{f_1\rho_{01}S_y^2C_0}{\left(f_1+f_2\right)S_x^2C_1},\ c_4 = -\frac{f_1\rho_{01}\rho_{12}S_y^2C_0}{\left(f_1+f_2\right)S_z^2C_2} \right\} \tag{31}$$

Substituting these optimum values of the derivatives d_2, d_4, c_2 and c_4 in equations (26) and (27), we have the expressions of minimum M. S. E. of the classes of estimators $T_i\left(i = 1, 2\right)$ as

$$\text{Min. } M\left(T_1\right) = \left\{ f^* - \frac{\left(f^*\rho_{01}\right)^2}{\left(f^*+f_2\right)} - f_2\frac{\left(f^*\rho_{01}\rho_{12}\right)^2}{\left(f^*+f_2\right)^2} \right\} C_0^2 S_y^4 \tag{32}$$

and $$\text{Min. } M\left(T_2\right) = \left\{ f^* - \frac{\left(f_1\rho_{01}\right)^2}{\left(f_1+f_2\right)} - f_2\frac{\left(f_1\rho_{01}\rho_{12}\right)^2}{\left(f_1+f_2\right)^2} \right\} C_0^2 S_y^4 . \tag{33}$$

Remark 4.1: It is to be noted from optimality conditions in equations (28) and (31) that the optimum values of derivatives of the proposed classes of estimators $T_i\left(i = 1, 2\right)$ depend on unknown population parameters such as $C_0, C_1, C_2, \rho_{01}, \rho_{12}, \rho_{02}, S_y^2$ and S_x^2. Thus, to use such estimators one has to use guessed or estimated values of them. Guessed values of population parameters can be obtained either from past data or experience gathered over time; for instance see Murthy (1967) and Tracy et al. (1996). If the guessed values are not known then it is advisable to use sample data to estimate these parameters as suggested by Singh *et al.* (2007) and Gupta and Shabbir (2008). It could be seen that the minimum mean square errors of the classes of estimators remain same up to the first order of approximations, even if population parameters are replaced by their respective sample estimates.

5. Efficiency Comparisons of the Proposed Classes of Estimators T_1 and T_2

To examine the performances of the proposed classes of estimators under two different cases of two-phase sampling set up as suggested in this paper, we have compared their efficiencies with some other estimators of population variance such as $s_{y_m}^{*2}$ (sample variance estimator in presence of random non-response), t_R^*, t_g^*, t_R^{**} and t_g^{**}. Proceeding as sections 3 and 4, the M. S. E.s/ minimum M. S. E.s of these estimators are derived up to the first order of approximations under the Cases I and II of the two phase-sampling set up and presented below.

Case I:

$$M(t_R^*) = S_y^4\left[f^*C_0^2 + f'C_1^2 - 2f'\rho_{01}C_0C_1\right] \tag{34}$$

$$\text{Min. } M(t_g^*) = S_y^4\left[f^*C_0^2 + w_1^2 f'C_x^2 + w_2^2 f'C_1^2 + 2w_1 f'\lambda_{210}C_x + 2w_2 f'\rho_{01}C_0C_1 + 2w_1w_2 f'\lambda_{030}C_x\right] \tag{35}$$

$$M(t_R^{**}) = S_y^4\left[f^*C_0^2 + f_3C_1^2 - 2f_3\rho_{01}C_0C_1\right] \tag{36}$$

$$\text{Min. } M(t_g^{**}) = S_y^4\left[f^*C_0^2 + w_1^2 f_3C_x^2 + f_3w_2^2C_1^2 + 2f_3w_1\lambda_{210}C_x + 2w_2f_3\rho_{01}C_0C_1 + 2w_1w_2f_3\lambda_{030}C_x\right] \tag{37}$$

Case II:

$$M(t_R^*) = S_y^4\left[f^*C_0^2 + \left(f^* + f_2\right)C_1^2 - 2f^*\rho_{01}C_0C_1\right] \tag{38}$$

$$\text{Min. } M(t_g^*) = S_y^4\left[f^*C_0^2 + w_1'^2\left(f^*+f_2\right)C_x^2 + w_2'^2\left(f^*+f_2\right)C_1^2 + 2w_1'f^*\lambda_{210}C_x + 2w_2'f^*\rho_{01}C_0C_1 + 2w_1'w_2'\left(f^*+f_2\right)\lambda_{030}C_x\right] \tag{39}$$

$$M(t_R^{**}) = S_y^4\left[f^*C_0^2 + \left(f_1+f_2\right)C_1^2 - 2f_1\rho_{01}C_0C_1\right] \tag{40}$$

$$\text{Min. } M(t_g^{**}) = S_y^4\left[f^*C_0^2 + v_1^2\left(f_1+f_2\right)C_x^2 + v_2^2\left(f_1+f_2\right)C_1^2 + 2v_1f_1\lambda_{210}C_x + 2v_2f_1\rho_{01}C_0C_1 + 2v_1v_2\left(f_1+f_2\right)\lambda_{030}C_x\right] \tag{41}$$

where

$$C_x = \frac{S_x}{\bar{\bar{X}}}, \qquad w_1 = \frac{\left\{\lambda_{030}\rho_{01}C_0 - \lambda_{210}C_1\right\}C_1}{C_x\left(C_1^2 - \lambda_{030}^2\right)}, \; w_2 = \frac{\left\{\lambda_{030}\lambda_{210} - \rho_{01}C_0C_1\right\}}{C_1^2 - \lambda_{030}^2},$$

$$w_1' = \frac{f^*\left\{\lambda_{030}\rho_{01}C_0 - \lambda_{210}C_1\right\}C_1}{\left(f^* + f_2\right)C_x\left(C_1^2 - \lambda_{030}^2\right)},$$

$$w_2' = \frac{f^*\left\{\lambda_{030}\lambda_{210} - \rho_{01}C_0C_1\right\}}{\left(f^*+f_2\right)\left(C_1^2 - \lambda_{030}^2\right)}, \; v_1 = \frac{f_1\left\{\lambda_{030}\rho_{01}C_0 - \lambda_{210}C_1\right\}C_1}{\left(f_1+f_2\right)C_x\left(C_1^2 - \lambda_{030}^2\right)} \text{ and } v_2 = \frac{f_1\left\{\lambda_{030}\lambda_{210} - \rho_{01}C_0C_1\right\}}{\left(f_1 + f_2\right)\left(C_1^2 - \lambda_{030}^2\right)}.$$

The variance of $s_{y_m}^{*2}$ can be obtained to the first order of approximation as

$$V\left(s_{y_m}^{*2}\right) = f^*C_0^2\, S_y^4. \tag{42}$$

The performances of the proposed classes of estimators T_1 and T_2 under their respective optimality conditions are compared with the other estimators considered in this paper and their dominance have been shown by empirical and graphical means of comparisons.

5.1. Numerical Illustration

We have chosen four natural population data sets to illustrate the efficacious performances of the proposed classes of estimators T_1 and T_2. 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: Cochran (1977, Page- 182)

y: Number of ‘placebo’ children.

x: Number of paralytic polio cases in the placebo group.

z: Number of paralytic polio cases in the ‘not inoculated’ group.

N= 34, n = 20, m = 12, $C_0 = 2.32188$, $C_1 = 1.82685$, $C_x = 1.2333$, $\rho_{01} = 0.6661$, $\rho_{02} = 0.5657$, $\rho_{12} = 0.6005$, $\lambda_{030} = 1.5224$ and $\lambda_{210} = 1.4083$.

Population II-Source: Murthy (1967, Page- 399)

y: Area under wheat in 1964.

x: Area under wheat in 1963.

z: Cultivated area in 1961.

N= 34, n = 20, m= 12, $C_0 = 1.6510$, $C_1 = 1.3828$, $C_x = 0.7205$, $\rho_{01} = 0.9218$, $\rho_{02} = 0.8914$, $\rho_{12} = 0.9346$, $\lambda_{030} = 0.9345$ and $\lambda_{210} = 1.0196$.

Population III- Source: Sukhatme (1970, Page- 185)

y: Area under wheat in 1937.

x: Area under wheat in 1936.

z:Total cultivated area in 1931.

N= 34, n = 20, m= 12, $C_0 = 1.5959$, $C_1 = 1.5105$, $C_x = 0.7678$, $\rho_{01} = 0.6251$, $\rho_{02} = 0.8007$, $\rho_{12} = 0.5342$, $\lambda_{030} = 1.0982$ and $\lambda_{210} = 0.8886$.

Population IV-Source: Murthy (1967, Page- 288)

y: Output.

x: Fixed Capital

z: Number of workers.

N= 80, n = 60, m= 40, $C_0 = 1.1255$, $C_1 = 1.6065$, $C_x = 0.9485$, $\rho_{01} = 0.7319$, $\rho_{02} = 0.7940$,

$\rho_{12} = 0.9716$, $\lambda_{030} = 1.2761$ and $\lambda_{210} = 0.5461$.

For different choices of non-response rate p, the performances of the proposed classes of estimators T_i $(i = 1, 2)$ under their respective optimality conditions are compared with the other estimators considered in this work. The performances of the proposed classes of estimators $T_i (i = 1, 2)$ have been shown in terms of the percent relative efficiencies and presented in Tables 1-2. The percent relative efficiencies of the proposed classes of estimators T_i with respect to an estimator t is defined as

$$\text{PRE} = \frac{M(t)}{\text{Min. } M(T_i)} \times 100 \tag{43}$$

where $M(t)$ denotes the M. S. E./ minimum M. S. E. of an estimator t.

Table 1: PREs of the class of estimators T_1 with respect to other estimators when non-response situations occur on the study variable y as well as on the auxiliary variable x at the second phase sample

Population I						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^*_R	t^*_g	$s^{*2}_{y_m}$	t^*_R	t^*_g
p = 0.05	166.5081	120.6259	118.1392	155.7111	123.3882	103.8387
P = 0.10	167.2997	119.5488	116.9608	156.8451	122.0540	103.7100
P = 0.15	168.0647	118.5078	115.8219	157.9718	120.7777	103.6017
P = 0.20	168.8045	117.5012	114.7207	159.0911	119.5555	103.5122
Population II						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^*_R	t^*_g	$s^{*2}_{y_m}$	t^*_R	t^*_g
p = 0.05	588.3309	269.9376	267.2372	434.57853	177.4551	162.5814
P = 0.10	592.7544	260.4811	257.6631	446.62791	175.1699	162.3549
P = 0.15	597.0525	251.2929	248.3605	458.74101	172.8349	162.1025
P = 0.20	601.2304	242.3616	239.3180	470.88961	170.4416	161.8155
Population III						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^*_R	t^*_g	$s^{*2}_{y_m}$	t^*_R	t^*_g
p = 0.05	192.4331	156.9049	141.3974	144.8091	149.5909	100.7661
P = 0.10	190.3244	153.9273	138.0406	145.7102	147.5202	100.6426
P = 0.15	188.3483	151.1370	134.8949	146.6061	145.5357	100.5384
P = 0.20	186.4927	148.5168	131.9409	147.4967	143.6320	100.4520
Population IV						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^*_R	t^*_g	$s^{*2}_{y_m}$	t^*_R	t^*_g
p = 0.05	229.5861	221.2544	139.8620	200.5576	313.9710	114.2487
P = 0.10	228.2167	219.6065	135.4945	202.2888	305.3023	113.3906
P = 0.15	226.9754	218.1129	131.5356	203.8629	297.1642	112.5688
P = 0.20	225.8451	216.7528	127.9307	205.2914	289.5040	111.7785

Table 2: PREs of the class of estimators T_2 with respect to other estimators when non-response situation occurs only on the study variable y at the second phase sample

Population I						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g
p = 0.05	159.1079	119.5632	117.4200	149.4356	122.4523	103.5997
P = 0.10	153.2846	117.6359	115.7038	144.8337	120.3623	103.2646
P = 0.15	148.0951	115.9182	114.1743	140.6857	118.4783	102.9626
P = 0.20	143.4411	114.3779	112.8027	136.9275	116.7715	102.6889
Population II						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g
p = 0.05	447.4754	229.0520	227.1995	351.2661	162.0725	148.8398
P = 0.10	365.7847	198.7121	197.2952	302.4260	150.0071	139.3465
P = 0.15	311.3616	178.4995	177.3726	266.8959	141.2298	132.4404
P = 0.20	272.5062	164.0686	163.1489	239.8874	134.5576	127.1906
Population III						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g
p = 0.05	183.7342	153.1377	139.7829	140.0319	147.1906	100.8178
P = 0.10	174.3510	147.1832	135.3248	136.5194	143.0499	100.7461
P = 0.15	166.2233	142.0253	131.4633	133.3175	139.2755	100.6807
P = 0.20	159.1150	137.5144	128.0860	130.3869	135.8209	100.6208
Population IV						
Estimators	Case I			Case II		
	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g	$s^{*2}_{y_m}$	t^{**}_R	t^{**}_g
p = 0.05	206.5460	200.0258	136.3308	182.3425	286.3165	112.6068
P = 0.10	188.2745	182.8725	130.1004	169.6207	257.5310	110.6591
P = 0.15	174.1499	169.6122	125.2841	159.4229	234.4563	109.0978
P = 0.20	162.9042	159.0548	121.4495	151.0657	215.5464	107.8183

5.2. Graphical Interpretation

For different choices of correlations ρ_{01}, ρ_{02} and ρ_{12}, we have demonstrated the performances of our proposed classes of estimators by pictorial representation. This could not only improve the readability of the results but also allow the comparison of a much denser grid for different correlation values. For different values of $\rho_{01}, \rho_{02}, \rho_{12}$, N = 500, n = 200, m= 100 and p = 0.10, the PREs of the classes of estimators T_1 and T_2 are derived with respect to $s^{*2}_{y_m}$ and shown in Figures 1-4.

Figure 1: PRE of T_1 under Case I

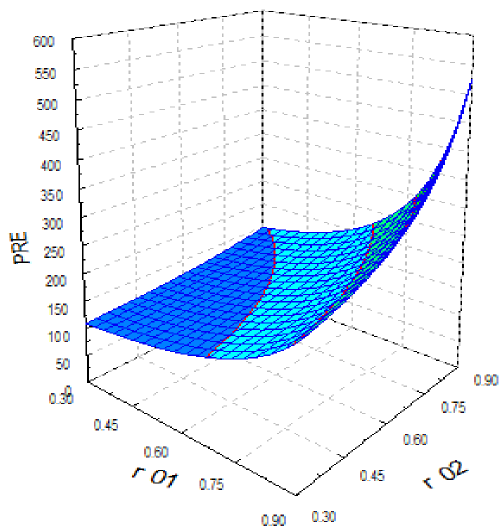


Figure 2: PRE of T_2 under Case I

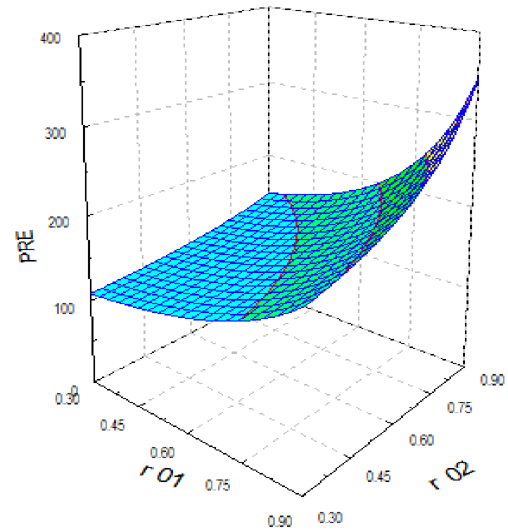


Figure 3: PRE of T_1 under Case II

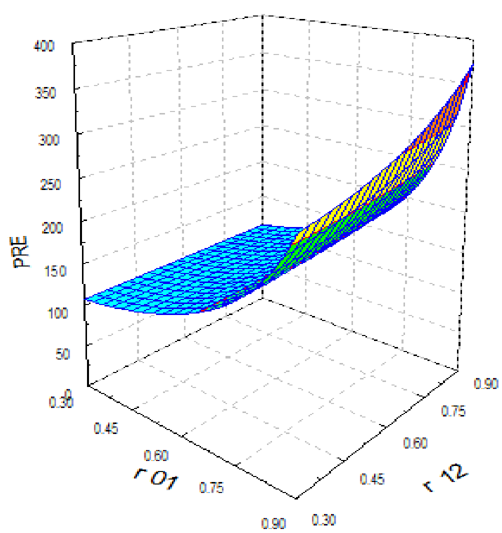
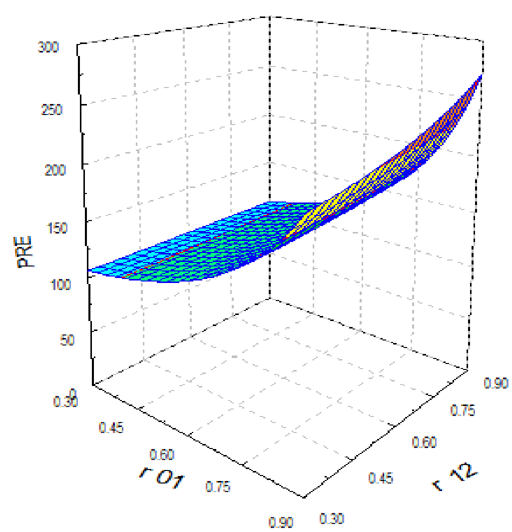


Figure 4: PRE of T_2 under Case II



Note: r 01, r 02 and r 12 denote ρ_{01}, ρ_{02} and ρ_{12} respectively in the Figures 1- 4.

6. Conclusions

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

1. From Tables 1 and 2, it is observed that

(a) For high positive values of the correlation coefficients ρ_{01}, ρ_{02} and ρ_{12} (specially for the populations II and IV), the proposed class of estimators T_1 yields impressive gains in efficiency over the other estimators $s^{*2}_{y_m}$, t^*_R and t^*_g and this behavior is visible from both the cases of the two-phase sampling set up as suggested in this work. Similar situations are also observed for the class of estimators T_2 when it dominates the estimators $s^{*2}_{y_m}$, t^{**}_R and t^{**}_g.

(b) For the different choices of non-response rate p, proposed classes of estimators $T_i\left(i = 1, 2\right)$ are more efficient than the other estimators considered in this work.

2. From Figures 1-4, it is noticed that

(a) The percent relative efficiencies of T_i $(i = 1, 2)$ are increasing with the increasing values of the correlation coefficients ρ_{01}, ρ_{02} and ρ_{12}. This phenomenon indicates that the proposed classes of estimators perform more precisely, if information on high positively correlated auxiliary variables is available.

Thus it is clear that the uses of auxiliary variables are highly rewarding in terms of the proposed classes of estimators. Hence, the propositions of the classes of estimators in the present study are highly justified as they unify several results. Therefore, the suggested classes of estimators are more attractive in comparison with the previous work of similar nature.

7. Recommendations of the Proposed Work for Real life Applications

In real life survey it may be found that the character of interest is sensitive or stigmatizing such as drinking alcohol, gambling habit, drug addiction, tax evasion, history of induced abortions etc. Hence, a direct survey is likely to yield unreliable responses because presence of random non – response situations in the sampled units. The suggested estimation strategies for estimating the character of interest are recommended to the survey statisticians to handle these realistic situations.

Acknowledgements: Authors are thankful to the reviewer for his valuable and constructive suggestions.

References

1. Ahmed, M. S., Abu-Dayyeh, W. and Hurairah, A. A. O. (2003): Some estimators for population variance under two phase sampling. Statistics in Transition 6 (1), 143–150.
2. Bose, C. (1943): Note on the sampling error in the method of double sampling. Sankhya, 6, 330.
3. Cochran, W. G. (1977): Sampling techniques. New-York: John Wiley and Sons.
4. Das, A. K. and Tripathi, T. P. (1978): Use of auxiliary information in estimating the finite population variance. Sankhya, C, 40, 139–148.
5. 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.
6. Isaki, C. T. (1983): Variance estimation using auxiliary information. Jour. Amer. Statist. Assoc, 78, 117–123.

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

8. Rubin, D. B. (1976): Inference and missing data. Biometrika, 63(3), 581–592.

9. Sukhatme, P. V. and Sukhatme, B. V. (1970): Sampling Theory of Surveys with Applications, Second Edition, Asia Publishing House, London.

10. Srivastava, S. K. and Jhajj, H. S. (1980): A class of estimators using auxiliary information for estimating finite population variance. Sankhya C, 42, 1–2, 87–96.

11. Singh, R. K. (1983): Estimation of finite population variance using ratio and product method of estimation. Biom. Jour., 25(2), 193–200.

12. Singh, S. and Joarder, A. H. (1998): Estimation of finite population variance using random non-response in survey sampling. Metrika, 98, 241–249.

13. 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 non-response. Test, 16: 565–597.

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

15. Singh, H. P., Tailor, R., Kim, J. M. and Singh, S. (2012): Families of estimators of finite population variance using a random non-response in survey sampling. The Korean Journal of Applied Statistics, 25(4), 681-695.

16. Tracy, D. S. and Osahan, S. S. (1994): Random non-response on study variable versus on study as well as auxiliary variables. Statistica, 54, 163–168.

17. 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.

18. Tripathi, T. P., Singh, H. P. and Upadhyaya, L. N. (1988): A generalized method of estimation in double sampling. Jour. Ind. Statist. Assoc., 26, 91– 101.

19. Upadhyaya, L. N. and Singh, H. P. (1983): Use of auxiliary information in the estimation of population variance. Mathematical Forum, 6, (2), 33–36.

20. Upadhyaya, L. N. and Singh, H. P. (2006): Almost unbiased ratio and product-type estimators of finite population variance in sample surveys. Statistics in Transition. 7(5), 1087-1096.