

Estimating the Problem of Non-Response and Measurement Error in Sample Survey

Sunil Kumar^{1,*}, Sandeep Bhougale², Vishwantra Sharma¹, Rahul Gupta¹ and J. P. Singh Joorel³

¹ Department of Statistics, University of Jammu, Jammu, India

² Department of Mathematics, Shri Mata Vaishno Devi, University, Kakryal, Jammu, India

³ Information and Library Network Centre, Gandhinagar, Gujrat, India

Received: 30 Apr. 2019, Revised: 19 Mar. 2020, Accepted: 17 Apr. 2020

Published online: 1 Nov. 2021

Abstract: In this paper, the problem of estimation of finite population mean of the study variable is discussed in the presence of non-response and measurement error using the auxiliary variable. Some realistic conditions have been obtained under which the proposed estimator is more efficient than usual unbiased estimator, ratio estimators and product estimators. An empirical study is also conducted to support the theoretical findings in different situations.

Keywords: Study Variable, Auxiliary Variable, Non-response, Measurement error

1 Introduction

Sample surveys are conducted to obtain data on a variety of matters in many fields of life. The use of auxiliary information which is correlated with the variable of interest may improve the efficiency of the estimators. If there is a positive correlation between the auxiliary variable in the study and the variable under study, the ratio method of estimation is used Cochran [1,2]. The product estimation method is used in situations where the correlation coefficient between the variable of interest and the auxiliary variable is negative Robson [3] and Murthy [4]. A survey usually encounters various technical difficulties. No survey is perfect in all regards. Generally errors are of two categories sampling and non-sampling errors. Sampling errors comprise of the differences between the sample and the population due solely to the particular units that have been selected. Non-sampling errors encompass all other things that contribute to survey error. Non-sampling errors are said to arise from wrongly conceived definitions, imperfections in the tabulation plans, failure to obtain response from all sample members, and so on (see, Ilves [5], Groves, [6]).

In practice, this ideal is not met and the researcher faces the problem of measurement error while collecting information from the individuals. Measurement error is the difference between the value which is recorded and the true value of a variable in the study. Many researchers, such as Cochran [7,8,9], Fuller [10], Shalabh [11], Manisha and Singh [12,13], Wang [14], Allen et al. [15], Singh and Karpe [16,17,18,19], Salas and Gregoire [20], Kumar et al. [21] and Shukla et al. [22] etc., have studied measurement errors.

In sample surveys, the term non-response refers to the failure to collect information from one or more respondents on one or more variables. The reasons why non-response occurs include non-availability of the respondents at home, refusal to answer the questionnaire, lack of information, etc. The problem of non-response was first studied by Hansen and Hurwitz [23]. They addressed incomplete samples in mail sample survey and estimated the sample mean of the responding individuals and the sample mean of the sub-sample drawn from the non-respondents. Other researchers who have studied non-response include El-Badry [24], Foradari [25], Srinath [26], Cochran [9], Rao [27,28,29], Khare and Srivastava [30,31,32,33], Sodipo and Obisesan [34], Singh and Kumar [35,36,37,38,39], Ismail et al. [40], Khare et al. [41,42], Kumar and Bhougale [43], Khare and Kumar [44], Singh et al. [45], Shabbir and Khan [46], Kumar [47], Kumar and Chatterjee [48], Sharma and Kumar [49] and many others.

* Corresponding author e-mail: sunilbhougale06@gmail.com

In practice, the researchers face the problem of measurement error and non-response while collecting information from individuals. Researchers who studied non-response have ignored the presence of possible measurement errors and vice – versa. Jackman [50] dealt with both non-response and measurement error simultaneously in the case of voter turnout. Furthermore, Dixon [51] studied the estimation of non-response bias and measurement error on the data from Consumer Expenditure Quarterly Interview Survey (CEQ), Current Population Survey (CPS) and National Health Interview Survey (NHIS), as an attempt to measure the differences in employment status of Washington. Azeem [52] suggested estimators for estimating the population mean of study variable in the presence of non-response and measurement error. Moreover, Kumar et al. [53] and Kumar and Choudhary [54] have suggested estimator for estimating the population mean of the study variable in the presence of non-response and measurement error in the study as well as the auxiliary variable. In this article, we have suggested a dual to Sahai's [55] estimator for estimating the population mean of the study variable in the presence of non-response and measurement error as well as auxiliary variable. An empirical study is conducted to judge the performance of the proposed estimator over other estimators of the population mean of the study variable.

2 Notations and Sampling Procedure

Most of the surveys yield estimates that are suspected to be biased, because of the presence of non-response and/or measurement bias. Numerous literature exist for eliminating the effect of non-response and measurement error independently in the process of estimating the population mean of the study variable Y . Cochran [2] established the use of auxiliary information to improve the precision of estimators of the population parameters.

Let us consider a population $U = U_1, U_2, \dots, U_N$ of N units. A simple random sample of size n is selected from the population without replacement on study variable Y and auxiliary variable X . In the situation, when there is presence of non-response and measurement error associated with the study variable $U_i^* = y_i^* - Y_i^*$, and in the presence of non-response on the auxiliary variable, let the measurement error associated with auxiliary variable be $V_i^* = x_i^* - X_i^*$. The measurement errors are random in nature and have mean zero and variances σ_U^2 and σ_V^2 respectively for the responding units and $\sigma_{U(2)}^2$ and $\sigma_{V(2)}^2$ respectively for the non-responding units of the population.

The classical ratio and product estimators for the population mean μ_y of the study variable y in the presence of non-response and measurement error are defined as

$$t_R = \hat{\mu}_y^* \left(\frac{\mu_x}{\hat{\mu}_x^*} \right) \quad (1)$$

and

$$t_P = \hat{\mu}_y^* \left(\frac{\hat{\mu}_x^*}{\mu_x} \right) \quad (2)$$

where $\hat{\mu}_y^*$ and $\hat{\mu}_x^*$ are the sample means of study and auxiliary variable in the presence of non-response and measurement error respectively, and μ_x is the population mean of the auxiliary variables.

Consider the transformation

$$x'_i = (1 + g)\mu_x - gx_i; i = 1, 2, 3, \dots, N, \quad (3)$$

where $g = \frac{n}{(N-n)}$.

Then

$$\hat{\mu}'_x = (1 + g)\mu_x - g\hat{\mu}_x^*, \quad (4)$$

is an unbiased estimator for the population mean μ_x of the auxiliary variable X and the correlation between μ_y and $\hat{\mu}'_x$ is negative. Using the transformation given in equation(3), Srivenkataramana [56] obtained dual to ratio estimator for the population mean μ_y as

$$t'_R = \hat{\mu}_y^* \left(\frac{\hat{\mu}'_x}{\mu_x} \right) \quad (5)$$

and

$$t'_P = \hat{\mu}_y^* \left(\frac{\mu_x}{\hat{\mu}'_x} \right). \quad (6)$$

In the present study, we have suggested a dual to Sahai's [34] estimator for estimating the population mean μ_y of the study variable y in the presence of non-response and measurement error in study as well as auxiliary variable. To the first degree of approximation, we have obtained the bias and mean squared error (MSE) of the proposed estimator. Furthermore, the conditions in which the proposed estimator is more efficient than the other existing estimators are obtained.

3 Proposed Dual to Sahai's Estimator

The following is the suggested dual to Sahai's [34] estimator for the population mean μ_y as

$$\hat{t}_s = \hat{\mu}_y^* \left(\frac{\mu_x + \delta \hat{\mu}_x}{\hat{\mu}_x + \delta \mu_x} \right), \tag{7}$$

where $\mu_x = \frac{N\mu_x - n\hat{\mu}_x^*}{N-n}$, and δ is a scalar used as the design parameter. It is interesting to note that

- a) For $\delta = 1, \hat{t}_s = \hat{\mu}_y^*$, usual unbiased estimator.
- b) For $\delta = 0, \hat{t}_s = \hat{\mu}_y^* \left(\frac{\mu_x}{\hat{\mu}_x} \right) = \hat{t}_P$, dual to product estimator.
- c) For δ very large, $\hat{t}_s = \hat{\mu}_y^* \left(\frac{\hat{\mu}_x}{\mu_x} \right) = \hat{t}_R$, dual to ratio estimator,

i.e.

$$\lim_{\delta \rightarrow \infty} \hat{t}_s = \lim_{\delta \rightarrow \infty} \hat{\mu}_y^* \left(\frac{\mu_x + \delta \hat{\mu}_x}{\hat{\mu}_x + \delta \mu_x} \right) = \hat{\mu}_y^* \lim_{\delta \rightarrow \infty} \left(\frac{\mu_x + \delta \hat{\mu}_x}{\hat{\mu}_x + \delta \mu_x} \right) \cong \hat{\mu}_y^* \left(\frac{\mu_x}{\hat{\mu}_x} \right) = \hat{t}_R.$$

To obtain the expressions of bias and MSE of the proposed estimator, let us assume the following

$$\omega_y^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n (y_i^* - \mu_Y), \omega_U^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n U_i^*, \omega_x^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i^* - \mu_X) \text{ and } \omega_V^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n V_i^*.$$

Adding ω_y^* and ω_U^* , we have $\omega_y^* + \omega_U^* = \frac{1}{\sqrt{n}} [\sum_{i=1}^n (y_i^* - \mu_Y) + \sum_{i=1}^n U_i^*]$.

Multiplying both sides by $\frac{1}{\sqrt{n}}$, we have

$$\frac{1}{\sqrt{n}} (\omega_y^* + \omega_U^*) = \left[\frac{1}{n} \sum_{i=1}^n (y_i^* - \mu_Y) + \frac{1}{n} \sum_{i=1}^n (y_i^* - Y_i^*) \right]$$

or

$$\frac{1}{\sqrt{n}} (\omega_y^* + \omega_U^*) = \hat{\mu}_y^* - \mu_Y,$$

or

$$\hat{\mu}_y^* = \mu_Y + \frac{1}{\sqrt{n}} (\omega_y^* + \omega_U^*) = \mu_Y + \omega_Y.$$

Similarly, one can obtain

$$\hat{\mu}_x^* = \mu_X + \frac{1}{\sqrt{n}} (\omega_x^* + \omega_V^*) = \mu_X + \omega_X.$$

Further

$$E \left(\frac{\omega_y^* + \omega_U^*}{\sqrt{n}} \right)^2 = \lambda_2 (\sigma_y^2 + \sigma_U^2) + \theta (\sigma_{y(2)}^2 + \sigma_{U(2)}^2)$$

$$E \left(\frac{\omega_x^* + \omega_V^*}{\sqrt{n}} \right)^2 = \lambda_2 (\sigma_x^2 + \sigma_V^2) + \theta (\sigma_{x(2)}^2 + \sigma_{V(2)}^2) = B_o(\text{say})$$

$$E \left[\left(\frac{\omega_y^* + \omega_U^*}{\sqrt{n}} \right) \left(\frac{\omega_x^* + \omega_V^*}{\sqrt{n}} \right) \right] = \lambda_2 \rho_{yx} \sigma_y \sigma_x + \theta \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} = A_o(\text{say}),$$

where $\mu_Y = \frac{1}{N} \sum_{i=1}^N y_i, \mu_X = \frac{1}{N} \sum_{i=1}^N x_i, \sigma_y^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \mu_Y)^2$ and $\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_X)^2$ denote the population mean and the population variance of the study variable y and auxiliary variable x. Let $\mu_{Y_1} = \frac{1}{N_1} \sum_{i=1}^{N_1} y_i$ and $\sigma_{y(1)}^2 = \frac{1}{(N_1-1)} \sum_{i=1}^{N_1} (y_i - \mu_{Y_1})^2$ denote the mean and variance of the response group. Similarly, let $\mu_{Y_2} = \frac{1}{N_2} \sum_{i=1}^{N_2} y_i$ and $\sigma_{y(2)}^2 = \frac{1}{(N_2-1)} \sum_{i=1}^{N_2} (y_i - \mu_{Y_2})^2$ denote the mean and variance of the non-response group of the study variable. Similarly, for auxiliary variable, $\mu_{X_1} = \frac{1}{N_1} \sum_{i=1}^{N_1} x_i, \sigma_{x(1)}^2 = \frac{1}{(N_1-1)} \sum_{i=1}^{N_1} (x_i - \mu_{X_1})^2, \mu_{X_2} = \frac{1}{N_2} \sum_{i=1}^{N_2} x_i$ and $\sigma_{x(2)}^2 = \frac{1}{(N_2-1)} \sum_{i=1}^{N_2} (x_i - \mu_{X_2})^2$ are the mean and variance of response and non-response group of auxiliary variable, respectively. Let $\rho_{yx} = \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_Y)(x_i - \mu_X)$ and $\rho_{yx(2)} = \frac{1}{N_2-1} \sum_{i=1}^{N_2} (y_i - \mu_Y)(x_i - \mu_X)$ are the coefficient of correlation between study and auxiliary variables for response and non-response, respectively; $\lambda_2 = \left(\frac{1}{n} - \frac{1}{N} \right); \theta = \frac{W_2(k-1)}{n}; W_1 = \frac{N_1}{N}$ and $W_2 = \frac{N_2}{N}; k$ is the inverse sampling ratio, $k > 1$. For bias of the proposed estimator \hat{t}_s , we have

$$\hat{\mu}_x = \frac{N\mu_x - n\hat{\mu}_x^*}{N-n} = \frac{N}{N-n} \mu_x - \frac{n}{N-n} \hat{\mu}_x^*$$

Let $g = \frac{n}{N-n} \Rightarrow g + 1 = \frac{N}{N-n}$, then

$$\hat{\mu}_x = (g + 1)\mu_x - g\hat{\mu}_x^*.$$

Thus, the proposed estimator \hat{t}_s under the transformation becomes

$$\hat{t}_s = \hat{\mu}_y^* \left[\frac{\mu_x + \delta \{ (1+g)\mu_x - g\hat{\mu}_x^* \}}{(1+g)\mu_x - g\hat{\mu}_x^* + \delta\mu_x} \right]. \quad (8)$$

Now, expressing equation (8) in terms of $\omega_i; i = x, y, U, V$; we have

$$\hat{t}_s = (\mu_y + \omega_y) \left[1 - \left(\frac{g\delta}{1+\delta} \right) \frac{\omega_x}{\mu_x} \right] \left[1 - \left(\frac{g}{1+\delta} \right) \frac{\omega_x}{\mu_x} \right]^{-1}. \quad (9)$$

Assume that $\left| \left(\frac{g\delta}{1+\delta} \right) \frac{\omega_x}{\mu_x} \right| < 1$ so that $\left[1 - \left(\frac{g}{1+\delta} \right) \frac{\omega_x}{\mu_x} \right]^{-1}$ is expandable.

Expanding the right hand side of equation (9), one can obtain

$$\hat{t}_s = (\mu_y + \omega_y) \left\{ 1 - \left(\frac{g\delta}{1+\delta} \right) \frac{\omega_x}{\mu_x} \right\} \left\{ 1 + \left(\frac{g}{1+\delta} \right) \frac{\omega_x}{\mu_x} - \left(\frac{g}{1+\delta} \right)^2 \frac{\omega_x^2}{\mu_x^2} + \dots \right\}$$

Neglecting terms of $\omega_i; i = x, y$, having power greater than two, we have

$$\hat{t}_s - \mu_y \cong \omega_y + \left(\frac{g}{1+\delta} \right) (1-\delta) \frac{\omega_x}{\mu_x} - \left(\frac{g}{1+\delta} \right)^2 \frac{\omega_x^2}{\mu_x^2} (\mu_y - \delta) + \left(\frac{g}{1+\delta} \right) (1-\delta) \frac{\omega_x \omega_y}{\mu_x}. \quad (10)$$

Assuming that $\Delta = \frac{1-\delta}{1+\delta} \Rightarrow \frac{1}{1+\delta} = \frac{1+\Delta}{2}$ in equation (10), we have

$$\hat{t}_s - \mu_y \cong \omega_y + g\Delta \frac{\omega_x}{\mu_x} - \frac{g^2(1+\Delta)^2}{4} \frac{\omega_x^2}{\mu_x^2} (\mu_y - \delta) + g\Delta \frac{\omega_x \omega_y}{\mu_x}. \quad (11)$$

Taking expectation of both sides of equation (11), one can obtain the bias of \hat{t}_s to the first degree of approximation

$$B(\hat{t}_s) = E(\hat{t}_s - \mu_y)$$

$$B(\hat{t}_s) = -\frac{g^2(1+\Delta)^2}{4} \frac{(\mu_y - \delta)}{\mu_x^2} B_o + \frac{g\Delta}{\mu_x} A_o. \quad (12)$$

Squaring both sides of equation (11) to the first degree of approximation, the MSE of \hat{t}_s is given by

$$(\hat{t}_s - \mu_y)^2 = \left(\omega_y + g\Delta \frac{\omega_x}{\mu_x} \right)^2 = \omega_y^2 + \frac{g^2\Delta^2}{\mu_x^2} \omega_x^2 + 2\frac{g\Delta}{\mu_x} \omega_y \omega_x. \quad (13)$$

Taking expectation on both sides of equation (13), we get the MSE of \hat{t}_s to the first degree of approximation as

$$\begin{aligned} MSE(\hat{t}_s) = & \left[\lambda_2 \left\{ \sigma_y^2 + \frac{g^2\Delta^2}{\mu_x^2} \sigma_x^2 + 2\frac{g\Delta}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + \frac{g^2\Delta^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 \right. \right. \\ & \left. \left. + \frac{g^2\Delta^2}{\mu_x^2} \sigma_{x(2)}^2 + 2\frac{g\Delta}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + g^2\Delta^2 \mu_x^2 \sigma_{V(2)}^2 \right) \right\} \right]. \quad (14) \end{aligned}$$

The MSE of \hat{t}_s is minimized when

$$\Delta = -\frac{\mu_x A_o}{g B_o} = \Delta_o(\text{say}). \quad (15)$$

Substitute the optimum value of Δ from (15) in (14) to obtain the optimum MSE of \hat{t}_s as

$$\begin{aligned} \min. MSE(\hat{t}_s) = & \left[\lambda_2 \left\{ \sigma_y^2 + \frac{A_o^2}{B_o^2} \sigma_x^2 - 2\frac{A_o}{B_o} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + \frac{A_o^2}{B_o^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 \right. \right. \\ & \left. \left. + \frac{A_o^2}{B_o^2} \sigma_{x(2)}^2 - 2\frac{A_o}{B_o} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + \frac{A_o^2}{B_o^2} \sigma_{V(2)}^2 \right) \right\} \right]. \quad (16) \end{aligned}$$

4 Efficiency Comparison

The MSE of the mentioned estimator are as follows

$$MSE(\hat{\mu}_y^*) = \lambda_2(\sigma_y^2 + \sigma_U^2) + \theta(\sigma_{y(2)}^2 + \sigma_{U(2)}^2), \tag{17}$$

$$MSE(t_R) = \left[\lambda_2 \left\{ \sigma_y^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_x^2 - 2 \frac{\mu_y}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_{x(2)}^2 - 2 \frac{\mu_y}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_{V(2)}^2 \right) \right\} \right], \tag{18}$$

$$MSE(t_P) = \left[\lambda_2 \left\{ \sigma_y^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_x^2 + 2 \frac{\mu_y}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_{x(2)}^2 + 2 \frac{\mu_y}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + \frac{\mu_y^2}{\mu_x^2} \sigma_{V(2)}^2 \right) \right\} \right], \tag{19}$$

$$MSE(t'_R) = \left[\lambda_2 \left\{ \sigma_y^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_x^2 - 2g \frac{\mu_y}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_{x(2)}^2 - 2g \frac{\mu_y}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_{V(2)}^2 \right) \right\} \right], \tag{20}$$

$$MSE(t'_P) = \left[\lambda_2 \left\{ \sigma_y^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_x^2 + 2g \frac{\mu_y}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_{x(2)}^2 + 2g \frac{\mu_y}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + g^2 \frac{\mu_y^2}{\mu_x^2} \sigma_{V(2)}^2 \right) \right\} \right], \tag{21}$$

$$MSE(\hat{t}_R) = \left[\lambda_2 \left\{ \sigma_y^2 + \frac{g^2}{\mu_x^2} \sigma_x^2 - 2 \frac{g}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + \frac{g^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 + \frac{g^2}{\mu_x^2} \sigma_{x(2)}^2 - 2 \frac{g}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + \frac{g^2}{\mu_x^2} \sigma_{V(2)}^2 \right) \right\} \right], \tag{22}$$

$$MSE(\hat{t}_P) = \left[\lambda_2 \left\{ \sigma_y^2 + \frac{g^2}{\mu_x^2} \sigma_x^2 + 2 \frac{g}{\mu_x} \rho_{yx} \sigma_y \sigma_x + \left(\sigma_U^2 + \frac{g^2}{\mu_x^2} \sigma_V^2 \right) \right\} + \theta \left\{ \sigma_{y(2)}^2 + \frac{g^2}{\mu_x^2} \sigma_{x(2)}^2 + 2 \frac{g}{\mu_x} \rho_{yx(2)} \sigma_{y(2)} \sigma_{x(2)} + \left(\sigma_{U(2)}^2 + \frac{g^2}{\mu_x^2} \sigma_{V(2)}^2 \right) \right\} \right]. \tag{23}$$

From (17-23) and equation (14), one can obtain the following

$$MSE(\hat{\mu}_y^*) - MSE(\hat{t}_s) \geq 0 \tag{24}$$

$$\text{if } \Delta \leq -\frac{2\mu_x A_o}{gB_o},$$

$$MSE(t_R) - MSE(\hat{t}_s) \geq 0 \tag{25}$$

$$\text{if } \frac{-2A_o + \mu_x B_o}{gB_o} \leq \Delta \leq -\frac{2\mu_y}{g},$$

$$MSE(t_p) - MSE(\hat{t}_s) \geq 0 \quad (26)$$

$$\text{if } \frac{-2A_o - \mu_x B_o}{gB_o} \leq \Delta \leq \frac{\mu_y}{g},$$

$$MSE(t'_R) - MSE(\hat{t}_s) \geq 0 \quad (27)$$

$$\text{if } \Delta \leq \frac{gB_o\mu_x - 2A_o\mu_x}{gB_o},$$

$$MSE(t'_P) - MSE(\hat{t}_s) \geq 0 \quad (28)$$

$$\text{if } \Delta \geq \frac{2A_o\mu_x - gB_o\mu_x}{gB_o},$$

$$MSE(\hat{t}_R) - MSE(\hat{t}_s) \geq 0 \quad (29)$$

$$\text{if } \Delta \geq \frac{gB_o - 2A_o\mu_x}{gB_o},$$

$$MSE(\hat{t}_P) - MSE(\hat{t}_s) \geq 0 \quad (30)$$

$$\text{if } \Delta \geq \frac{2A_o\mu_x - gB_o}{gB_o}.$$

If the conditions (24-30) holds true, the proposed estimator \hat{t}_s is more efficient than other mentioned estimators when there is non-response and measurement error on the study as well as auxiliary variables.

5 Empirical Study

In this section, we demonstrate the performance of different estimators over the usual unbiased estimators by generating four populations from normal distribution with different choices of parameters using R language program. The auxiliary information on variable 'X' has been generated from N (5, 10) population. This type of population is very relevant in mostly all socio-economic situations with one study and one auxiliary variable.

Population I $X = N(5, 10); Y = X + N(0, 1); y = Y + N(1, 3); X = X + N(1, 3); N = 5000; \mu_Y = 4.927167; \mu_X = 4.924306; \sigma_y^2 = 102.0075; \sigma_x^2 = 101.4117; \sigma_U^2 = 8.862114; \sigma_V^2 = 9.001304; \rho_{yx} = 0.995059$

N_1	N_2	$\sigma_{y(2)}^2$	$\sigma_{x(2)}^2$	$\sigma_{U(2)}^2$	$\sigma_{V(2)}^2$	$\rho_{yx(2)}$
4500	500	99.99174	99.87471	9.150544	8.756592	0.994916
4250	750	100.9428	100.8224	9.053862	8.766538	0.994916
4000	1000	104.2711	103.2349	8.821278	8.339179	0.995472

Population II $X = N(5, 10); Y = X + N(0, 1); y = Y + N(1, 5); X = X + N(1, 5); N = 5000; \mu_Y = 4.996681; \mu_X = 5.013507; \sigma_y^2 = 97.12064; \sigma_x^2 = 95.95803; \sigma_U^2 = 23.96055; \sigma_V^2 = 24.19283; \rho_{yx} = 0.994822$

N_1	N_2	$\sigma_{y(2)}^2$	$\sigma_{x(2)}^2$	$\sigma_{U(2)}^2$	$\sigma_{V(2)}^2$	$\rho_{yx(2)}$
4500	500	97.02783	94.54578	22.80557	25.43263	0.994546
4250	750	98.27616	97.42674	23.27837	24.13829	0.994992
4000	1000	96.09359	94.71923	24.42978	23.03076	0.99467

Population III $X = N(5, 10); Y = X + N(0, 1); y = Y + N(2, 3); X = X + N(2, 3); N = 5000; \mu_Y = 4.730993; \mu_X = 4.741928; \sigma_y^2 = 101.2633; \sigma_x^2 = 100.2288; \sigma_U^2 = 9.1025; \sigma_V^2 = 9.052019; \rho_{yx} = 0.995187$

N_1	N_2	$\sigma_{y(2)}^2$	$\sigma_{x(2)}^2$	$\sigma_{U(2)}^2$	$\sigma_{V(2)}^2$	$\rho_{yx(2)}$
4500	500	102.7504	101.2097	9.095136	8.8123	0.995045
4250	750	99.55993	99.49764	9.233619	8.805872	0.995314
4000	1000	105.4334	103.8947	9.277715	9.072151	0.995105

Population IV $X = N(5, 10); Y = X + N(0, 1); y = Y + N(2, 5); X = X + N(2, 5); N = 5000; \mu_Y = 4.961081; \mu_X = 4.96178; \sigma_y^2 = 102.2408; \sigma_x^2 = 100.8680; \sigma_U^2 = 25.94111; \sigma_V^2 = 25.03951; \rho_{yx} = 0.394221$

N_1	N_2	$\sigma_{y(2)}^2$	$\sigma_{x(2)}^2$	$\sigma_{U(2)}^2$	$\sigma_{V(2)}^2$	$\rho_{yx(2)}$
4500	500	103.5361	102.1031	25.31099	22.84483	0.394622
4250	750	103.6790	102.7446	24.6859	26.12337	0.395036
4000	1000	100.1031	99.31665	25.80394	24.50468	0.394778

Table 1: Percent relative efficiencies of the estimators with respect to the usual unbiased estimator for Population I

N_1	N_2	k	$PRE(\hat{t}_s)$	$PRE(\hat{\mu}_v^*)$	$PRE(t_R)$	$PRE(t_P)$	$PRE(t'_R)$	$PRE(t'_P)$	$PRE(\hat{t}_R)$	$PRE(\hat{t}_P)$
4500	500	2	611.3859	100	586.1473	26.155	123.5517	82.28832	104.2404	95.99756
4250	750	2	612.4949	100	587.2272	26.15129	123.5588	82.285	104.2414	95.99667
4000	1000	2	618.4895	100	593.7605	26.16283	123.5637	82.28579	104.2418	95.99656
4500	500	3	610.4908	100	585.2622	26.15738	123.5467	82.29064	104.2397	95.9982
4250	750	3	612.5058	100	587.2254	26.15069	123.5594	82.28464	104.2416	95.99658
4000	1000	3	622.7055	100	598.3392	26.117023	123.5677	82.286	104.2421	95.9964
4500	500	4	609.7453	100	584.525	26.15938	123.5423	82.29258	104.239	95.99872
4250	750	4	612.5143	100	587.224	26.15023	123.5599	82.28436	104.2416	95.9965
4000	1000	4	625.8268	100	601.7272	26.17563	123.5706	82.28615	104.2423	95.99628
4500	500	5	609.1146	100	583.9015	26.16107	123.5389	82.29423	104.2385	95.99917
4250	750	5	612.5211	100	587.2229	26.14985	123.5603	82.28413	104.2417	95.99645
4000	1000	5	628.2307	100	593.3356	26.17976	123.5729	82.28626	104.2425	95.99619

Table 2: Percent relative efficiencies of the estimators with respect to the usual unbiased estimator for Population II

N_1	N_2	k	$PRE(\hat{t}_s)$	$PRE(\hat{\mu}_y^*)$	$PRE(t_R)$	$PRE(t_P)$	$PRE(t'_R)$	$PRE(t'_P)$	$PRE(\hat{t}_R)$	$PRE(\hat{t}_P)$
4500	500	2	272.9998	100	246.7997	28.02771	119.5495	84.18319	103.5922	96.55749
4250	750	2	274.0183	100	247.8102	28.01111	119.579	84.16816	103.5967	96.55358
4000	1000	2	273.284	100	247.5285	28.0755	119.5172	84.20558	103.5866	96.56272
4500	500	3	272.8247	100	246.5344	28.01987	119.5527	84.18034	103.5929	96.55688
4250	750	3	274.6241	100	248.3268	27.99153	119.6041	84.1543	103.6007	96.55008
4000	1000	3	273.3338	100	247.8097	28.10205	119.4975	84.21871	103.5832	96.56582
4500	500	4	272.679	100	246.3135	28.01332	119.5554	84.177978	103.5934	96.55636
4250	750	4	275.0968	100	248.7298	27.97633	119.6236	84.14354	103.6037	96.54736
4000	1000	4	273.371	100	248.0165	28.12157	119.4831	84.22835	103.5808	96.5681
4500	500	5	272.5558	100	246.1266	28.00778	119.5577	84.17596	103.5939	96.55593
4250	750	5	275.4759	100	249.0539	27.96419	119.6391	84.13493	103.6062	96.54519
4000	1000	5	273.3998	100	248.1751	28.13653	119.472	84.23573	103.5789	96.56984

Table 3: Percent relative efficiencies of the estimators with respect to the usual unbiased estimator for Population III

N_1	N_2	k	$PRE(\hat{t}_s)$	$PRE(\hat{\mu}_y^*)$	$PRE(t_R)$	$PRE(t_P)$	$PRE(t'_R)$	$PRE(t'_P)$	$PRE(\hat{t}_R)$	$PRE(\hat{t}_P)$
4500	500	2	601.5921	100	579.6432	26.33387	123.3383	82.40343	104.3867	95.86719
4250	750	2	600.0930	100	577.75412	26.31837	123.3498	82.39595	104.3887	95.86531
4000	1000	2	602.8336	100	580.9905	26.33597	123.3398	82.40336	104.3868	95.86710
4500	500	3	602.5760	100	580.7790	26.33891	123.3360	82.40535	104.3862	95.86764
4250	750	3	599.8685	100	577.3757	26.31135	123.3563	82.39209	104.3898	95.86431
4000	1000	3	604.5177	100	582.8746	26.34164	123.3389	82.40492	104.3865	95.86742
4500	500	4	603.3976	100	581.7269	26.34311	123.3341	82.40694	104.3858	95.86801
4250	750	4	599.6940	100	577.0808	26.30589	123.3614	82.38907	104.3907	95.86353
4000	1000	4	605.7497	100	584.2524	26.34576	123.3382	82.40605	104.3863	95.86765
4500	500	5	604.0940	100	582.5301	26.34665	123.3325	82.40829	104.3855	95.86832
4250	750	5	599.5544	100	576.8447	26.30151	123.3655	82.38666	104.3914	95.86291
4000	1000	5	606.6901	100	585.3037	26.34889	123.3377	82.40692	104.3861	95.86782

Table 4: Percent relative efficiencies of the estimators with respect to the usual unbiased estimator for Population IV

N_1	N_2	k	$PRE(\hat{t}_s)$	$PRE(\hat{\mu}_y^*)$	$PRE(t_R)$	$PRE(t_P)$	$PRE(t'_R)$	$PRE(t'_P)$	$PRE(\hat{t}_R)$	$PRE(\hat{t}_P)$
4500	500	2	354.8893	100	328.7961	27.14476	121.2849	83.33059	103.9072	96.28215
4250	750	2	301.8073	100	274.5276	27.52986	120.3716	83.75985	103.7712	96.39880
4000	1000	2	271.2105	100	244.0121	27.93889	119.5888	84.14978	103.6517	96.50254
4500	500	3	270.4614	100	243.5957	27.98911	119.5372	84.18031	103.6432	96.51016
4250	750	3	226.8626	100	198.3265	28.61770	118.1727	84.85689	103.4347	96.69096
4000	1000	3	204.2863	100	175.7478	29.26783	117.0891	85.43617	103.2636	96.84161
4500	500	4	228.7139	100	201.1703	29.71607	118.1516	84.88476	103.4294	96.69652
4250	750	4	193.4410	100	163.9436	29.49841	116.5639	85.70503	103.1819	96.91288
4000	1000	4	176.0908	100	146.5301	30.29490	115.3703	86.37732	103.9887	97.08473
4500	500	5	203.9089	100	175.7708	29.34852	117.0261	85.47826	103.2526	96.85169
4250	750	5	174.6404	100	144.3668	30.22601	115.3357	86.38035	103.9850	96.08718
4000	1000	5	160.6783	100	130.3080	31.11247	114.1159	87.09577	103.7838	96.26760

Tables 1 to 4 demonstrate that the proposed estimator \hat{t}_s performs better compared to the other estimators considered in the study in terms of gaining efficiency. according to the Tables 1 and 2, it is envisaged that the percent relative efficiency (PRE) of the proposed estimator \hat{t}_s decreases with the increase in the value k, when $N_1 = 4500$; $N_2 = 500$. However, PRE of \hat{t}_s increases with the increase in the value of k, when $N_1 = 4250$; $N_2 = 750$ and $N_1 = 4000$; $N_2 = 1000$ for population I and II. Table 3 shows that the PRE of the proposed estimator \hat{t}_s decreases with the increase in the value of k when $N_1 = 4250$; $N_2 = 750$, while it increases when $N_1 = 4500$; $N_2 = 500$ and $N_1 = 4000$; $N_2 = 1000$. Furthermore, Table 4 indicates that the PRE of \hat{t}_s decreases with the increase in the value of k in all cases.

6 Conclusion

In the present paper, we have proposed an estimator for estimating the population mean of the study variable. The suggested estimator used auxiliary information to improve efficiencies in the situation when there are non-response and measurement errors on study variable and auxiliary variable. The relative performance of the proposed estimators was compared with the conventional estimators. The proposed estimator performs better than the usual unbiased estimator, ratio estimator and product estimators, transformed ratio and product estimators and dual to ratio and product estimators in the presence of measurement and non-response error. The study was supported by empirical study based on four populations. We recommend a proposed estimator for future study to investigate the characteristics of the variable of interest when there is measurement and non-response occurs in the survey.

Acknowledgement

The authors are grateful to the reviewers for their valuable suggestions which improved the quality of the paper.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- [1] W.G. Cochran, The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, T. J. Agric. Sc., **30**, 262-275 (1940).
- [2] W.G. Cochran, Sampling theory when the sampling units are of unequal sizes, J. Amer. Statist. Assoc., **37**, 199-212 (1942).
- [3] D.S. Robson, Applications of multivariate polykeys to the theory of unbiased ratio-type estimation, J. Amer. Statist. Assoc., **52(280)**, 511-522, (1957).
- [4] M.N. Murthy, Product method of estimation, Sankhya, **26**, A, 294-307, (1964).
- [5] M. Ilves, Estimation in the presence of non-response and measurement errors, Proceedings of the third Baltic – Nortic conference in Survey Statistics, June 13-17, (2011), Norrfallsviken, Sweden.
- [6] R.M. Groves, Survey errors and survey costs, New York, NY: Wiley – Interscience (1989).
- [7] W.G. Cochran, *Sampling Techniques*. John Wiley, New York (1963).
- [8] W.G. Cochran, Errors of measurement in statistics, Technometrics, **10(4)**, 637-666 (1968).
- [9] W.G. Cochran, *Sampling Techniques*, 3rd edition, John Wiley and Sons. New York (1977).
- [10] W.A. Fuller, Estimation in the presence of measurement error, International Statistical Review, **63(2)**, 121-147 (1995).
- [11] Shalabh, Ratio method of estimation in the presence of measurement errors, Journal of the Indian Society of Agricultural Statistics, **50(2)**, 150-155, (1997).
- [12] Manisha and R.K. Singh, An estimation of population mean in the presence of measurement errors, Journal of the Indian Society of Agricultural Statistics, **54(1)**, 13-18, (2001).
- [13] Manisha and Singh, R.K. (2002), *Role of regression estimator involving measurement errors*. Brazilian Journal of Probability and Statistics, **16**, 39-46.
- [14] L. Wang, A simple adjustment for measurement errors in some dependent variable models, Statistics & Probability Letters, **58**, 427-433, (2002).
- [15] J. Allen, H.P. Singh, and F. Smarandache, A family of estimators of population mean using multi-auxiliary information in presence of measurement errors, Int. J. Soc. Econ., **30(7)**, 837-849 (2003).
- [16] H.P. Singh and N. Karpe, Effect of measurement errors on a class of estimators of population mean using auxiliary information in sample surveys, Journal of Statistical Research, Iran, **4**, 175-189, (2007).
- [17] H.P. Singh and N. Karpe, Estimation of population variance using auxiliary information in the presence of measurement errors, Statistics in Transition-New Series, **9(3)**, 443-470, (2008).
- [18] H.P. Singh and N. Karpe, On the estimation of ratio and product of two population means using supplementary information in presence of measurement errors, Statistica, Anno **LXIX**, N.1, 27-47, (2009).

- [19] H.P. Singh and N. Karpe, Estimation of mean, ratio and product using auxiliary information in the presence of measurement errors in sample surveys, *Journal of Statistical Theory and Practice*, **4(1)**, 111-136, (2010).
- [20] C. Salas, and T. G. Gregoire, Statistical analysis of ratio estimators and their estimators of variances when the auxiliary variate is measured with error, *European Journal of Forest Research*, **129**, 847-861, (2010).
- [21] M. Kumar, R. Singh, N. Sawan, and P. Chauhan, Exponential ratio method of estimation in the presence of measurement errors, *Int. J. Agri. Statist. Sci.*, **7(2)**, 457-461, (2011).
- [22] D. Shukla, S. Pathak, and N.S. Thakur, An estimator for mean estimation in presence of measurement error, *Research and Reviews: A Journal of Statistics*, **1(1)**, 1-8, (2012).
- [23] M.H. Hansen and W.N. Hurwitz, The problem of non-response in sample surveys, *J. Amer. Statist. Assoc.*, **41**, 517- 529 (1946).
- [24] M.A. El-Badry, M. A., A sampling procedure for mailed questionnaires, *J. Amer. Statist. Assoc.*, **51**, 209-227 (1956).
- [25] G.T. Foradari, Some non-response sampling theory for two stage designs, Mimeo. 297, North Carolina State College, Raleigh (1961).
- [26] K.P. Srinath, Multiphase sampling in non-response problems, *J. Amer. Statist. Assoc.*, **66**, 583-586, (1971).
- [27] P.S.R.S. Rao, Ratio estimation with sub sampling the non-respondents, *Survey Methodology*, **12(2)**, 217-230, (1986).
- [28] P.S.R.S. Rao, Ratio and Regression estimates with sub sampling of non-respondents, In Paper presented at a Special Contributed Session of the International Statistical Association Meetings, Sept., 2-16, (1987), Tokyo, Japan.
- [29] P.S.R.S. Rao, Regression estimators with sub sampling of non-respondents. In-Data Quality Control, Theory and Pragmatics, (Gunar E. Liepins and V.R.R. Uppuluri, eds) Marcel Dekker, New York, 191-2008, (1990).
- [30] B. B. Khare and S. Srivastava, Estimation of population mean using auxiliary character in presence of non-response, *Nat. Acad. Sci. Lett., India*, **16**, 111-114, (1993).
- [31] B. B. Khare and S. Srivastava, Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response, *Proc. Indian Nat. Sci. Acad.*, **65**, 195-203, (1995).
- [32] B. B. Khare and S. Srivastava, Transformed ratio type estimators for the population mean in the presence of non response, *Comm. Statist. Theory Methods*, **26**, 1779-1791, (1997).
- [33] B. B. Khare and S. Srivastava, Generalized two phase estimators for the population mean in the presence of non-response, *Aligarh J. Statist.*, **30**, 39-54, (2010).
- [34] A.A. Sodipo and K.O. Obisesan, Estimation of the population mean using difference cum ratio estimator with full response on the auxiliary character, *Res. J. App. Sci.*, **2**, 769-772, (2007).
- [35] H.P. Singh and S. Kumar, Estimation of mean in presence of non-response using two-phase sampling scheme, *Statistical Papers*, **50(3)**, 559-582, (2010).
- [36] H.P. Singh and S. Kumar, A regression approach to the estimation of the finite population mean in the presence of non-response, *Aust. N.Z.J. Statist.*, **50(4)**, 395-408, (2008a).
- [37] H.P. Singh and S. Kumar, A general family of estimators of finite population ratio, product and mean using two-phase sampling scheme in the presence of non-response, *J. Statist. Theo. Practice* **2**, 677-692, (2008b).
- [38] H.P. Singh and S. Kumar, A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information, *SORT*, **33**, 71-84, (2009).
- [39] H.P. Singh and S. Kumar, Combination of regression and ratio estimate in presence of non-response, *Braz. J. Statist. Assoc.*, **25(2)**, 205-217, (2011).
- [40] M. Ismail, M. Hanif, and M.S. Shahbaz, A new cost-effective estimator in the presence of non-response for two-phase sampling, *Journal Technology*, **63(2)**, 1-4 (2013).
- [41] B. B. Khare, S. Srivastava, and K. Kumar, An improved class of estimators for population mean using auxiliary character in the presence of non-response. *Natl. Acad. Sci. Lett.*, **35(5)**, 361-366, (2012a). DOI 10.1007/ s40009-012-0059-1.
- [42] B. B. Khare, S. Srivastava, and K. Kumar, Chain type estimators of two population means using auxiliary characters in the presence of non- response, *J. Sci. Res. Banaras Hindu Univ. Varanasi*, **56**, 183-196, (2012b).
- [43] S. Kumar and S. Bhogal, Estimation of the population mean in presence of non-response, *Commun. Korean Statist. Soc.*, **18(1)**, 1-12, (2011).
- [44] B. B. Khare and K. Kumar, Transformed two phase sampling ratio and product type estimators for population mean in the presence of non-response, *Aligarh J. Statist.*, **29**, 91-106, (2009).
- [45] H.P. Singh, S. Kumar and M. Kozak, Improved estimation of finite population mean using sub-sampling to deal with the non-response in two-phase sampling scheme, *Commun. Statist. Theo. Methods*, **39(5)**, 791-802, (2010).
- [46] J. Shabbir, and N.S. Khan, Some modified exponential ratio-type estimators in the presence of non-response under two-phase sampling scheme, *Electron. J. App. Statist. Anal.*, **6(1)**, 1-17, (2013).
- [47] S. Kumar, Improved estimation of population mean in presence of nonresponse and measurement error, *Journal of Statistical Theory and Practice*, **10(4)**, 707-720, (2016).
- [48] S. Kumar, and K. Chatterjee, A simulation study for unit non-response under double sampling. *Journal of Statistics Applications and Probability Letters*, **2(1)**, 37-50 (2015).
- [49] V. Sharma and S. Kumar, Estimation of population mean using auxiliary variable and non-response, *Revista Investigacion Operacional*, **41(3)**, 438-444, (2020).
- [50] S. Jackman, Correcting surveys for non-response and measurement error using auxiliary information, *Electoral Studies*, **18(1)**, 7-27. (1999).

- [51] J. Dixon, Assessing non-response bias and measurement error using statistical matching, Section on Survey Research Methods, (2010).
- [52] M. Azeem, On estimation of population mean in the presence of measurement error and non-response, Unpublished Ph.D. thesis. National college of Business Administration and Economics, Lahore(2014).
- [53] S. Kumar, S. Bhogal, N. S. Nataraja and M. Viswanathaiah, Estimation of population mean in the presence of non-response and measurement error, Colombian Journal of Statistics, **38(1)**, 145-161, (2015).
- [54] S. Kumar and M. Choudhary, Estimation of Population Product in the Presence of Non-Response and Measurement Error in Successive Sampling, **10(3)**, 71-83, (2021).
- [55] A. Sahai, An efficient variant of the product and ratio estimator, Statistica Neerlandica, **33**, 27-35, (1979).
- [56] T. Srivenkataramana, A dual to ratio estimator in sample surveys, Biometrika, **67(1)**, 199-204, (1980).
-