

Non-Response in Sample Surveys - A New Approach

Hassan Zeb¹, Muhammad Azeem², Sareer Badshah³

^{1,2,3} Department of Statistics, Islamia College Peshawar, Pakistan.

Email: ¹hassanzeb@icp.edu.pk, ²azeem_stats@icp.edu.pk, ³sareerbadshah@gamil.com

Received: 21 August, 2018 / Accepted: 22 November, 2018 / Published online: 26 March, 2019

Abstract. In performing sample surveys, the investigator often faces the problem of non-response and the possibility of bias as a result of this phenomenon. This article discusses the issue of estimating the population mean in the presence of non-response under simple random sampling scheme using two auxiliary variables. These missing observations may occur on auxiliary, study or both of the variables. A weighted ratio-cum-exponential factor-type estimator is suggested and its properties have been studied under large sample approximation. The mean square error of suggested estimator is minimum as compared to selected estimators in the study. The theoretical results are verified through a simulation study.

AMS (MOS) Subject Classification Codes: 62D05

Key Words: Auxiliary; bias; factor-type; investigator; mean square error.

1. INTRODUCTION

The precision of the estimators is improved when, in sample surveys, information on the auxiliary variables is incorporated only at the estimation stage, designing stage or both stages provided that study and auxiliary variables are correlated. Gain in efficiency is obtained when information on suitable supplementary variables is properly utilized. When the study and supplementary variables are positively correlated, then ratio type of estimation is used Cochran [6]. When negative correlation exists between study and auxiliary variables, then Murthy [18] and Robson [23] proposed product type of estimation. Under various sampling schemes Dayyeh et al. [7], Ahmad [1], Azeem and Hanif [4], Chaudhary and Kumar [5], Gupta and Shabbir [9], Gupta and Shabbir [10], Azaz et al. [2], Kadilar and Cingi [12], Khan and Khan [13], Khare and Sinha [14], Kumar and Viswanathaiah [15], Muneer et al. [17], Noor-ul-Amin [20], Ozturk [21], Shabbir and Gupta [25], Srivastava and Jhaji [28],

Tripathy [29] and Verma and Singh [30] have proposed a large number of estimators of population parameters using knowledge on auxiliary variable(s).

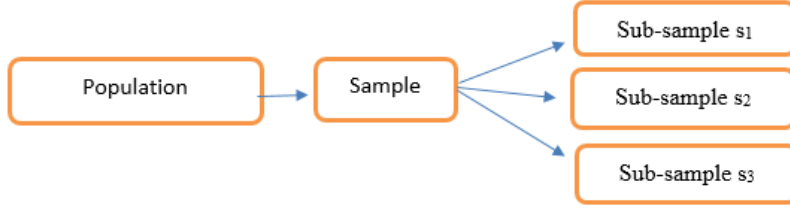
Under two phase sampling, Shukla [26] studied factor-type estimator and expression for minimum mean square error (MSE) is derived for a fixed cost. Estimating population mean under simple random sampling scheme, Shukla et al. [27] proposed three classes of estimators utilizing knowledge on two auxiliary variables and derived expressions for minimum MSE. Under two-phase sampling, Audu and Adewara [3] developed a modified factor-type estimator and obtained the expression of minimum mean square error, first and second-phase sample sizes for a (a) fixed cost and (b) fixed variance. For the first time, the seriousness of missing observations was dealt by Hansen and Hurwitz [11] in estimating population mean. An alternative approach to Hansen and Hurwitz [11] approach was developed by Grover and Kaur [8]. The selected respondents are divided into three mutually exclusive sub-groups i.e.

- (i) responded to both study and auxiliary variables, (sub-sample s_1)
- (ii) responded to auxiliary variable, only (sub-sample s_2)
- (iii) responded to study variable, only (sub-sample s_3).

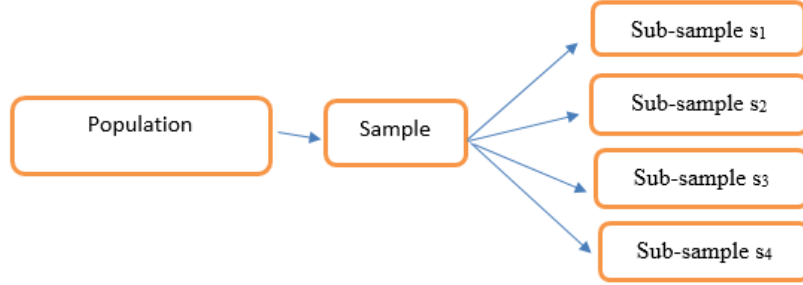
In the present article, a new method has been proposed where selected respondents are divided into four mutually exclusive sub-groups, s_1 , s_2 , and s_3 as defined in (i), (ii), and (iii) above and s_4 defined as: did not respond to any of the variables in the study. In this method, a questionnaire is initially mailed to the selected sample and those who did not answer to any variable are directly contacted with the assumption that all respondents provide information through direct contact method.

2. PROPOSED SAMPLING PLAN AND NOTATIONS

Let there be N units (U_1, U_2, \dots, U_N) in the population. Let Y and X_1, X_2 be the characteristics on study and auxiliary variables respectively. For i^{th} unit of population, let the triplet $(y_i, x_{1i}, x_{2i}); i = 1, 2, 3 \dots N$ designates the values of study and auxiliary variables respectively. From this population a sample “ s ” of n units is drawn using simple random sampling without replacement (SRSWOR). The population mean \bar{Y} is estimated utilizing known means \bar{X}_1 and \bar{X}_2 of auxiliary variables. A questionnaire is mailed to whole of the n units selected in sample “ s ”. Let in the sub-sample “ s_1 ”, for $(n - p - q - r)$ units full response occurs on all three variables, in sub-sample “ s_2 ”, p units responded on both auxiliary variables X_1, X_2 but no-reply on study variable Y , in sub-sample “ s_3 ”, q units responded on study variable but did not respond on both auxiliary variables and for r units in “ s_4 ”, information is absent on all variables. As in sub-sample “ s_4 ”, information is not available on all three variables, information is obtained by using a face-to-face interview method.

Figure 1: Grover and Kaur (2014) Method

$$s = s_1 \cup s_2 \cup s_3, \quad s_i \cap s_k = \emptyset \text{ for } j, k = 1, 2, 3 \text{ and } j \neq k$$

Figure 2: Suggested Method

$$s = s_1 \cup s_2 \cup s_3 \cup s_4, \quad s_i \cap s_k = \emptyset \text{ for } j, k = 1, 2, 3, 4 \text{ and } j \neq k$$

For different sub-samples, the sample means of study and auxiliary variables are;

$$\bar{x}_l = \frac{\sum_{i \in s_1} x_{li}}{(n-p-q-r)}, \quad \bar{y} = \frac{\sum_{i \in s_1} y_i}{(n-p-q-r)}, \quad \bar{x}'_l = \frac{\sum_{i \in s_2} x'_{li}}{p}, \quad \bar{y}'' = \frac{\sum_{i \in s_3} y''_i}{q}, \quad \bar{x}'''_l = \frac{\sum_{i \in s_4} x'''_{li}}{r},$$

$$\bar{y}''' = \frac{\sum_{i \in s_4} y'''_i}{r}, \quad \text{for } l = 1, 2. \quad (2.1)$$

$$\bar{x}_{lA} = \frac{(n-p-q-r)\bar{x}_l + p\bar{x}'_l + r\bar{x}'''_l}{(n-q)} \quad (\text{pooled mean of } X_l \text{ based on } S_1 \cap S_2 \cap S_4),$$

$$\bar{y}_A = \frac{(n-p-q-r)\bar{y} + q\bar{y}'' + r\bar{y}'''}{(n-p)} \quad (\text{pooled mean of } Y \text{ based on } S_1 \cap S_3 \cap S_4).$$

$$f_0 = \frac{1}{(n-p-q-r)} - \frac{1}{N}, \quad f_1 = \frac{1}{(n-p)} - \frac{1}{N}, \quad f_2 = \frac{1}{(n-q)} - \frac{1}{N}, \quad f_3 = \frac{1}{(n-r)} - \frac{1}{N}, \quad (2.2)$$

$$\bar{Y} = N^{-1} \left(\sum_{i=1}^N Y_i \right), \quad \bar{X}_l = N^{-1} \left(\sum_{i=1}^N X_{li} \right), \quad S_0^2 = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N-1}, \quad S_l^2 = \frac{\sum_{i=1}^N (X_{li} - \bar{X}_l)^2}{(N-1)},$$

$$S_{0l} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})(X_{li} - \bar{X}_l)}{(N-1)}, \quad S_{12} = \frac{\sum_{i=1}^N (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{(N-1)}, \quad C_0 = \frac{S_y}{\bar{Y}}, \quad C_l = \frac{S_l}{\bar{X}_l}, \quad (2.3)$$

$$\rho_{0l} = \frac{S_{0l}}{S_0 S_l}, \quad \rho_{12} = \frac{S_{12}}{S_1 S_2}.$$

3. SOME AVAILABLE ESTIMATORS

Selected estimators of population mean and their corresponding MSE's are given as follows.

(i) The mean per unit estimator of population mean is:

$$t_0 = \bar{y}_A \quad (3.1)$$

The mean square error of t_0 is:

$$MSE(t_0) \cong \bar{Y}^2 f_1 C_0^2 \quad (3.2)$$

(ii) Sanaullah et al. [24] suggested exponential estimator-I

$$t_1 = \bar{y}_A \exp \left[\alpha \left(\frac{\bar{X}_2 - \bar{x}_{2A}}{\bar{X}_2 + \bar{x}_{2A}} \right) - (1 - \alpha) \left(\frac{\bar{X}_1 - \bar{x}_{1A}}{\bar{X}_1 + \bar{x}_{1A}} \right) \right] \quad (3.3)$$

where α is the optimization constant. The expression of mean square error is:

$$MSE(t_1) \approx \bar{Y}^2 \left[\begin{array}{l} f_1 \{ C_0^2 + (1 - \alpha) H_{01} C_1^2 - \alpha H_{02} C_2^2 \} \\ + f_2 \left\{ \frac{(1-\alpha)^2}{4} C_1^2 + \frac{\alpha^2}{4} C_2^2 - \frac{\alpha(1-\alpha)}{2} H_{12} C_2^2 \right\} \end{array} \right] \quad (3.4)$$

(iii) Vishwakarma and Kumar [31] exponential chain ratio-product type estimator:

$$t_2 = \bar{y}_A \left[k \exp \left\{ \frac{\left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2 - \bar{X}_1}{\left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2 + \bar{X}_1} \right\} + (1 - k) \exp \left\{ \frac{\bar{X}_1 - \left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2}{\bar{X}_1 + \left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2} \right\} \right] \quad (3.5)$$

where k is the optimization constant. The expression of mean square error is given by:

$$MSE(t_2) \approx \bar{Y}^2 \left[\begin{array}{l} f_1 \{ C_0^2 + (2k - 1) (H_{01} C_1^2 - H_{02} C_2^2) \} \\ + f_2 \frac{(2k-1)^2}{4} (C_1^2 + C_2^2 - 2H_{12} C_2^2) \end{array} \right] \quad (3.6)$$

(iv) Khan and Khan [13] proposed the following estimator

$$t_3 = \left\{ \left[\bar{y}_A + \alpha_1 (\bar{x}_{1A} - \bar{X}_1) \right] * \left[\alpha_2 \exp \left(\frac{\left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2 - \bar{X}_1}{\left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2 + \bar{X}_1} \right) + (1 - \alpha_2) \exp \left(\frac{\bar{X}_1 - \left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2}{\bar{X}_1 + \left(\frac{\bar{x}_{1A}}{\bar{x}_{2A}} \right) \bar{X}_2} \right) \right] \right\} \quad (3.7)$$

The mean square error of t_3 is given by:

$$MSE(t_3) \approx \bar{Y}^2 \left\{ \begin{array}{l} f_1 \left[C_0^2 + H_{01} C_1^2 (2\alpha_1 R_1 + (2\alpha_2 - 1)) - (2\alpha_2 - 1) H_{02} C_2^2 \right] \\ + f_2 C_1^2 \left[\alpha_1^2 R_1^2 + \frac{(2\alpha_2 - 1)^2}{4} + \alpha_1 (2\alpha_2 - 1) R_1 \right] + \\ f_2 C_2^2 \left[\frac{(2\alpha_2 - 1)^2}{4} - \frac{(2\alpha_2 - 1)^2}{2} H_{12} - \alpha_1 (2\alpha_2 - 1) R_1 H_{12} \right] \end{array} \right\} \quad (3.8)$$

(v) Riaz et al. [22] proposed a regression-cum-exponential estimator

$$t_4 = [a_1 \bar{y}_A + a_2 (\bar{X}_2 - \bar{x}_{2A})] \exp \left[\left(\frac{\bar{X}_1 - \bar{x}_{1A}}{\bar{X}_1 + \bar{x}_{1A}} \right) \right] \tag{3.9}$$

where α_1 and α_2 are real positive constants.

$$MSE(t_4) \approx \bar{Y}^2 \left[\begin{aligned} &(a_1 - 1)^2 + f_1 \{ \alpha_1^2 C_0^2 - a_1^2 H_{01} C_1^2 - 2a_1 a_2 R_2 H_{02} C_2^2 \} \\ &+ f_2 \left\{ \left(\frac{a_1}{2} \right)^2 C_1^2 + (a_2 R_2)^2 C_2^2 + a_1 a_2 R_2 H_{12} C_2^2 \right\} \end{aligned} \right] \tag{3.10}$$

$$t_5 = [\alpha_1 \bar{y}_A + \alpha_2 (\bar{X}_2 - \bar{x}_{2A})] \exp \left[- \frac{(\bar{X}_1 - \bar{x}_{1A})}{(\bar{X}_1 + \bar{x}_{1A})} \right] \tag{3.11}$$

$$MSE(t_5) \approx \bar{Y}^2 \left[\begin{aligned} &(a_1 - 1)^2 + f_1 \{ \alpha_1^2 C_0^2 + a_1^2 H_{01} C_1^2 - 2a_1 a_2 R_2 H_{02} C_2^2 \} \\ &+ f_2 \left\{ \left(\frac{a_1}{2} \right)^2 C_1^2 + (a_2 R_2)^2 C_2^2 - a_1 a_2 R_2 H_{12} C_2^2 \right\} \end{aligned} \right] \tag{3.12}$$

where $R_2 = \frac{\bar{X}_2}{\bar{Y}}$.

(vi) Muneer et al. [17] Estimator

$$t_6 = \left\{ \left[k_3 \bar{y}_A + k_4 (\bar{x}_{1A} - \bar{X}_1) \right] * \left[w \left(2 - \exp \left(\frac{\bar{x}_{2A} - \bar{X}_2}{\bar{x}_{2A} + \bar{X}_2} \right) \right) + (1-w) \exp \left(\frac{\bar{X}_2 - \bar{x}_{2A}}{\bar{X}_2 + \bar{x}_{2A}} \right) \right] \right\} \tag{3.13}$$

$$MSE(t_6) \approx \bar{Y}^2 \left\{ \begin{aligned} &(k_3 - 1)^2 + f_1 \left[k_3^2 C_0^2 - 2k_3 k_4 R_1 H_{01} C_1^2 - k_3^2 H_{02} C_2^2 \right] \\ &+ f_2 \left[k_4^2 R_1^2 C_1^2 + \frac{k_3^2}{4} C_2^2 + k_3 k_4 R_1 H_{12} C_2^2 \right] \end{aligned} \right\}. \tag{3.14}$$

4. PROPOSED ESTIMATOR

The suggested weighted ratio-cum-exponential factor-type estimator of finite population mean is based upon, $s_1 \cup s_2 \cup s_3 \cup s_4$, which is as follows:

$$t_p = \bar{y} \left\{ w \left[\frac{N_1}{D_1} \exp \left(\frac{N_2 - D_2}{N_2 + D_2} \right) \right] + (1 - w) \left[\frac{N_2}{D_2} \exp \left(\frac{N_1 - D_1}{N_1 + D_1} \right) \right] \right\}. \tag{4.1}$$

where $0 \leq w \leq 1$ and

$$\begin{aligned} N_l &= (A_l + C_l) \bar{X}_l + f B_l \bar{x}_{lA}, D_l = (A_l + f B_l) \bar{X}_l + C_l \bar{x}_{lA}, A_l = (k_l - 1)(k_l - 2), \\ B_l &= (k_l - 1)(k_l - 4), C_l = (k_l - 2)(k_l - 3)(k_l - 4), f = \frac{n}{N}, \\ \alpha_l &= \frac{f B_l}{A_l + f B_l + C_l}, \beta_l = \frac{C_l}{A_l + f B_l + C_l}, \theta_l = \alpha_l - \beta_l, \end{aligned} \tag{4.2}$$

where $l=1, 2$.

Using (4.2) in (4.1), we get:

$$t_p = \bar{y}_A \left[\begin{array}{l} w \left(\frac{\{(A_1 + C_1) \bar{X}_1 + f B_1 \bar{x}_{1A}\}}{\{(A_1 + f B_1) \bar{X}_1 + C_1 \bar{x}_{1A}\}} \right) * \\ \exp \left\{ \frac{\{(A_2 + C_2) \bar{X}_2 + f B_2 \bar{x}_{2A}\} - \{(A_2 + f B_2) \bar{X}_2 + C_2 \bar{x}_{2A}\}}{\{(A_2 + C_2) \bar{X}_2 + f B_2 \bar{x}_{2A}\} + \{(A_2 + f B_2) \bar{X}_2 + C_2 \bar{x}_{2A}\}} \right\} \\ + (1 - w) \left(\frac{\{(A_2 + C_2) \bar{X}_2 + f B_2 \bar{x}_{2A}\}}{\{(A_2 + f B_2) \bar{X}_2 + C_2 \bar{x}_{2A}\}} \right) * \\ \exp \left\{ \frac{\{(A_1 + C_1) \bar{X}_1 + f B_1 \bar{x}_{1A}\} - \{(A_1 + f B_1) \bar{X}_1 + C_1 \bar{x}_{1A}\}}{\{(A_1 + C_1) \bar{X}_1 + f B_1 \bar{x}_{1A}\} + \{(A_1 + f B_1) \bar{X}_1 + C_1 \bar{x}_{1A}\}} \right\} \end{array} \right]. \quad (4.3)$$

5. BIAS AND MSE OF SUGGESTED ESTIMATOR

To obtain expressions of bias and MSE of proposed estimator up to first degree of approximation, we define the following error terms:

$$\begin{aligned} e_0 &= \frac{\bar{y}}{Y} - 1, e_l = \frac{\bar{x}_l}{X_l} - 1, e'_l = \frac{\bar{x}'_l}{X'_l} - 1, e''_0 = \frac{\bar{y}''}{Y} - 1, e'''_0 = \frac{\bar{y}'''}{Y} - 1, e'''_l = \frac{\bar{x}'''}{X'_l} - 1. \\ E(e_0) &= E(e_l) = E(e'_l) = E(e''_0) = E(e'''_0) = E(e'''_l) = 0, E(e_0^2) = f_0 C_0^2, E(e_l^2) = f_0 C_l^2, \\ E(e_l'^2) &= (r^{-1} - N^{-1}) C_l^2, E(e_0 e_l) = f_0 \rho_{0l} C_0 C_l, E(e_1 e_2) = f_0 \rho_{12} C_1 C_2, \\ E(e'_1 e'_2) &= (r^{-1} - N^{-1}) \rho_{12} C_1, E(e''_0 e''_l) = (r^{-1} - N^{-1}) \rho_{0l} C_0 C_l, \\ E(e''_0 e''_l) &= (r^{-1} - N^{-1}) \rho_{12} C_1 C_2, E(e_0 e'_l) = E(e_0 e''_0) = E(e_0 e'''_0) = \\ E(e_0 e'''_0) &= E(e'_l e''_0) = E(e'_l e'''_0) = E(e''_0 e''_l) = E(e''_0 e'''_l) = 0, \\ H_{0l} &= \rho_{0l} \frac{C_0}{C_l}, H_{12} = \rho_{12} \frac{C_1}{C_2}. \end{aligned} \quad (5.1)$$

For bias and MSE of the suggested estimator, replacing \bar{y}_A and \bar{x}_{lA} by corresponding expressions from (2.1) and then using notations from (4.2), we get:

$$t_p = \bar{Y} \left\{ (1 + \Omega_0) \left[\begin{array}{l} w \left(\frac{\left(\frac{1 + \alpha_1 \Omega_1}{1 + \beta_1 \Omega_1} \right) \exp \left(\frac{\langle 1 + \alpha_2 \Omega_2 \rangle - \langle 1 + \beta_2 \Omega_2 \rangle}{\langle 1 + \alpha_2 \Omega_2 \rangle + \langle 1 + \beta_2 \Omega_2 \rangle} \right)}{\left(\frac{1 + \alpha_2 \Omega_2}{1 + \beta_2 \Omega_2} \right) \exp \left(\frac{\langle 1 + \alpha_1 \Omega_1 \rangle - \langle 1 + \beta_1 \Omega_1 \rangle}{\langle 1 + \alpha_1 \Omega_1 \rangle + \langle 1 + \beta_1 \Omega_1 \rangle} \right)} \right) \right] \right\}, \quad (5.2)$$

where $\Omega_0 = \left\{ \left(\frac{n-p-q-r}{n-p} \right) e_0 + \left(\frac{q}{n-p} \right) e''_0 + \left(\frac{r}{n-p} \right) e'''_0 \right\}$,

$\Omega_1 = \left\{ \left(\frac{n-p-q-r}{n-q} \right) e_1 + \frac{p}{n-q} e'_1 + \frac{r}{n-q} e''_1 \right\}$ and $\Omega_2 = \left\{ \left(\frac{n-p-q-r}{n-q} \right) e_2 + \frac{p}{n-q} e'_2 + \frac{r}{n-q} e''_2 \right\}$.

$$t_p - \bar{Y} = \bar{Y} \left(\Omega_0 + b_1 \Omega_1 + b_2 \Omega_2 + b_1 \Omega_0 \Omega_1 + b_2 \Omega_0 \Omega_2 + b_3 \Omega_1 \Omega_2 + b_4 \Omega_1^2 + b_5 \Omega_2^2 \right), \quad (5.3)$$

where

$$\begin{aligned} b_1 &= \frac{(1+w)}{2} \theta_1, b_2 = \frac{(2-w)}{2} \theta_2, b_3 = \frac{\theta_1 \theta_2}{2}, \\ b_4 &= (1-w) \left\{ \frac{\theta_1^2}{8} - \frac{\theta_1}{4} (\alpha_1 + \beta_1) \right\} - w \theta_1 \beta_1, \\ b_5 &= \left[w \left\{ \frac{\theta_2^2}{8} - \frac{\theta_2}{4} (\alpha_2 + \beta_2) \right\} - (1-w) \theta_2 \beta_2 \right]. \end{aligned} \quad (5.4)$$

Taking expectation of (5.3), the bias of t_p becomes:

$$\text{Bias}(t_p) \cong \bar{Y} \left[f_1 \left(b_1 H_{01} C_1^2 + b_2 H_{02} C_2^2 \right) + f_2 \left(b_3 H_{12} C_2^2 + b_4 C_1^2 + b_5 C_2^2 \right) \right], \quad (5.5)$$

Squaring (5.3), taking expectation, the MSE of t_p is:

$$MSE(t_p) \cong \bar{Y}^2 \left[\frac{f_1 (C_0^2 + 2b_1 H_{01} C_1^2 + 2b_2 H_{02} C_2^2) +}{f_2 (b_1^2 C_1^2 + b_2^2 C_2^2 + 2b_1 b_2 H_{12} C_2^2)} \right]. \quad (5.6)$$

For minimum MSE, differentiate (5.6) with respect to b_1 and b_2 respectively and equating to zero, b_{1m} and b_{2m} are:

$$b_{1m} = \frac{f_1 (H_{02} H_{12} C_2^2 - H_{01} C_1^2)}{f_2 (C_1^2 - H_{12}^2 C_2^2)}, \quad b_{2m} = \frac{f_1 C_1^2 (H_{01} H_{12} - H_{02})}{f_2 (C_1^2 - H_{12}^2 C_2^2)}. \quad (5.7)$$

Rearranging (5.7), θ_{1m} and θ_{2m} are given by:

$$\theta_{1m} = \frac{2f_1 (H_{02} H_{12} C_2^2 - H_{01} C_1^2)}{(1+w) f_2 (C_1^2 - H_{12}^2 C_2^2)}, \quad \theta_{2m} = \frac{2f_1 C_1^2 (H_{01} H_{12} - H_{02})}{(2-w) f_2 (C_1^2 - H_{12}^2 C_2^2)}. \quad (5.8)$$

In terms of characterization scalar k_l , (5.8) can be rearranged in the form:

$$\left[\begin{array}{l} (1+\theta_{1m}) k_1^3 + (f\theta_{1m} - f - 8\theta_{1m} - 9) k_1^2 + (23\theta_{1m} \\ - 5f\theta_{1m} + 5f + 26) k_1 + (4f\theta_{1m} - 4f - 22\theta_{1m} - 24) \end{array} \right] = 0, \\ \left[\begin{array}{l} (1+\theta_{2m}) k_2^3 + (f\theta_{2m} - f - 8\theta_{2m} - 9) k_2^2 + (23\theta_{2m} \\ - 5f\theta_{2m} + 5f + 26) k_2 + (4f\theta_{2m} - 4f - 22\theta_{2m} - 24) \end{array} \right] = 0. \quad (5.9)$$

Solving (5.9), gives multiple choices of k_1 and k_2 combinations, which result multiple choices of minimum mean square error.

6. COMPARISON

The suggested estimator is compared to the selected estimators on the basis of MSE and conditions are obtained under which the proposed estimator outperforms the selected estimators.

- (i) From (5.6) and (3.2), $MSE(t_p) < MSE(t_0)$, if $\omega_1 C_1^2 + \omega_2 C_1^2 + \omega_3 C_2^2 + \omega_4 C_2^2 < 0$, where $\omega_1 = 2f_1 b_1 H_{01}$, $\omega_2 = f_2 b_1^2$, $\omega_3 = 2f_1 H_{02}$ and $\omega_4 = f_2 (b_2^2 + 2b_1 b_2 H_{12})$.
- (ii) From (5.6) and (3.4), $MSE(t_p) < MSE(t_2)$, if $\lambda_1 C_1^2 + \lambda_2 C_1^2 + \lambda_3 C_2^2 + \lambda_4 C_2^2 < 0$, where $\lambda_1 = f_1 \{2b_1 - (1-\alpha)\} H_{01}$, $\lambda_2 = f_2 \left\{ b_1^2 - \frac{(1-\alpha)^2}{4} \right\}$, $\lambda_3 = f_1 (2b_2 + \alpha) H_{02}$, and $\lambda_4 = f_2 \left\{ \left(b_2^2 - \frac{\alpha^2}{4} \right) + \left(2b_1 b_2 + \frac{\alpha(1-\alpha)}{2} \right) H_{12} \right\}$.
- (iii) From (5.6) and (3.6), $MSE(t_p) < MSE(t_2)$, if $\pi_1 C_1^2 + \pi_2 C_1^2 + \pi_3 C_2^2 + \pi_4 C_2^2 < 0$, where $\pi_1 = f_1 \{2b_1 - (2k-1)\} H_{01}$, $\pi_2 = f_2 \left\{ b_1^2 - \frac{(2k-1)^2}{4} \right\}$, $\pi_3 = f_1 \{2b_2 + (2k-1)\} H_{02}$, and $\pi_4 = f_2 \left\{ \left(b_2^2 - \frac{(2k-1)^2}{4} \right) + \left(2b_1 b_2 + \frac{(2k-1)^2}{2} \right) H_{12} \right\}$.
- (iv) From (5.6) and (3.8), $MSE(t_p) < MSE(t_3)$, if $\phi_1 C_1^2 + \phi_2 C_1^2 + \phi_3 C_2^2 + \phi_4 C_2^2 < 0$, $\phi_1 = f_1 \{2b_1 - 2\alpha_1 R_1 - (2\alpha_2 - 1)\} H_{01}$,

$$\phi_2 = f_2 \left\{ b_1^2 - \alpha_1^2 R_1^2 - \frac{(2\alpha_2 - 1)^2}{4} - \alpha_1 (2\alpha_2 - 1) R_1 \right\}, \phi_3 = f_1 \{ 2b_2 + (2\alpha_2 - 1) \} H_{02},$$

$$\phi_4 = f_2 \left\{ \left(b_2^2 - \frac{(2\alpha_2 - 1)^2}{4} \right) + \left(2b_1 b_2 + \frac{(2\alpha_2 - 1)^2}{2} + \alpha_1 (2\alpha_2 - 1) R_1 \right) H_{12} \right\}$$

- (v) From (5.6) and (3.10), $MSE(t_p) < MSE(t_4)$, if
 $\{(\psi_1 C_0^2 + \psi_2 C_1^2 + \psi_3 C_1^2 + \psi_4 C_2^2 + \psi_5 C_2^2) - (\alpha_1 - 1)^2\} < 0$,
 where $\psi_1 = f_1 (1 - \alpha_1^2)$, $\psi_2 = f_1 (2b_1 + \alpha_1^2) H_{01}$, $\psi_3 = f_2 \left\{ b_1^2 - \left(\frac{\alpha_1}{2} \right)^2 \right\}$,
 $\psi_4 = 2f_1 (b_2 + \alpha_1 \alpha_2 R_2) H_{02}$ and $\psi_5 = f_2 \left\{ (b_2^2 - \alpha_2^2 R_2^2) + (2b_1 b_2 - \alpha_1 \alpha_2 R_2) H_{12} \right\}$.
- (vi) From (5.6) and (3.12), $MSE(t_p) < MSE(t_5)$, if
 $\{(\eta_1 C_0^2 + \eta_2 C_1^2 + \eta_3 C_1^2 + \eta_4 C_2^2 + \eta_5 C_2^2) - (\alpha_1 - 1)^2\} < 0$, where $\eta_1 = f_1 (1 - \alpha_1^2)$,
 $\eta_2 = f_1 (2b_1 - \alpha_1^2) H_{01}$, $\eta_3 = f_2 \left\{ b_1^2 - \left(\frac{\alpha_1}{2} \right)^2 \right\}$,
 $\eta_4 = 2f_1 (b_2 + \alpha_1 \alpha_2 R_2) H_{02}$ and $\eta_5 = f_2 \left\{ (b_2^2 - \alpha_2^2 R_2^2) + (2b_1 b_2 + \alpha_1 \alpha_2 R_2) H_{12} \right\}$.
- (vii) From (5.6) and (3.14), $MSE(t_p) < MSE(t_6)$, if
 $\{(\tau_1 C_0^2 + \tau_2 C_1^2 + \tau_3 C_1^2 + \tau_4 C_2^2 + \tau_5 C_2^2) - (k_3 - 1)^2\} < 0$,
 where $\tau_1 = f_1 (1 - k_3^2)$, $\tau_2 = 2f_1 (b_1 + k_3 k_4 R_1) H_{01}$, $\tau_3 = f_2 (b_1^2 - k_4^2 R_1^2)$,
 $\tau_4 = f_1 (2b_2 + k_3^2) H_{02}$ and $\tau_5 = f_2 \left\{ \left(b_2^2 - \frac{k_4^2}{4} \right) + (2b_1 b_2 - k_3 k_4 R_1) H_{12} \right\}$.

7. SIMULATION RESULTS AND DISCUSSION

To make the results more encouraging, simulation study has been conducted. From the following fixed populations, 10,000 samples were generated using (SRSWOR) scheme.

Population 1. Source: Murthy [19]

Y = wheat cultivated area in 1964, X_1 = wheat cultivated area in 1963, X_2 = Cultivated land in 1961.

$N = 34, \bar{Y} = 199.44, \bar{X}_1 = 208.89, \bar{X}_2 = 747.59, C_0^2 = 0.5673, C_1^2 = 0.5191, C_2^2 = 0.35274,$
 $\rho_{01} = 0.9801, \rho_{02} = 0.9043, \rho_{12} = 0.9097$.

Population 2. Source: Kutner et al. [16]

Y = No. of females employed, X_1 = No. of active physicians, X_2 = No. of females who are educated. $N = 61, \bar{Y} = 7.46, \bar{X}_1 = 5.31, \bar{X}_2 = 179, C_0^2 = 0.7103, C_1^2 = 0.7587, C_2^2 = 0.2515, \rho_{01} = 0.7373, \rho_{02} = -0.207, \rho_{12} = -0.0033$.

Population 3. Source: Kutner et al. [16]

Y = Total cost of claims by subscriber, X_1 = subscriber's age (year), X_2 = Treatment duration in days.

$N = 788, \bar{Y} = 2799.955, \bar{X}_1 = 58.72, \bar{X}_2 = 164.03, C_0^2 = 5.70931, C_1^2 = 0.013231, C_2^2 = 0.543397, \rho_{01} = -0.059, \rho_{02} = 0.176, \rho_{12} = 0.136$.

Under the proposed sampling plan and with different missingness rates in each variable, two different sample sizes (30% and 20% of the corresponding population sizes) were taken. For optimal mean square error (MSE), θ_{1m} and θ_{2m} are obtained in (5.8). Rearranging (5.8), two polynomials of degree 3 in k_1 and k_2 are obtained in (5.9). Solving equation (5.9), we get multiple choices for minimum MSE of the suggested estimator, which depend on the number of combinations of real roots of k_1 and k_2 . (see Table 1, 2 and 3, in Appendix). In population-1, at serial no. 5, 9 and 10 both k_1 and k_2 have three real roots so there are nine possibilities for minimum MSE in each case. For rest of the quadruplets (n, p, q, r)

observed, there is only one real and two imaginary roots for k_1 and all roots of k_2 are real so the estimator attains three minimum values of MSE in each case. For population 2, at serial no.8, k_1 has only one real and two complex roots and all roots of k_2 are real so there are three possible values of minimum MSE and for remaining combinations of quadruplet, all roots of k_1 and k_2 are real so there are nine possibilities for minimum MSE in each case. In population 3, k_1 has only one real and two complex roots for all quadruplets (n, p, q, r) examined, while all roots of k_2 are real so in each case, there are three possible values of minimum MSE. The percent relative efficiency (PRE) is computed by using the following expression:

$$PRE = Var(t_0)/Var(t_i) * 100,$$

for $i = 0, 1, 2, 3, 4, 5, 6$. and results are displayed in Tables 4, 5 and 6 in appendix.

8. CONCLUSION

A pooled ratio-cum-exponential factor-type estimator is proposed for estimating the population mean under non-response incorporating knowledge on two supplementary variables, and comparison of suggested estimator with the chosen estimators is made on the basis of percent relative efficiency. From Table 1, 2 and 3, (Appendix), number of real roots of k_1 and k_2 reveal that the selected estimator provides a wide range for optimal MSE, which depends on the number of combinations of real roots of the characterizing constant k_l ($l = 1, 2$). The compiled values of PRE in Table 4 (Appendix) reveal that the suggested estimator outperforms all the estimators chosen, for high positive correlations between variables (Population-1) under both large and small sample sizes and different missingness rates. Similarly results of PRE in Table 5 and 6 (Appendix) reveal that the suggested estimator outperforms under low correlation structures (Population-2 and 3) for both large and small sample sizes and different missingness rates. Thus, due to the dominant nature of the proposed estimator, it may be recommended for practical use.

APPENDIX

TABLE 1. Real Roots of k_1 & k_2

S.No	Population-I					
	SS*		MR**		Real Roots	
	n	p	q	r	k_1	k_2
1	14	3	3	1	2.417	1.89
					————	3.8826
					————	4.5206
2	14	3	2	1	2.459	1.3471
					————	3.6326
					————	4.0048
3	14	3	1	1	2.9905	1.605
					————	3.9521
					————	4.696
4	14	2	3	1	2.0236	1.1489
					————	4.014
					————	4.0679
5	14	1	3	1	1.252	1.0404
					2.0326	3.9324
					20.358	5.9947
6	10	3	3	1	2.0538	1.8819
					————	3.3938
					————	4.2241
7	10	3	2	1	3.1645	1.7774
					————	3.4736
					————	4.2457
8	10	3	1	1	2.5912	1.6309
					————	2.9118
					————	7.1187
9	10	2	3	1	1.0024	1.9801
					2.6624	3.2576
					46.3146	4.8124
10	10	1	3	1	2.2029	1.7764
					2.7004	3.9985
					12.7291	5.0065

*SS= Sample size

*MR=Missingness rate

TABLE 2. Real Roots of k_1 & k_2

S.No	Population-2					
	SS	MR			Real Roots	
	n	p	q	r	k_1	k_2
1	24	4	4	2	1.5935	1.5416
					2.9321	3.0708
					16.8056	6.6751
2	24	4	3	2	1.8923	1.2408
					2.8071	2.9008
					30.215	15.0743
3	24	4	2	2	-11.911	1.5687
					0.7704	3.1773
					2.8995	6.9621
4	24	3	4	2	1.0579	1.4447
					2.998	2.9012
					13.0561	9.8998
5	24	2	4	2	1.0468	1.8716
					1.9049	3.1094
					12.0641	5.4489
6	12	4	4	2	-4.7028	1.6773
					1.2117	2.7932
					2.5801	6.3889
7	12	4	3	2	-6.3356	1.699
					0.7811	2.0912
					2.5002	5.5432
8	12	4	2	2	2.4333	1.0143
					————	2.5674
					————	7.4151
9	12	3	4	2	1.323	1.7439
					2.328	2.2366
					14.5108	5.1487
10	12	2	4	2	1.4884	1.7556
					2.2208	2.5112
					9.2163	5.1003

TABLE 3. Real Roots of k_1 & k_2

S.No	Population-3					
	SS		MR		Real Roots	
	n	p	q	r	k_1	k_2
1	315	49	49	25	2.9408	-56.8404
					————	1.961
					————	2.8347
2	315	49	48	25	2.21644	-25.2419
					————	1.9853
					————	2.8453
3	315	49	47	25	2.039	-6.2396
					————	0.8991
					————	2.8629
4	315	48	49	25	2.2952	1.0756
					————	3.8895
					————	50.8061
5	315	47	49	25	2.0341	-24.3362
					————	0.3227
					————	2.1712
6	158	49	49	25	1.9035	1.6588
					————	2.1808
					————	5.0833
7	158	49	48	25	2.0926	1.102
					————	2.5007
					————	19.0513
8	158	49	47	25	2.0806	1.958
					————	2.3201
					————	15.0032
9	158	48	49	25	2.0914	1.4461
					————	2.5774
					————	60.3952
10	158	47	49	25	2.0918	1.3554
					————	2.7354
					————	9.6115

TABLE 4. PRE of the selected estimators with respect to mean per unit estimator for Population-1

S.No	SS	MR	Percentage Relative Efficiency									
			n	p	q	r	t_o	t_1	t_2	t_3	t_4	t_5
1	12	3	3	1	100	90.7	98.5	99.8	106.7	102.7	103.2	155.6
2	12	3	2	1	100	105.5	103.7	117.6	119.3	100.7	116.5	168.2
3	12	3	1	1	100	118.1	133.3	107.3	129.9	133.4	107.3	202.3
4	12	2	3	1	100	117.5	130.5	107.7	130.6	127.3	125.8	230.2
5	12	1	3	1	100	135.4	113.4	145.6	141.1	106.4	137.4	336.8
6	10	3	3	1	100	73.6	93.8	93.2	90.8	81.1	91.4	119.3
7	10	3	2	1	100	82.1	97.5	105.8	96.3	97.8	98.5	130.2
8	10	3	1	1	100	103	101.7	117	124.1	100.6	123.5	178.1
9	10	2	3	1	100	96.1	93.8	115.5	110.3	104.2	113.2	173.8
10	10	1	3	1	100	127.2	110.5	131.6	128.2	94	122.1	336.1

TABLE 5. PRE of the selected estimators with respect to mean per unit estimator for Population 2

S.No	SS	MR	p	q	r	Percentage Relative Efficiency	t_0	t_1	t_2	t_3	t_4	t_5	t_6	t_p
1	18	4	4	4	2	100	103.6	105.3	102.6	107.4	62.8	102.7	230.3	
2	18	4	3	2		100	103.1	107.2	106.3	113.5	63.5	109.7	204.7	
3	18	4	2	2		100	99.9	108.7	119.5	129.7	71.8	123.8	176.9	
4	18	3	4	2		100	105.2	110.3	111.5	113.8	61.5	109	226.7	
5	18	2	4	2		100	103.7	109.3	124.8	128.6	66.4	125.6	205.5	
6	12	4	4	2		100	62.9	76.3	82.8	85.7	76	93.9	104.8	
7	12	4	3	2		100	72.2	94	88.5	106.6	69.4	97.6	110.2	
8	12	4	2	2		100	97.9	101.1	88.2	102.8	62.8	98.9	203.3	
9	12	3	4	2		100	80.3	98.1	99	103.8	64.8	102.1	129.3	
10	12	2	4	2		100	104.6	108.3	104.2	100.2	60.4	96.4	293.2	

TABLE 6. PRE of the selected estimators with respect to mean per unit estimator for Population 3

S.No	SS	MR	Percentage Relative Efficiency									
			n	p	q	r	t_o	t_1	t_2	t_3	t_4	t_5
1	236	49	49	25	100	100.6	101.4	100.4	91.7	92.9	92	164.2
2	236	49	48	25	100	100.9	102.9	101.2	113.5	115.4	114.2	174.9
3	236	49	47	25	100	100.7	101.8	101	113.3	114.8	113.9	175.3
4	236	48	49	25	100	100.8	101.8	101	113.6	114.3	112.8	175.8
5	236	47	49	25	100	100.6	101.6	100.6	91.1	92.4	91.7	168.4
6	158	49	49	25	100	100.1	100.8	99.8	90.9	92	91.2	132.9
7	158	49	48	25	100	100	100.7	99.7	92	93.2	92.5	129
8	158	49	47	25	100	99.5	100.6	99.3	119.3	120.8	119.5	139.4
9	158	48	49	25	100	99.9	100.9	99.7	92	93.1	92.4	129.3
10	158	47	49	25	100	99.6	100.3	99.2	119.4	120.6	119.1	140.1

REFERENCES

- [1] Z. Ahmad, M. Hanif and M. Ahmad, *Generalized multi-phase multivariate ratio estimators for partial information case using multi-auxiliary variables*, Commun. Korean Statist Soc. **17**, No. 5 (2010) 625-637.
- [2] A. Ahmad, A. Sanaullah and M. Hanif, *Modified maximum likelihood integrated robust ratio estimator in simple random sampling*, Punjab Univ. j. math. **51**, No. 3 (2019) 133-145.
- [3] A. Audu and A. A. Adewara, *A modified factor-type estimator under two-phase sampling*, Punjab Univ. j. math. **49**, No. 2 (2017) 59-73.
- [4] M. Azeem and M. Hanif, *Joint influence of measurement error and non-response on estimation of population mean*, Communications in Statistics-Theory and Methods **46**, No. 4 (2017) 1679-1693.
- [5] M. K. Chaudhary and A. Kumar, *Estimation of mean of finite population using double sampling scheme under non-response*, J. Stat. Appl. Pro. **5**, No. 2 (2016) 287-297.
- [6] W. Cochran, *The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce*, The Journal of Agricultural Science **30**, No. 2 (1940) 262-275.
- [7] W. A. A.-Dayyeh, M. S. Ahmad, R. A. Ahmad and H. A. Muttlak, *Some estimators of a finite population mean using auxiliary information*, Appl. Math. Comput. **139**, (2003) 287-298.
- [8] L. K. Grover and P. Kaur, *Exponential ratio type estimators of population mean under non-response*, Open Journal of Statistics **4**, (2014) 97-110.
- [9] G. Gupta and J. Shabbir, *On improvement in estimating the population mean in simple random sampling*, J. Appl Stat. **35**, No. 5 (2008a) 559-566.
- [10] G. Gupta and J. Shabbir, *Variance estimation in simple random sampling using auxiliary information*, Hacettepe J. Math Stat. **37**, No. 1 (2008b) 57-67.
- [11] M. H. Hansen and W. N. Hurwitz, *The problem of non-response in sampling surveys*, Journal of the American Statistical Association **41**, (1946) 517-529.
- [12] C. Kadilar and H. Cingi, *Ratio estimators in simple random sampling*, Appl. Math. and Comput. **151**, (2004) 893-902.
- [13] L. Khan and J. S. Khan, *An efficient class of estimators of population mean in two-phase sampling using two auxiliary variables*, Pak. J. Stat. Oper. Res. **11**, No. 3 (2015) 397-408.
- [14] B. B. Khare and R. R. Sinha, *On class of estimators for population mean using multi-auxiliary character in presence of non-response*, Statistics in Transition-new series **10**, (2009) 3-14.

- [15] S. Kumar and M. Viswanathaiah, *A Generalized family of transformed estimators in the presence of non-response in sample surveys*, Columbia International Publishing Journal of Advanced Computing **2**, No. 3 (2013) 99-109.
- [16] M. H. Kutner, C. Nachtsheim and J. Neter, *Applied linear regression models*, 4th edition, McGraw-Hill/Irwin 1989.
- [17] S. Muneer, J. Shabbir and A. Khalil, *Estimation of finite population mean in simple random sampling and stratified random sampling using two auxiliary variables*, Commun. Stat. Theory Methods **45**, No. 5 (2017) 2181-2192.
- [18] M. N. Murthy, *Product method of estimation*, Sankhya, Series A. **26**, (1964) 294-307.
- [19] M. N. Murthy, *Sampling theory and methods*, Statistical Publishing Society: Calcutta, India (1967).
- [20] M. Noor-ul-Amin, M. Tayyab and M. Hanif, *Mean estimation using even order ranked set sampling*, Punjab Univ. j. math. **51**, No. 1 (2019) 91-99.
- [21] O. Ozturk, *Ratio estimators based on a ranked set sample in a finite population setting*, Journal of the Korean Statistical Society 2018.
- [22] N. Riaz, M. F. Tabasum, M. Hanif, N. Amin and A. Sana, *Generalized exponential type regression estimator by using two auxiliary information*, Sci. Int. (Lahore) **27**, No. 1 (2015) 383-388.
- [23] D. S. Robson, *Application of multivariate polykays to the theory of unbiased ratio-type estimation*, The Journal of the American Statistical Association **59**, (1957) 1225-1226.
- [24] A. Sanaullah, H. Khan, H. A. Ali and R. Singh, *Improved exponential ratio-type estimators in survey sampling*, Journal of Reliability and Statistical Studies **5**, No. 2 (2012) 119-132.
- [25] J. Shabbir and S. Gupta, *On estimating finite population mean in simple and stratified random sampling*, Commun. Stat. Theory Methods **40**, (2010) 199-212.
- [26] D. Shukla, *F-T estimator under two-phase sampling*, Metron **59**, No. 1-2 (2002) 253-263.
- [27] D. Shukla, S. Pathak and N. S. Thakur, *Estimation of population mean using two auxiliary sources in sample surveys*, Statistics in Transition **13**, No. 1 (2012) 21-36.
- [28] S. K. Srivastava and H. S. Jhaji, *A class of estimators of the population mean in survey sampling using auxiliary information*, Biometrika **68**, (1981) 341-343.
- [29] T. P. Tripathy, *A general class of estimators of population ratio*, Sankhya Series C. **42**, (1980) 45-63.
- [30] H. K. Verma and R. Singh, *A family of efficient estimator in circular systematic sampling*, Journal of Advanced Computing **3**, No. 2 (2014) 56-68.

- [31] G. K. Vishwakarma and M. Kumar, *An improved class of chain ratio-product type estimators in two-phase sampling using two auxiliary variables*, Journal of Probability and Statistics Article ID 939701 (2014).