

EFFECT OF MEASUREMENT ERROR AND NON-RESPONSE ON ESTIMATION OF POPULATION MEAN

Poonam Singh*¹, Rajesh Singh* and Carlos N. Bouza**

*Department of statistics, Banaras Hindu University (U.P), India

**Facultad MATCOM, Universidad de La Habana, Cuba

ABSTRACT

In this paper effect of measurement error and non response error is examined on estimation of unknown population mean of study variable. We have obtained the expression of the MSE (mean square error) of the proposed estimator up to first order of approximation. We have shown theoretically and empirically that the proposed estimator performs better than other estimators considered in this article. For empirically study we have used for different data sets.

KEYWORDS: Auxiliary Variable, Measurement Error, Non-Response, Bias, Mean Square Error (MSE).

MSC: 62D05

RESUMEN

En este paper es examinado el efecto de error de medición de respuesta en la estimación de la media desconocida de la variable de estudio. Hemos obtenido la expresión del MSE (error cuadrático medio) del estimador propuesto hasta el primer orden de aproximación. Hemos mostrado teóricamente y empíricamente que el propuesto estimador se comporta mejor que los otros considerados en este artículo. Para los estudios empíricos usamos diferentes data.

PALABRAS CLAVE: Variable Auxiliar, Error de Medida, Sesgo de No-Respuesta, Error Cuadrático Medio

MSC: 62D05

1. INTRODUCTION

In a perfect world survey has no non-response, all selected element will participate and provide all of the requested information. However, today reality is very different. Missing data due to non-response is a normal although undesirable feature of any survey. In theory of sample surveys, auxiliary variables play important role. Auxiliary information is used to increase precision of an estimator. Error free measurement of the auxiliary variables on the population frame would thus seems critical for making appropriate finite population inferences. Unfortunately, there has been very little research examining the impact of measurement error in the auxiliary variable on estimation of parameters. Measurement errors occur when answer provided by respondents departs from the true value on the measurement (e.g. failure to report correctly whether he visited doctor in last six months). Measurement errors include observational error, instrument error, respondent error etc. Fuller (1987), Corrol, Ruppert and Stefanski (1995), Meijer (2000), Bound, Brown and Mathiowetz (2001), Hausman (2001), Srivastava and Shalabh(2001), Manisha and Singh (2002), Singh and Karpe (2008,2009), Kumar et al.(2011), Shukla et al.(2012) and Singh and Sharma (2015) are the few references who have studied problem of measurement error.

Besides measurement errors, non-response has always been a matter of concern in sample surveys. Non-response is the failure to get information from some units of the population due to various reasons like unavailability of respondents, lack of information and refusals etc.

Description of non-response error and its effect is described in Cochran (1977). Kalton and Karsprzyk (1986), Merg (1995), Rubin (1996), Kenward and Carpenter (2007) etc. gave several approaches to address non-response in sample surveys. Non-response problem is studied to:-

- Avoid non-response before it has occurred.
- Develop techniques required in estimation when non-response has occurred.

¹ Corresponding Author, ppiyu123@gmail.com

Hansen and Hurwitz (1946) were the first who pioneered technique for estimation of mean when non-response is present in surveys. He simply drawn a simple random sample of size n and mailed questionnaire to sampled units. Then re-contacted some of the non-responding units by drawing a sub-sample from the non-responding units in the initial first attempt.

Cochran (1977) applied Hansen and Hurwitz technique to formulate a ratio estimator of the population mean. Similarly, Rao (1986), Ofkar and Lee (2000), Tabasum and Khan (2004,2006), Sodipo and Obisesan (2007), Singh and Kumar (2008), Singh et al. (2014), Chaudhary et al. (2014), Singh and Singh (2015) and Sharma and Singh (2015) considered the problem of estimating mean in presence of non-response/ measurement error.

Problem of measurement error and non-response error are studied by many researchers separately. But these problems may creep into survey sampling at the same time. If these errors are small and negligible they can be ignored but if these errors are not negligible, inferences may lead to undesirable consequences. In this paper we will study how both the errors affect efficiency of estimators.

2. NOTATIONS

Let us consider a finite population $(U = U_1, U_2, \dots, U_N)$ of size N such that Y be study variable and X any be auxiliary variable. We draw a sample of size n from a population by using simple random sampling without replacement scheme.

Suppose that N_1 units respond for the survey questions and N_2 units do not respond. Then by following

Hansen Hurwitz (1946) sampling plan, a sub-sample of size $k = \frac{r_2}{h} (h > 1)$ from N_2 non-respondents

is selected at random and is re-contacted for their direct interview.

Here it is assumed that r units respond to the survey.

Let (x_i^*, y_i^*) be the observed values and (X_i^*, Y_i^*) be the true values of the study variable Y and auxiliary variable X. where $(i=1, 2, \dots, n)$ unit in the sample.

Then measurement error is given by-

$$u_i^* = y_i^* - Y_i^* \quad \text{And} \quad v_i^* = x_i^* - X_i^* \quad (2.1)$$

Where (u_i^*, v_i^*) are random in nature and both are uncorrelated with mean zero and variance S_U^2 and S_V^2 are associated with measurement error in study variable Y and auxiliary variable X respectively for the responding part of the population. $S_{U(2)}^2$ and $S_{V(2)}^2$ are the variances associated with measurement error in study variable Y and auxiliary variable X respectively for the non-responding part of the population.

We further assume that mean of study variable Y is unknown and auxiliary variable X is known.

Following symbols have their meaning given below:

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i = \text{population mean of Y}$$

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i = \text{population mean of X}$$

\bar{y} and \bar{x} are the sample means of y and x respectively. S_Y^2 And S_X^2 are the population variances of Y and X respectively for the responding part of the population. $S_{Y(2)}^2$ And $S_{X(2)}^2$ are the variances of Y and X respectively for non-responding part of the population.

ρ And $\rho_{(2)}$ are the population correlation coefficient between X and Y for the responding and non-

responding part of the population respectively. C_Y And $C_{Y(2)}$ are the coefficients of variation of Y for the responding and non-responding part of the population respectively. Similarly C_X and $C_{X(2)}$ are the

coefficients of variation of X for the responding and non-responding part of the population respectively.

In order to obtain MSE of the estimators in presence of non-response and measurement error, following notations are used:

Let

$$w_Y^* = \sum_{i=1}^n (Y_i^* - \bar{Y}) \quad (2.2)$$

and

$$w_X^* = \sum_{i=1}^n (X_i^* - \bar{X}) \quad (2.3)$$

Then

$$w_U^* = \sum_{i=1}^n U_i^* \quad (2.4)$$

and

$$w_V^* = \sum_{i=1}^n V_i^* \quad (2.5)$$

Adding (2.2) and (2.4), dividing both sides by n, we have

$$\frac{1}{n}(w_Y^* + w_U^*) = \frac{1}{n} \sum_{i=1}^n (Y_i^* - \bar{Y}) + \frac{1}{n} \sum_{i=1}^n (y_i^* - Y_i^*) \quad (2.6)$$

Or

$$\frac{1}{n}(w_Y^* + w_U^*) = \frac{1}{n} \sum_{i=1}^n y_i^* - \bar{Y} = \xi_0 \text{ (say)} \quad (2.7)$$

Similarly adding (2.3) and (2.5), dividing both sides by n we get

$$\frac{1}{n}(w_X^* + w_V^*) = \frac{1}{n} \sum_{i=1}^n X_i^* - \bar{X} = \xi_1 \text{ (say)} \quad (2.8)$$

On simplification, we get

$$\bar{y}^* = \bar{Y} + \xi_0 \quad (2.9)$$

$$\bar{x}^* = \bar{X} + \xi_1 \quad (2.10)$$

Further,

$$E(\xi_0^2) = \lambda_2 (S_Y^2 + S_U^2) + \theta (S_{Y(2)}^2 + S_{U(2)}^2) = \nabla_0^2 \text{ (say)} \quad (2.11)$$

$$E(\xi_1^2) = \lambda_2 (S_X^2 + S_V^2) + \theta (S_{X(2)}^2 + S_{V(2)}^2) = \nabla_1^2 \text{ (say)} \quad (2.12)$$

$$E(\xi_0 \xi_1) = \lambda_2 \rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)} = \nabla_0 \nabla_1 \text{ (say)} \quad (2.13)$$

3. EXISTING ESTIMATORS

Hansen Hurwitz (1946) estimator for estimating population mean in presence of non-response and measurement error is given by-

$$\bar{y}^* = \left(\frac{n_1}{n} \right) \bar{y}_{n1} + \left(\frac{n_2}{n} \right) \bar{y}_r = t_a \text{ (say)} \quad (3.1)$$

$$\text{where } \bar{y}_{n1} = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i \text{ and } \bar{y}_r = \frac{1}{r} \sum_{i=1}^r y_i.$$

Expression (2.1) can be written as:

$$t_a = \bar{Y} + \xi_0 \quad (3.2)$$

Subtracting \bar{Y} from both the sides of equation (3.2) and taking expectation, we get bias of estimator t_a given as:

$$\text{Bias}(t_a) = 0 \quad (3.3)$$

Subtracting \bar{Y} from both the sides of equation (3.2) and squaring, we get

$$(t_a - \bar{Y})^2 = \xi_0^2 \quad (3.4)$$

Taking expectations, mean square error of the estimator t_a is obtained up to first order of approximation as:

$$\text{MSE}(t_a) = \lambda_2(S_Y^2 + S_U^2) + \theta(S_{Y(2)}^2 + S_{U(2)}^2) \quad (3.5)$$

In case when measurement error is zero, then

$$\text{MSE}(t_a) = \lambda_2 S_Y^2 + \theta S_{Y(2)}^2 \quad (3.6)$$

Contribution of measurement error to MSE of estimator t_a is:

$$\text{ME}(t_a) = \lambda_2 S_U^2 + \theta S_{U(2)}^2 \quad (3.7)$$

Cochran (1977) ratio estimator in presence of non-response and measurement error is given by-

$$t_r = \frac{\bar{y}^*}{\bar{x}^*} \bar{X} \quad (3.8)$$

Expressing the estimator t_r in terms of ξ 's, and then simplifying, we get:

$$t_r = \bar{Y} + \xi_0 - \frac{\bar{Y}}{\bar{X}} \xi_1 - \frac{\xi_0 \xi_1}{\bar{X}} + \frac{\bar{Y}^2}{\bar{X}^2} \xi_1^2 \quad (3.9)$$

Subtracting \bar{Y} from both the sides of equation (3.9), we get

$$(t_r - \bar{Y}) = \xi_0 - \frac{\bar{Y}}{\bar{X}} \xi_1 - \frac{\xi_0 \xi_1}{\bar{X}} + \frac{\bar{Y}^2}{\bar{X}^2} \xi_1^2 \quad (3.10)$$

Taking expectation of equation (3.10), we get bias of the estimator t_r estimator as

$$\text{Bias}(t_r) = \frac{\bar{Y}^2}{\bar{X}^2} \left\{ \lambda_2(S_X^2 + S_V^2) + \theta(S_{X(2)}^2 + S_{V(2)}^2) \right\} - \frac{1}{\bar{X}} \left\{ \lambda_2(\rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)}) \right\} \quad (3.11)$$

Squaring equation (3.10) and then taking expectations, MSE of the estimator t_r is given by

$$\begin{aligned} \text{MSE}(t_r) &= \lambda_2(S_Y^2 + S_U^2) + \theta(S_{Y(2)}^2 + S_{U(2)}^2) + \frac{\bar{Y}^2}{\bar{X}^2} \left\{ \lambda_2(S_X^2 + S_V^2) + \theta(S_{X(2)}^2 + S_{V(2)}^2) \right\} \\ &\quad - \frac{2\bar{Y}}{\bar{X}} \left\{ \lambda_2(\rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)}) \right\} \end{aligned} \quad (3.12)$$

In case when measurement error is zero or negligible,

$$\begin{aligned} \text{MSE}(t_r) &= \lambda_2 S_Y^2 + \theta S_{Y(2)}^2 + \frac{\bar{Y}^2}{\bar{X}^2} \left\{ \lambda_2 S_X^2 + \theta S_{X(2)}^2 \right\} \\ &\quad - \frac{2\bar{Y}}{\bar{X}} \left\{ \lambda_2(\rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)}) \right\} \end{aligned} \quad (3.13)$$

The contribution of measurement error to the MSE of the estimator t_r is:

$$\text{MSE}(t_r) = \lambda_2 \bar{Y}^2 \left(\frac{S_U^2}{\bar{Y}^2} + \frac{S_V^2}{\bar{X}^2} \right) + \theta \bar{Y}^2 \left(\frac{S_{U(2)}^2}{\bar{Y}^2} + \frac{S_{V(2)}^2}{\bar{X}^2} \right) \quad (3.14)$$

Rao's (1991) estimator under non-response and measurement error is given by-

$$t_{ra} = \left[w_1 (\bar{X} - \bar{x}^*) + w_2 \bar{y}^* \right] \quad (3.15)$$

Expressing t_{ra} in terms of ξ 's, subtracting \bar{Y} and then squaring we get:

$$\text{MSE}(t_{ra}) = E \left[-w_1 \xi_1 + w_2 \xi_0 + w_2 \bar{Y} - \bar{Y} \right]^2 \quad (3.16)$$

Or

$$\text{MSE}(t_{ra}) = E \left[w_1^2 \xi_1^2 + w_2^2 \xi_0^2 + w_2^2 \bar{Y}^2 - 2w_1 w_2 \xi_0 \xi_1 - 2w_2 \bar{Y}^2 + \bar{Y}^2 \right] \quad (3.17)$$

Equation (3.17) can be written as

$$\text{MSE}(t_{ra}) = \left[w_1^2 A_r + w_2^2 B_r - 2w_1 w_2 D_r - 2w_2 C_r + \bar{Y}^2 \right] \quad (3.18)$$

Where

$$A_r = \nabla_1^2 \quad (3.19)$$

$$B_r = \bar{Y}^2 + \nabla_0^2 \quad (3.20)$$

$$C_r = \bar{Y}^2 \quad (3.21)$$

$$D_r = \nabla_0 \nabla_1 \quad (3.22)$$

Differentiating equation (3.19) with respect to w_1 and w_2 we get

$$w_1 = \frac{CD}{AB - D^2} \quad \text{and} \quad w_2 = \frac{AC}{AB - D^2} \quad (3.23)$$

Therefore substituting values of equation (3.23) in equation (3.18) MSE becomes

$$MSE(t_{ra}) = \bar{Y}^2 + \frac{AC}{AB - D^2} \quad (3.24)$$

Bahl and Tuteja (1991) estimator in presence of non-response and measurement error is given by-

$$t_{bt} = \bar{y}^* \exp \left[\frac{\bar{X} - \bar{x}^*}{\bar{X} + \bar{x}^*} \right] \quad (3.25)$$

Expressing equation (3.25) in terms of ξ 's, we get

$$\bar{y}_{bt} = \bar{Y} + \xi_0 - \frac{\bar{Y}}{2\bar{X}} \xi_1 - \frac{\xi_0 \xi_1}{2\bar{X}} + \frac{3}{8} \frac{\bar{Y}}{\bar{X}^2} \xi_1^2 \quad (3.26)$$

Subtracting from both sides of equation and squaring it, MSE of estimator t_{ra} is obtained as:

$$MSE(t_{bt}) = E \left[\xi_0^2 + \frac{\bar{Y}^2}{4\bar{X}^2} \xi_1^2 - \frac{\bar{Y}}{\bar{X}} \xi_0 \xi_1 \right] \quad (3.27)$$

Or

$$MSE(t_{bt}) = \left[\begin{aligned} & \left\{ \lambda_2 (S_Y^2 + S_U^2) + \theta (S_{Y(2)}^2 + S_{U(2)}^2) \right\} + \frac{\bar{Y}^2}{4\bar{X}^2} \left\{ \lambda_2 (S_X^2 + S_V^2) + \theta (S_{X(2)}^2 + S_{V(2)}^2) \right\} \\ & - \frac{\bar{Y}}{\bar{X}} (\lambda_2 \rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)}) \end{aligned} \right] \quad (3.28)$$

Regression Estimator under measurement error and non-response is given by-

$$t_{lr} = \bar{y}^* + b(\bar{X} - \bar{x}^*) \quad (3.29)$$

MSE of estimator t_{lr} is obtained as:

$$MSE(t_{lr}) = \lambda_2 (S_Y^2 + S_U^2) + \theta (S_{Y(2)}^2 + S_{U(2)}^2) + b^2 \left\{ \lambda_2 (S_X^2 + S_V^2) + \theta (S_{X(2)}^2 + S_{V(2)}^2) \right\} - 2b \left\{ \lambda_2 \rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)} \right\} \quad (3.30)$$

Differentiating MSE of t_{lr} with respect to b and equating it to zero, we get

$$b = \frac{\lambda_2 \rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)}}{\lambda_2 (S_X^2 + S_V^2) + \theta (S_{X(2)}^2 + S_{V(2)}^2)} = b_0 \text{ (say)} \quad (3.31)$$

The minimum MSE of the estimator t_{lr} is given by

$$\min MSE(t_{lr}) = \lambda_2 (S_Y^2 + S_U^2) + \theta (S_{Y(2)}^2 + S_{U(2)}^2) + b_0^2 \left\{ \lambda_2 (S_X^2 + S_V^2) + \theta (S_{X(2)}^2 + S_{V(2)}^2) \right\} - 2b_0 \left\{ \lambda_2 \rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)} \right\} \quad (3.32)$$

The contribution of measurement error in MSE of Regression estimator is:

$$ME(t_{lr}) = \lambda_2 S_U^2 + \theta S_{V(2)}^2 + b_0^2 \left\{ \lambda_2 S_V^2 + \theta S_{V(2)}^2 \right\} \quad (3.33)$$

4. PROPOSED ESTIMATOR

We propose estimator t_{sp} in presence of non-response and measurement error as

$$t_{sp} = \left[\frac{1}{2} \left\{ \bar{y}^* \exp\left(\frac{\bar{X} - \bar{x}^*}{\bar{X} + \bar{x}^*}\right) + \bar{y} \exp\left(\frac{\bar{x}^* - \bar{X}}{\bar{x}^* + \bar{X}}\right) \right\} + \alpha_1(\bar{X} - \bar{x}^*) + \alpha_2 \bar{y}^* \right] \exp\left[\frac{\bar{X}'' - \bar{x}^{**}}{\bar{X}'' + \bar{x}^{**}}\right] \quad (4.1)$$

where $\bar{X}'' = \bar{X}p$ and $\bar{x}^{**} = \bar{x}^* + \bar{X}(p-1)$

Further $p = \frac{\rho_{xy} + 1}{4}$ is a suitably chosen constant.

Expressing equation (3.1) in terms of ξ 's we have

$$t_{sp} = \left[(\bar{Y} + \xi_0 - \frac{\bar{Y}}{8\bar{X}^2} \xi_1^2) - \alpha_1 \xi_1 + \alpha_2 \bar{Y} + \alpha_2 \xi_0 \right] \exp\left[\frac{\xi_1}{2\bar{X}p} - \frac{\xi_1^2}{4\bar{X}^2 p^2}\right] \quad (4.2)$$

Simplifying equation (4.2) we get:

$$t_{sp} = \left[\left(\bar{Y} + \xi_0 + \frac{\bar{Y}}{8\bar{X}^2} \xi_1^2 - \frac{\bar{Y}}{2\bar{X}} \frac{\xi_1}{p} - \frac{\xi_0 \xi_1}{2\bar{X}p} + \frac{3}{8} \frac{\bar{Y}}{\bar{X}^2} \frac{\xi_1^2}{p^2} \right) - \alpha_1 \left(\xi_1 - \frac{\xi_1^2}{2\bar{X}p} \right) + \alpha_2 \left(\bar{Y} + \xi_0 - \frac{\bar{Y}}{2\bar{X}p} \xi_1 - \frac{\xi_0 \xi_1}{2\bar{X}p} + \frac{3}{8} \frac{\bar{Y}^2}{\bar{X}^2 p^2} \xi_1^2 \right) \right] \quad (4.3)$$

Or

$$t_{sp} - \bar{Y} = \left[\left(\xi_0 + \frac{\bar{Y}}{8\bar{X}^2} \xi_1^2 - \frac{\bar{Y}}{2\bar{X}} \frac{\xi_1}{p} - \frac{\xi_0 \xi_1}{2\bar{X}p} + \frac{3}{8} \frac{\bar{Y}}{\bar{X}^2} \frac{\xi_1^2}{p^2} \right) - \alpha_1 \left(\xi_1 - \frac{\xi_1^2}{2\bar{X}p} \right) + \alpha_2 \left(\bar{Y} + \xi_0 - \frac{\bar{Y}}{2\bar{X}p} \xi_1 - \frac{\xi_0 \xi_1}{2\bar{X}p} + \frac{3}{8} \frac{\bar{Y}^2}{\bar{X}^2 p^2} \xi_1^2 \right) \right] \quad (4.4)$$

Taking expectation on both the sides of equation (4.4), we get the bias of the estimator t_{sp} as:

$$Bias(t_{sp}) = \alpha_2 \bar{Y} + \left(\frac{\bar{Y}}{8\bar{X}^2} + \frac{\alpha_1}{2\bar{X}p} + \frac{3}{8} \frac{\bar{Y}}{\bar{X}^2 p^2} + \frac{3}{8} \frac{\alpha_2 \bar{Y}}{\bar{X}^2 p^2} \right) \nabla_1^2 - \left(\frac{1}{2\bar{X}p} + \frac{\alpha_2}{2\bar{X}p} \right) \nabla_0 \nabla_1 \quad (4.5)$$

Squaring equation (4.4) and taking expectations on both the sides, MSE of the estimator t_{sp} is given by

$$MSE(t_{sp}) = \alpha_1^2 \nabla_1^2 + \alpha_2^2 \left(\bar{Y}^2 + \nabla_0^2 + \frac{\bar{Y}^2}{\bar{X}^2 p^2} \nabla_1^2 - \frac{2\bar{Y}}{\bar{X}p} \nabla_0 \nabla_1 \right) - 2\alpha_1 \alpha_2 \left(\nabla_0 \nabla_1 - \frac{\bar{Y}}{\bar{X}p} \nabla_1^2 \right) - 2\alpha_1 \left(\nabla_0 \nabla_1 - \frac{\bar{Y}}{2\bar{X}p} \right) + 2\alpha_2 \left(\frac{\bar{Y}}{8\bar{X}^2} \nabla_1^2 - \frac{3}{2} \frac{\bar{Y}}{\bar{X}p} \nabla_0 \nabla_1 + \nabla_0^2 + \frac{5}{8} \frac{\bar{Y}^2}{\bar{X}^2} \frac{\nabla_1^2}{p^2} \right) + \left(\nabla_0^2 + \frac{\bar{Y}^2}{4\bar{X}^2} \frac{\nabla_1^2}{p^2} - \frac{\bar{Y}}{\bar{X}} \frac{\nabla_0 \nabla_1}{p} \right) \quad (4.6)$$

Where

$$\nabla_0^2 = E(\xi_0^2) = \lambda_2(S_Y^2 + S_U^2) + \theta(S_{Y(2)}^2 + S_{U(2)}^2) \quad (4.7)$$

$$\nabla_1^2 = E(\xi_1^2) = \lambda_2(S_X^2 + S_V^2) + \theta(S_{X(2)}^2 + S_{V(2)}^2) \quad (4.8)$$

$$\nabla_0 \nabla_1 = E(\xi_0 \xi_1) = \lambda_2 \rho_{YX} S_Y S_X + \theta \rho_{YX(2)} S_{Y(2)} S_{X(2)} \quad (4.9)$$

Equation (4.6) can be written as:

$$MSE(t_{sp}) = \alpha_1^2 A_s + \alpha_2^2 B_s - 2\alpha_1 \alpha_2 C_s - 2\alpha_1 D_s + 2\alpha_2 E_s + F_s \quad (4.10)$$

Where

$$A_s = \nabla_1^2 \quad (4.11)$$

$$B_s = \left(\bar{Y}^2 + \nabla_0^2 + \frac{\bar{Y}^2}{\bar{X}^2 p^2} \nabla_1^2 - 2 \frac{\bar{Y}}{\bar{X}p} \nabla_0 \nabla_1 \right) \quad (4.12)$$

$$C_s = \left(\nabla_0 \nabla_1 - \frac{\bar{Y}}{\bar{X}p} \nabla_1^2 \right) \quad (4.13)$$

$$D_s = \left(\nabla_0 \nabla_1 - \frac{\bar{Y}}{2\bar{X}p} \nabla_1^2 \right) \quad (4.14)$$

$$E_s = \left(\frac{\bar{Y}}{8\bar{X}} \nabla_1^2 - \frac{3}{2} \frac{\bar{Y}}{\bar{X}p} \nabla_0 \nabla_1 + \nabla_0^2 + \frac{5}{8} \frac{\bar{Y}^2}{\bar{X}^2} \frac{\nabla_1^2}{p^2} \right) \quad (4.15)$$

$$F_s = \left(\nabla_0^2 - \frac{\bar{Y}}{\bar{X}p} \nabla_0 \nabla_1 + \frac{\bar{Y}^2}{4\bar{X}^2} \frac{\nabla_1^2}{p^2} \right) \quad (4.16)$$

Differentiating equation (4.10) with respect to α_1 , α_2 and equating it to zero we get:

$$\alpha_1 A_s - \alpha_2 C_s = D_s \quad (4.17)$$

$$-\alpha_1 C_s + \alpha_2 B_s = -E_s \quad (4.18)$$

Solving equation (4.17) and (4.18) we get optimum values of α_1 and α_2 :

$$\alpha_1 = \frac{B_s D_s - C_s E_s}{A_s B_s - C_s^2} = \alpha_{10} \quad (4.19)$$

$$\alpha_2 = \frac{A_s E_s - C_s D_s}{A_s B_s - C_s^2} = \alpha_{20} \quad (4.20)$$

Substituting these optimum values of α_1 and α_2 in equation (4.10), min MSE of estimator t_{sp} is:

$$\min MSE(t_{sp}) = \bar{Y}^2 + \left[\frac{2C_s D_s E_s - B_s D_s^2 - A_s E_s^2}{A_s B_s - C_s^2} \right] \quad (4.21)$$

5. EMPIRICAL STUDY

We use following data sets for empirical study: Source: Muhammad Azeem & Muhammad Hanif (2016)

Population I:

$N=5000$, $\bar{Y} = 4.9271$, $\bar{X} = 4.9243$, $S_Y^2 = 102.007$, $S_X^2 = 101.411$, $S_U^2 = 8.8621$, $S_V^2 = 9.0013$, $\rho_{XY} = 0.9950$

N_1	N_2	$S_{Y(2)}^2$	$S_{X(2)}^2$	$S_{U(2)}^2$	$S_{V(2)}^2$	$\rho_{YX(2)}$
4500	500	99.99174	99.8747	9.1505	8.756	0.9949
4250	750	100.8224	100.822	9.05382	8.766	0.9955
4000	1000	103.2349	103.234	8.8212	8.339	0.9954

Population II:

$N=5000$, $\bar{Y} = 4.9966$, $\bar{X} = 5.0135$, $S_Y^2 = 97.1206$, $S_X^2 = 95.9580$, $S_U^2 = 23.96055$, $S_V^2 = 24.1928$, $\rho_{XY} = 0.9948$

N_1	N_2	$S_{Y(2)}^2$	$S_{X(2)}^2$	$S_{U(2)}^2$	$S_{V(2)}^2$	$\rho_{YX(2)}$
4500	500	97.0278	94.5457	22.8055	25.4326	0.9945
4250	750	98.2761	97.4267	23.2783	24.1382	0.9949
4000	1000	96.0935	94.7192	24.4297	23.0307	0.9946

Population III:

$$N=5000, \bar{Y} = 4.7309, \bar{X} = 4.7419, S_Y^2 = 101.2633, S_X^2 = 100.2288, S_U^2 = 9.1025, \\ S_V^2 = 9.0520, \rho_{XY} = 0.9951$$

N_1	N_2	$S_{Y(2)}^2$	$S_{X(2)}^2$	$S_{U(2)}^2$	$S_{V(2)}^2$	$\rho_{YX(2)}$
4500	500	102.75	101.2097	9.0951	8.8123	0.9950
4250	750	99.559	99.4976	9.2336	8.8058	0.9953
4000	1000	105.433	103.89	9.2777	9.0721	0.9951

Population IV:

$$N=5000, \bar{Y} = 1.9600, \bar{X} = 1.9433, S_Y^2 = 25.441, S_X^2 = 100.228, S_U^2 = 6.0404, \\ S_V^2 = 6.2244, \rho_{XY} = 0.9808$$

N_1	N_2	$S_{Y(2)}^2$	$S_{X(2)}^2$	$S_{U(2)}^2$	$S_{V(2)}^2$	$\rho_{YX(2)}$
4500	500	24.527	23.6120	6.3354	5.5894	0.97911
4250	750	28.596	27.553	6.1242	6.2996	0.9821
4000	1000	25.877	25.213	5.9383	6.2722	0.9825

Since real data set was not available for this problem so using above four populations from Muhammad Azeem & Muhammad Hanif (2016) mean square errors of the estimators in presence of non-response and measurement error were computed. We have also computed the percent relative efficiencies (PREs) of the various estimators with respect to usual unbiased estimator \bar{y}^* by using the formula:

$$PRE(t, \bar{y}^*) = \frac{MSE(\bar{y}^*)}{MSE(t)} \times 100, \quad (5.1)$$

where $t = t_r, t_{ra}, t_{bt}, t_{lr}, t_{sp}$

The findings are presented in the following tables :

Table 5.1: PRE's of estimators with respect to \bar{y}^* for population I.

N_1	N_2	Estimators	PRE with measurement error			PRE without measurement error		
			h=2	h=4	h=8	h=2	h=4	h=8
4500	500	\bar{y}^*	100	100	100	100	100	100
		t_r	586.0808	584.4539	582.4221	10,087.16	10,039.54	9,980.247
		t_{ra}	612.2445	610.7793	609.0847	10,096.39	10,051.12	9,995.222
		t_{bt}	297.5699	297.3415	297.055	390.6204	390.8468	391.132
		t_{lr}	611.3325	609.6874	607.6332	10,095.55	10,050.11	9,993.889
		t_{sp}	636.6662	639.9666	647.8538	13,360.24	14,161.74	16,129.83
		\bar{y}^*	100	100	100	100	100	100

4250	750	t_r	587.1388	587.1021	587.0624	10,248.54	10,427.13	10,627.41
		t_{ra}	613.3782	613.637	614.1666	10,258.23	10,439.78	10,644.02
		t_{bt}	297.7318	297.7451	297.7594	390.8179	391.2838	391.7889
		t_{lr}	612.4202	612.4071	612.3929	10,257.35	10,438.65	10,642.39
		t_{sp}	639.1152	646.7875	662.2565	13,868.23	15,764.65	20,800.16
4000	1000	\bar{y}^*	100	100	100	100	100	100
		t_r	593.61	601.4833	609.1461	10,254.12	10,419.62	10,581.45
		t_{ra}	619.36	626.9757	634.7779	10,260.05	10,423.84	10,584.87
		t_{bt}	298.1597	298.6841	299.1832	390.2786	390.0507	389.835
		t_{lr}	618.3516	625.5946	632.6514	10259.12	10422.57	10,582.91
		t_{sp}	646.6072	664.6857	693.5958	14,084.42	16,555.63	24,029.92

Table 5.2: PRE's of estimators with respect to \bar{y}^* for population II.

N_1	N_2	Estimators	PRE with measurement error			PRE with no measurement error		
			h=2	h=4	h=8	h=2	h=4	h=8
4500	500	\bar{y}^*	100	100	100	100	100	100
		t_r	246.8139	246.3256	245.7158	9,605.463	9,516.803	9,408.218
		t_{ra}	273.9812	273.851	273.8332	9,627.647	9,548.192	9,453.234
		t_{bt}	219.3226	219.2884	219.2455	384.3351	383.5184	382.5021
		t_{lr}	273.0122	272.69	272.2883	9,626.869	9,547.259	9,451.989
		t_{sp}	284.9563	287.0743	291.5858	12,010.92	12,431.58	13,462.85
4250	750	\bar{y}^*	100	100	100	100	100	100
		t_r	247.8297	248.7546	249.7567	9,712.634	9,775.975	9,844.793
		t_{ra}	275.0537	276.4282	278.1861	9,728.151	9,789.833	9,857.214
		t_{bt}	219.7493	220.2895	220.8733	385.095	385.3791	385.6841
		t_{lr}	274.0347	275.1171	276.2908	9,727.333	9,788.778	9,855.687
		t_{sp}	286.643	291.5057	300.313	12,355.68	13,484.17	16,238.12
4000	1000	\bar{y}^*	100	100	100	100	100	100
		t_r	244.4634	245.8319	247.1578	9,720.503	9,787.652	9,853.052
		t_{ra}	271.9547	273.1309	274.6739	9,739.698	9,809.953	9,878.888
		t_{bt}	218.5734	218.5575	218.5423	384.7701	384.6362	384.5076
		t_{lr}	270.8887	271.6787	272.4493	9,738.844	9,808.791	9,877.11
		t_{sp}	284.0312	289.4657	299.6114	12,498.39	13,965.25	17,801.28

Table 5.3: PRE's of estimators with respect to \bar{y}^* for population III.

N_1	N_2	Estimators	PRE with measurement error			PRE without measurement error		
			h=2	h=4	h=8	h=2	h=4	h=8
4500	500	\bar{y}^*	100	100	100	100	100	100
		t_r	579.6391	581.718	584.3259	10372.41	10322.25	10260.54
		t_{ra}	602.5743	604.575	607.2361	10381.95	10334.08	10275.63
		t_{bt}	294.0979	294.1538	294.2234	386.6437	386.3398	385.9626
		t_{lr}	601.5868	603.3876	605.649	10381.04	10332.99	10274.18
		t_{sp}	627.3482	634.4098	647.2312	13920.59	14780.38	16897.79
4250	750	\bar{y}^*	100	100	100	100	100	100
		t_r	577.7208	576.9875	576.1954	10441.01	10488	10539.43
		t_{ra}	601.0964	600.9336	601.0279	10446.07	10490.43	10541.19
		t_{bt}	294.1564	294.2815	294.4171	387.3762	388.0877	388.8621
		t_{lr}	600.063	599.6086	599.1197	10445.12	10489.22	10539.44
		t_{sp}	627.0971	634.4653	649.7234	14305.99	16064.17	20923.44
4000	1000	\bar{y}^*	100	100	100	100	100	100
		t_r	580.9977	584.2545	587.352	10368.01	10323.23	10281.49
		t_{ra}	603.9336	607.2549	610.8464	10378.7	10337.44	10299.79
		t_{bt}	294.1975	294.3565	294.5062	386.515	386.1235	385.7562
		t_{lr}	602.841	605.7524	608.5238	10377.69	10336.06	10297.66
		t_{sp}	631.4019	645.205	669.8931	14405.81	16541.35	23228.66

Table 5.4: PRE's of estimators with respect to \bar{y}^* for population IV.

N_1	N_2	Estimators	PRE with measurement error			PRE without measurement error		
			h=2	h=4	h=8	h=2	h=4	h=8
4500	500	\bar{y}^*	100	100	100	100	100	100
		t_r	236.0505	236.1577	236.2926	2604.001	2570.671	2529.699
		t_{ra}	264.4871	264.4013	264.5374	2611.026	2578.072	2537.761
		t_{bt}	215.3613	214.8848	214.2881	364.6081	364.2541	363.8071
		t_{lr}	262.8514	262.4443	261.9378	2609.706	2576.497	2535.675
		t_{sp}	285.011	288.7275	296.554	3010.02	3051.212	3167.222
4250	750	\bar{y}^*	100	100	100	100	100	100
		t_r	238.9165	242.8124	246.9824	2654.337	2690.853	2729.431
		t_{ra}	267.7365	272.083	277.2414	2661.282	2698.123	2737.467
		t_{bt}	216.8776	218.4431	220.088	365.1299	365.4987	365.8784
		t_{lr}	265.9904	269.7947	273.8686	2659.867	2696.261	2734.712

		t_{sp}	290.0738	301.9229	322.4102	3108.209	3322.39	3756.298
4000	1000	\bar{y}^*	100	100	100	100	100	100
		t_r	236.8203	237.8343	238.8079	2667.974	2719.149	2769.596
		t_{ra}	266.0461	268.056	270.6769	2676.232	2729.543	2782.804
		t_{bt}	216.3226	217.0976	217.8408	365.9304	367.2356	368.4837
		t_{lr}	264.2399	265.5873	266.8832	2674.771	2727.543	2779.726
		t_{sp}	289.2459	300.5426	322.2754	3149.877	3442.595	4052.483

From the Tables (5.1), (5.2), (5.3) and (5.4), it is clear that the proposed estimator t_{sp} has largest PRE (percentage relative efficiency) i.e. t_{sp} is the most efficient estimator among all other estimators considered in this paper. We have also computed PRE for estimators in presence of measurement error and in absence of measurement error. From our findings we conclude that estimators show unexpected increase in efficiency when measurement error is not considered. The PRE's (in case of without measurement error) becomes approximately double than the PRE's (in case of with measurement error).

6. CONCLUSIONS

In this paper, we have proposed an estimator t_{sp} in presence of non-response and measurement error and derived its MSE (mean square error) up to first order of approximation. We have also compared efficiency of estimators (in case of without measurement error) with efficiency of estimators (in case of with measurement error). From our empirical study we conclude that PRE (percentage relative efficiency) of our proposed estimator t_{sp} is maximum among all the estimators that we have considered here. We have also found that measurement error and non-response error effects PRE of estimators at high rate. As we can observe in Table (5.1), (5.2), (5.3) and (5.4) that PRE of estimators $t_r, t_{ra}, t_{lr}, t_{sp}$ has unexpectedly declined when measurement error is taken into account and it has approximately fallen down to half of the case when measurement error is not considered in estimation. However estimator t_{bt} is not as much affected as other estimators and it shows small decline in PRE in presence of measurement error. Hence we conclude that estimation of population parameters under the assumption that all the data is observable without any measurement error and non-response error is highly incorrect. As it is very clear in the tables (5.1), (5.2), (5.3) and (5.4) that the measurement error and non-response errors are heavily affecting estimators considered here. Our proposed estimator t_{sp} is efficient estimator as compared to usual estimator \bar{y}^* as well as other estimators considered here.

RECEIVED: NOVEMBER, 2016.

REVISED: FEBRUARY, 2017

Acknowledgement: This paper is an improvement of a previous version which was benefited from the suggestions of 2 anonymous referees. One of the authors report his acknowledgment of the support of Red CYTED 516RT0513 "RED IBEROAMERICANA DE AGROBIGDATA Y "DECISION SUPPORT SYSTEMS" (DSS) PARA UN SECTOR AGROPECUARIO SOSTENIBLE".

REFERENCES

- [1] BAHL, S. and TUTEJA, R.K. (1991): Ratio and product type exponential estimator. **Journal of Information and Optimization Sciences**, 12, 159-163.
- [2] BOUND, J., C. BROWN, and N. Mathiowetz (2001): Measurement Error in Survey Data, in Handbook of Econometrics, Vol. 5, ed. by J. J. Heckman, and E. E. Leamer. North Holland-n-response. **Assam Statist. Review**, 20, 73-83.
- [3] CARROLL, R. J., D. RUPPERT, and L. A. STEFANSKI (1995): **Measurement Error in Nonlinear Models: A Modern Perspective**. Chapman and Hall, London.
- [4] CHAUDHARY, M. K., PRAJAPATI, A., SINGH, R. and FLORENTIN, S. (2014) : **Two-phase sampling in estimation of population mean in the presence of non-response. The efficient use of supplementary information in finite population sampling**. Education Publishing, USA.

- [5] COCHRAN, W.G. (1968): Errors of measurement in statistics. **Technometrics**, 10, 637-666.
- [6] COCHRAN, W.G. (1977): **Sampling Techniques**, 3rd Edition. John Wiley & Sons, New York.
- [7] FULLER, W. (1987): Measurement Error Models. **Review of Economic Studies**, 72, 343-366
- [8] FULLER, W.A. (1995): Estimation in the presence of measurement error. **Int. Statist. Review**, 63, 121-147.
- [9] HANSEN, M. H. and HURWITZ, W. N. (1946): The problem of non-response in sample surveys. **Jour. of the Amer. Stat. Assoc.**, 41, 517-529.
- [10] KALTON, G. and KASPRZYK, D., (1986): The Treatment of Missing Data. *Survey Methodology*, 12, 1-16.
- [11] KAUFFMAN, G.M. and KING, B., (1973): Bayesian analysis of non-response in dichotomous processes. **J. Amer. Stat. Assoc.**, 68: 670-678
- [12] KENDALL, M. and STUART, A., (1977): **The advanced theory of statistics: distribution theory**, vol. I, Charles Griffin and co., London
- [13] KENWARD, M.G. and CARPENTER, J., (2007): Multiple imputations. **Current perspectives**. 16, 199-218
- [14] KHAN, S. A., ALI, H., MANZOOR, S. and ALAMGIR (2015): A class of transformed efficient ratio estimators of finite population mean. **Pakistan Journal of Statistics**, 31, 353-362.
- [15] Khare, B.B. and Sinha, R.R., 2004. Estimation of finite population ratio using two-phase sampling scheme in the presence of non-response. **Aligarh J. Stat.**, 24: 43-56
- [16] KHARE, B.B. and SRIVASTAVA, S. (1993): Estimation of population mean using auxiliary character in presence of non-response. **Nat. Acad. Sci. Lett. India**, 16, 111-114.
- [17] KHARE, B.B. and SRIVASTAVA, S. (1997): Transformed ratio type estimators for the population mean in the presence of non-response. **Comm. Statist. - Theory Methods** 26, 1779-1791.
- [18] KUMAR, M., SINGH, R., SAWAN, N. and CHAUHAN, P. (2011): Exponential ratio method of estimators in the presence of measurement errors. **International Journal of Agricultural and Statistical Sciences**, 7, 457-461.
- [19] KUMAR, M., SINGH, R., SINGH, A. K. and SMARANDACHE, F. (2011): Some ratio type estimators under measurement errors. **World Applied Sciences Journal**, 14, 272-276.
- [20] LIU, C. (1996): Bayesian robust multivariate linear regression with incomplete data. **J. Amer. Stat. Assoc.**, 91, 43-56
- [21] MANISHA, A. and SINGH, R. K. (2002): Role of regression estimator involving measurement errors. **Brazilian Journal of Probability and Statistics**, 16, 39-46.
- [22] OKAFOR, F. C. and LEE, H. (2000): Double sampling for ratio and regression estimation with sub-sampling the non-respondents. **Survey Methodology**, 26, 183-188.
- [23] OKAFOR, F. C. and LEE, H., (2000): Double sampling for ratio and regression estimation with sub sampling the non-respondent. **Survey Methodology**, 26, 183-188. 5, 1219-1227
- [24] RAO, P. S. R. S. (1986): Ratio Estimation with sub sampling the non-respondents. **Survey Methodology** 12, 217-230.
- [25] RAO, P. S. R. S. (1986): Ratio estimation with sub sampling the non respondents. **Survey Methodology**, 12, 217-230.
- [26] RAO, P. S. R. S. (1987): Ratio and regression estimates with sub-sampling the non respondents, Paper presented at a special contributed session of the **International Statistical Association Meetings**, September, Tokyo, Japan, 2-16.
- [27] RAO, T. (1991): On certain methods of improving ratio and regression estimators. **Communications in Statistics-Theory and Methods**, 20, 3325-3340.
- [28] SHABBIR, J. and KHAN, N.S. (2013): Some modified exponential ratio-type estimators in the presence of non-response under two-phase sampling scheme. **Electron. J. A Stat. Anal.**, 6, 1-17.
- [29] SHALAB, H (1997): Ratio method of estimation in the presence of measurement errors. **J. Ind. Soc. Agri. Statist.**, 50, 150-155.
- [30] SHARMA, P. and SINGH, R. (2015) : Method of estimation in the presence of non-response and measurement errors simultaneously. **Jour. of Moder. Appl. Stat. Meth.**, 14,1,107-121.
- [31] SINGH, H.P. and KARPE, N. (2008): Ratio-product estimator for population mean in presence of measurement errors. **J. Appl. Statist. Sci.**, 16, 49-64.
- [32] SINGH, H.P. and KUMAR, S. (2011): Combination of regression and ratio estimate in presence of non-response. **Braz. J. Statist. Assoc.**, 25, 205-217.
- [33] SINGH, R., MALIK, S. and KOHSHNIVESAN, M. (2014) : An alternative estimator for estimating the finite population mean in presence of measurement errors with the view to financial modeling. **Sci. Jour. Of Appl. Math. and Stat.**, 2, 107-111.
- [34] SINGH, V.K. and SINGH, R. (2015) : Estimation of mean using dual-to-ratio and difference type estimators under measurement error model. **JSAPL** 2, 137-148.

- [35] SODIPO, A. A. and OBISESAN, K. O., 2007. Estimation of the population mean using difference cum ratio estimator with full response on the auxiliary character. **Res. J. Applied Sci.**, 2, 769-772.
- [36] SRIVASTAVA, A., K., and SHALAB, H (2001): Effect of measurement errors on the regression method of estimation in survey sampling. **Journal of Statistical Research**, 35, 35-44..
- [37] SUD, U. C. and SRIVASTAVA, A. K. (2000): Estimation of population mean in repeated surveys in the presence of measurement errors. **Journal of the Indian Society of Agricultural Statistics**, 53, 125-133.
- [38] TABASUM, R. and KHAN, I.A. (2004): Double sampling for ratio estimation with non-response. **J. Ind. Statist. Assoc.**, 33, 33-54.