Pakistan Journal of Statistics and Operation Research

Estimation of Population Mean Using Three-Stage Optional RRT Model in the Presence of Measurement Errors under Stratified Two-Phase Sampling

Ronald Onyango[1*], Brian Oduor[2], Francis Odundo[3]

* Corresponding Author

1. Department of Applied Statistics, Financial Mathematics and Actuarial Science, Jaramogi Oginga Odinga University of Science and Technology, Kenya, assangaronald@gmail.com
2. Department of Applied Statistics, Financial Mathematics and Actuarial Science, Jaramogi Oginga Odinga University of Science and Technology, Kenya, oduobriano@gmail.com
3. Department of Applied Statistics, Financial Mathematics and Actuarial Science, Jaramogi Oginga Odinga University of Science and Technology, Kenya, okellof2001@yahoo.com

Abstract

In the present study, the problem of estimation of the finite population mean of a sensitive study variable using the three-stage optional Randomized Response Technique (RRT) model under measurement errors is addressed. A generalized class of estimators is proposed using a mixture of auxiliary attribute and variable. Some members of the proposed generalized class of estimators are identified and studied. The bias and mean square error expressions for the proposed estimators are correctly derived up to first order Taylor's series of approximation. The proposed estimator's efficiency is investigated theoretically and numerically using both real data and simulated data. From the numerical study, the proposed estimators outperforms other existing estimators of the finite population mean. Furthermore, efficiencies of the proposed estimators of the finite population mean decreases as sensitivity level of the survey question increases.

Key Words: Auxiliary attributes; Measurement errors; Sensitivity level; Bias

Mathematical Subject Classification: 62D05

1. Introduction

In a survey, a researcher faces the problem of estimation of the finite population mean of a study variable using auxiliary variable in the presence of measurement errors. Measurement errors are the differences between the true value of a variable and the value recorded in a survey. Under-reporting, over-reporting, memory loss, prestige bias, processing errors, and incorrect respondent values are some of the causes of measurement errors in a survey. Shalabh (1997), Diwakar et al. (2012), Yadav et al. (2017), Vishwakarma et al. (2020), Singh and Karpe (2010), Kumar et al. (2011), and Shukla et al. (2012) all discuss the problem of estimation of the finite population mean of a non-sensitive variable using an auxiliary variable. Aside from measurement errors, researchers must contend with the issue of estimating the population mean of a sensitive survey question with a social stigmatizing characteristic. Personal income, alcohol consumption, abortion, tax evasion, number of sexual partners, negative web site usage, homosexuality, reckless driving, indiscriminate gambling, domestic violence, and illicit drug use are just a few

 DOI: http://dx.doi.org/10.18187/pjsor.v19i1.3929

examples. Correct responses to such sensitive study variables are difficult to obtain in personal interviews that involve direct questioning of people, because the respondent's privacy is not protected. In reality, the majority of respondents are always hesitant to provide honest response to a contentious topic for fear of embarrassment or loss of status. As a result, the respondent will either refuse to answer the question or will purposefully provide an incorrect response.

Warner (1965) developed the Randomized Response Technique (RRT) to reduce response bias in surveys involving a sensitive study variable by protecting respondents' anonymity. Randomized Response Technique (RRT) uses a scrambling variable that is independent of the survey and auxiliary variables to estimate the mean of a sensitive study variable. The respondent must provide a genuine response to a non-sensitive auxiliary variable while providing a scrambled response to the study variable. In additive scrambling Randomized Response Technique (RRT) model (Pollock and Bek 1976), the respondent is expected to scramble the genuine answer to a sensitive question by adding a random integer. The value-added is unknown to the survey practitioners, but the probability distribution of the scrambled response is assumed to be known. The optional Randomized Response Technique (ORRT) was pioneered by Chaudhuri and Mukharjee (1988). If a respondent believes the question is sensitive, the strategy involves giving them the option of providing a direct or scrambled response. Gupta et al (2006) proposed a one-stage optional Randomized Response Technique (ORRT) model in which the respondent provides a direct response if the question is not sensitive and a scrambled response otherwise. Gupta et al. (2012) proposed a two-stage optional Randomized Response Technique (ORRT) model to increase respondent participation and privacy.

Mehta et al. (2012) proposed a three-stage optional Randomized Response Technique (ORRT) model to encourage respondent cooperation and privacy. In the first stage, a predetermined number of respondents, t_h are asked to provide a direct response to a sensitive subject. Thereafter, another predetermined proportion, f_h is asked to scramble their response in the second stage. The remaining proportion, $(1 - t_h - f_h)$ is then given the option of providing a direct or scrambled response. According to Neeraj and Mehta (2017) the three-stage optional RRT model ensures greater respondent cooperation and privacy.

Mushtaq and Noor-Ul-Amin (2020), Shabbir and Naeem (2018), and Shabbir and Zahid (2019) discuss the problem of estimation of the finite population mean of a sensitive variable using non-sensitive auxiliary variable based on non-optional Randomized Response Technique (RRT) model. Khalil (2018) proposed a generalized estimator in the presence of measurement errors based on non-optional Randomized Response Technique (RRT) model. Under simple random sampling, Khalil et al. (2019) studied the problem of estimation of the finite population mean of a sensitive study variable in the presence of measurement errors using a one-stage optional Randomized Response Technique (ORRT) model. Onyango et al. (2021) recently studied the problem of estimation of the finite population mean under measurement errors using the non-optional Randomized Response Technique (RRT) model in stratified two-phase sampling.

In the literature, little work has been done on the topic of estimation of the finite population mean of a sensitive study variable in the presence of measurement errors using the three-stage optional RRT model. Additionally, the impact of measurement errors and three-stage optional RRT model on estimation of the finite population mean has not been investigated.

Rest of the article is organized in the following way. Section 2 of this paper provides a detailed description of the population under study. Section 3 discusses some of the existing estimators of population mean in the literature. Section 4 describes the properties of the proposed estimators of population mean. Section 5 investigates the theoretical efficiency of the proposed estimator. A numerical study of the performance of the proposed estimators is done in section 6. Finally, section 7 contains the conclusions of the study.

2. Population description and notations

Consider a heterogeneous population $U = U_1, U_2, \dots, U_N$ of size N that is divided into L homogeneous strata, each of which contains N_h units. The population is made up of a sensitive study variable, auxiliary variable and scrambled response denoted as Y, X, and Z respectively. Let $\bar{Z}_h \; and \; \bar{X}_h$ denote the population means of the scrambled response and auxiliary variable in the h^{th} stratum respectively. Furthermore, let A_{hj} denote the value of j^{th} attribute for i^{th} unit in the h^{th} stratum. The auxiliary attribute takes the values 1 and 0 if i^{th} population unit possesses and does not possess an attribute respectively. Furthermore, let $A_{hj} = \sum_{h}^{N_h} \tau_{hij}$ and $P_h = \frac{A_{hj}}{N_h}$, denote the total number of units that have an attribute and proportion of units possessing an attribute in the h^{th} stratum respectively. In addition, let $S_{Zh}^2, S_{Ph}^2, and\ S_{Xh}^2$ denote the population variances of the scrambled response, auxiliary attribute and variable in the h^{th} stratum respectively. Let $S_{Xh}, S_{ZPh}, and\ S_{XPh}$ denote the population covariance's between their subscripts in the h^{th} stratum. Moreover, let $\rho_{ZXh}, \rho_{Zph}, and\ \rho_{XPh}$ denote the population coefficient of correlation between their subscripts in the h^{th} stratum. Let $(z_{hi}^*, x_{hi}^*)\ and\ (Z_{hi}^*, X_{hi}^*)$ denote the observed and true values respectively for the scrambled response and auxiliary variable in the h^{th} stratum in the presence of measurement errors. Let $T_{hi}^* = z_{hi}^* - Z_{hi}^*$, and $V_{hi}^* = x_{hi}^* - X_{hi}^*$ denote the measurement errors associated with the scrambled response and auxiliary variable in the h^{th} stratum. The measurement errors are assumed to occur randomly and to be independent with a mean of zero. Additionally, the measurement errors are independent of the sensitive study and non-sensitive auxiliary variables. Let $S_{Th}^2\ and\ S_{Vh}^2$ denote the population variances of the measurement errors associated with the scrambled response and auxiliary variable in the h^{th} stratum respectively.

In the literature researchers have used conventional and non-conventional measures of auxiliary variable to develop efficient estimators of population mean in the presence of extreme values. For more details see Almanjahie et al. (2021), Subzar et al. (2018), and Shabbir et al. (2021).

The coefficient of variation of an auxiliary variable is given as $C_{Xh} = \frac{S_{Xh}}{\bar{X}_h}$.

The coefficient of skewness is defined as $\beta_{1h}(x) = \frac{N_h \sum_{h=1}^{L}(x_{hi} - \bar{X}_h)^3}{(N_h - 1)(N_h - 2)S_{Xh}^3}$

The coefficient of kurtosis is defined as $\beta_{2h}(x) = \frac{N_h(N_h+1)\sum_{h=1}^{L}(x_{hi} - \bar{X}_h)^4}{(N_h - 1)(N_h - 2)(N_h - 3)S_{Xh}^3} - \frac{3(N_h - 1)^2}{(N_h - 2)(N_h - 3)}$

The mid-range is given as $MR_h(x) = \frac{\left(x_{h(1)} + x_{h(N_h)}\right)}{2}$, where $x_{h(1)}$ is the smallest value and $x_{h(N_h)}$ is the largest value in a data set.

Tri-mean was proposed by Turkey (1970) and is defined as $TM_h(x) = \frac{Q_{1h}(x) + 2Q_{2h}(x) + Q_{3h}(x)}{4}$, where $Q_{1h}(x)$ and $Q_{2h}(x), and\ Q_{3h}(x)$ are first, second and third quartiles respectively.

The quartile deviation is defined as $QD_h(x) = \frac{(Q_{3h}(x) - Q_{1h}(x))}{2}$.

The Hodge-Lehmann (1963) estimator is defined as $H_{Lh}(x) = median \frac{x_h(j) + x_h(k)}{2}$, where $1 \le j \le k \le N_h$.

A relatively large sample of size of n_h' is drawn from the h^{th} stratum using simple random sampling without replacement (SRSWOR). Let $\bar{x}_h' = \frac{1}{n_h'}\sum_{i=1}^{n_h'} x_{hi}$ and $p_h' = \frac{a_{hj}}{n_h'}$ be the sample mean of the auxiliary variable and the proportion of units possessing an auxiliary attribute in the first phase sample respectively. A second phase random sample of size n_h is drawn from the first phase sample using a simple random sampling without replacement (SRSWOR). Furthermore, let $\bar{z}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} z_{hi}, \bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$ and $p_h = \frac{a_{hj}}{n_h}$ be the sample mean of scrambled response, auxiliary variable, and the proportion of units in the second phase sample that have an auxiliary attribute respectively.

3. Some Existing Estimators

(i) The ordinary estimator is defined as

$$t_0 = \sum_{h=1}^{L} W_h \bar{z}_h \tag{1}$$

The variance of the estimator is given as

$$Var(t_0) \cong \sum_{h=1}^{L} W_h^2 B_h \tag{2}$$

, where $B_h = \theta_h(S_{Zh}^2 + S_{Th}^2)$ and $\theta_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$.

(ii) The usual ratio estimator is defined as

$$t_R = \sum_{h=1}^{L} W_h \bar{z}_h \frac{\bar{x}'_h}{\bar{x}_h} \tag{3}$$

The bias and MSE are given as

$$Bias(t_R) \cong \sum_{h=1}^{L} \frac{W_h}{\bar{X}_h}\left[\frac{9}{8} R_h (A_h - C_h) - (E_h - D_h)\right] \tag{4}$$

, and

$$MSE(t_R) \cong \sum_{h=1}^{L} W_h^2 [B_h + R_h^2 (A_h - C_h) - 2R_h (E_h - D_h)] \tag{5}$$

, respectively, where $A_h = \theta_h (S_{Xh}^2 + S_{Vh}^2)$, $C_h = \theta'_h S_{Xh}^2$, $D_h = \theta'_h S_{ZXh}^2$, $E_h = \theta_h S_{ZXh}$, and $\theta'_h = \left(\frac{1}{n'_h} - \frac{1}{N_h}\right)$.

(iii) Bahl and Tuteja (1991) exponential-ratio type estimator is defined as

$$t_{ER} = \sum_{h=1}^{L} W_h \bar{z}_h exp\left(\frac{\bar{x}'_h - \bar{x}_h}{\bar{x}'_h + \bar{x}_h}\right) \tag{6}$$

The bias and MSE are given as

$$Bias(t_{ER}) \cong \sum_{h=1}^{L} \frac{W_h}{2\bar{X}_h}\left[\frac{3}{4} R_h (A_h - C_h) - (E_h - D_h)\right] \tag{7}$$

, and

$$MSE(t_{ER}) \cong \sum_{h=1}^{L} W_h^2 \left[B_h + \frac{1}{4} R_h^2 (A_h - C_h) - R_h (E_h - D_h)\right] \tag{8}$$

, respectively

4. Proposed Strategy of Estimation of Finite Population Mean

In the three-stage optional RRT model, a respondent is required to provide a scrambled response defined as

$$Z_{hi} = \begin{cases} Y_{hi} \;\; with\, pobability\; t_h + (1 - t_h - f_h)(1 - \psi_h) \\ Y_{hi} + S_{hi}, with\, probability\; f_h + (1 - t_h - f_h) \end{cases} \tag{9}$$

, where ψ_h and S_{hi} are the sensitivity level and scrambling variable respectively. We assume that $S_{hi} \sim N(0, S_{Sh}^2)$. The mean of the scrambled response is given as

$$E(Z_{hi}) = E[(Y_{hi})(1 - \varphi_h) + (Y_{hi} + S_{Shi})\varphi_h] \tag{10}$$

, where $\varphi_h = f_h + (1 - t_h - f_h)$

$$E(Z_{hi}) = E[\,(Y_{hi}) - Y_{hi}\varphi_h + Y_{hi}\varphi_h + S_{Shi}\varphi_h] \tag{11}$$

$$E(Z_{hi}) = E(Y_{hi}) \tag{12}$$

The variance of the scrambled response is given as

$$S_{Zh}^2 = E(Z_{hi}^2) - [E(Z_{hi})]^2 \tag{13}$$

$$S_{Zh}^2 = E(Y_{hi}^2)(1 - \varphi_h) + E[(Y_{hi} + S_{Shi})^2]\varphi_h - [E(Z_{hi})]^2 \tag{14}$$

$$S_{Zh}^2 = E(Y_{hi}^2) - E(Y_{hi}^2)\varphi_h + E(Y_{hi}^2)\varphi_h + 2E(Y_{hi}S_{Shi})\varphi_h + E(S_{Shi})^2\varphi_h - [E(Z_{hi})]^2 \tag{15}$$

$$S_{Zh}^2 = S_{Yh}^2 + \varphi_h S_{Sh}^2 \tag{16}$$

The three stage optional RRT model reduces to one stage optional RRT model and two stage optional RRT model when $\psi_h = 0$ *and* $\psi_h = 1$ respectively.

4.1 Proposed generalized class of estimators

Let W_h denote the weight of the h^{th} stratum. Further let $\bar{x}'_h$ and p'_h denote the mean of an auxiliary variable and proportion of units possessing an attribute in the first phase h^{th} stratum sample respectively. In addition, let

 DOI: http://dx.doi.org/10.18187/pjsor.v19i1.3929

$\bar{z}_h, \bar{x}_h, and\ p_h$ denote the means of the scrambled response, auxiliary variable and proportion of units possessing an attribute in the second phase in h^{th} stratum sample respectively. The proposed generalized class of estimators is defined as

$$t_g = \sum_{h=1}^{L} W_h[\bar{z}_h + \alpha_h(\bar{x}'_h - \bar{x}_h) + \beta_h(p'_h - p_h)]exp\left(\frac{a_h(\bar{x}'_h - \bar{x}_h)}{a_h(\bar{x}'_h - \bar{x}_h) + 2b_h}\right) \quad (17)$$

, where $\alpha_h\ and\ \beta_h$ are suitable constants, a_h and $b_h$are either real numbers or known population parameters of an auxiliary variable.

To obtain the expressions for the bias and MSE of the proposed estimator, let

$$\sigma_{X1h} = \bar{x}'_h - \bar{X}_h, \quad (18)$$

$$\sigma_{P1h} = p'_h - P_h, \quad (19)$$

$$\sigma_{Xh} = \bar{x}_h - \bar{X}_h, \quad (20)$$

$$\sigma_{Ph} = p_h - P_h, \quad (21)$$

$$\sigma_{Zh} = \bar{z}_h - \bar{Z}_h \quad (22)$$

Take expectations on both sides of equations (18) - (22) to obtain

$$E(\sigma_{Zh}) = E(\sigma_{Xh}) = E(\sigma_{X1h}) = E(\sigma_{Ph}) = E(\sigma_{P1h}) = 0 \quad (23)$$

Square both sides of equations (18) - (22) and then introduce expectations to obtain

$$\mathrm{E}(\sigma^2_{\mathrm{Xh}}) = \theta_{\mathrm{h}}(\mathrm{S}^2_{\mathrm{Xh}} + \mathrm{S}^2_{\mathrm{Vh}}) = \mathrm{A_h} \quad (24)$$

$$\mathrm{E}(\sigma^2_{\mathrm{Zh}}) = \theta_{\mathrm{h}}(\mathrm{S}^2_{\mathrm{Zh}} + \mathrm{S}^2_{\mathrm{Th}}) = \mathrm{B_h} \quad (25)$$

$$\mathrm{E}(\sigma^2_{\mathrm{X1h}}) = \mathrm{E}(\sigma_{X1h}\sigma_{Xh}) = \theta'_{\mathrm{h}}\mathrm{S}^2_{\mathrm{Xh}} = \mathrm{C_h} \quad (26)$$

$$\mathrm{E}(\sigma_{X1h}\sigma_{Zh}) = \theta'_{\mathrm{h}}\mathrm{S}^2_{\mathrm{ZXh}} = \mathrm{D_h} \quad (27)$$

$$\mathrm{E}(\sigma_{Xh}\sigma_{Zh}) = \theta_{\mathrm{h}}\mathrm{S}^2_{\mathrm{ZXh}} = \mathrm{E_h} \quad (28)$$

$$\mathrm{E}(\sigma^2_{\mathrm{Ph}}) = \theta_{\mathrm{h}}\mathrm{S}^2_{\mathrm{Ph}} = \mathrm{F_h} \quad (29)$$

$$\mathrm{E}(\sigma^2_{\mathrm{P1h}}) = \mathrm{E}(\sigma_{P1h}\sigma_{Ph}) = \theta'_{\mathrm{h}}\mathrm{S}^2_{\mathrm{Ph}} = \mathrm{G_h} \quad (30)$$

$$\mathrm{E}(\sigma_{Ph}\sigma_{Zh}) = \theta_{\mathrm{h}}\mathrm{S}^2_{\mathrm{ZPh}} = \mathrm{H_h} \quad (31)$$

$$\mathrm{E}(\sigma_{P1h}\sigma_{Zh}) = \theta'_{\mathrm{h}}\mathrm{S}^2_{\mathrm{ZPh}} = \mathrm{I_h} \quad (32)$$

$$\mathrm{E}(\sigma_{Ph}\sigma_{Xh}) = \theta_{\mathrm{h}}\mathrm{S}^2_{\mathrm{PXh}} = \mathrm{J_h} \quad (33)$$

$$\mathrm{E}(\sigma_{P1h}\sigma_{Xh}) = \mathrm{E}(\sigma_{Ph}\sigma_{X1h}) = \mathrm{E}(\sigma_{X1h}\sigma_{P1h}) = \theta'_{\mathrm{h}}\mathrm{S}^2_{\mathrm{XPh}} = \mathrm{L_h} \quad (34)$$

Substitute equations (18) - (22) in (17) and expand using Taylor's approximation while ignoring terms of order greater than two to obtain

(35)

$$t_g = \sum_{h=1}^{L} W_h \begin{bmatrix} \bar{Z}_h - \frac{1}{2}\bar{Z}_h\lambda_h\sigma_{Xh} + \frac{1}{2}\bar{Z}_h\lambda_h\sigma_{X1h} + \frac{3}{8}\bar{Z}_h\lambda_h^2\sigma_{Xh}^2 - \frac{1}{4}\bar{Z}_h\lambda_h^2\,\sigma_{Xh}\sigma_{X1h} - \frac{1}{8}\,\bar{Z}_h\lambda_h^2\sigma_{X1h}^2 \\ +\sigma_{Zh} - \frac{1}{2}\lambda_h\sigma_{Xh}\sigma_{Zh} + \frac{1}{2}\lambda_h\sigma_{X1h}\sigma_{Zh} + \alpha_h\sigma_{X1h} - \frac{1}{2}\lambda_h\alpha_h\sigma_{Xh}\sigma_{X1h} + \\ \frac{1}{2}\lambda_h\alpha_h\sigma_{X1h}^2 - \alpha_h\sigma_{Xh} + \frac{1}{2}\lambda_h\alpha_h\sigma_{Xh}^2 - \frac{1}{2}\lambda_h\alpha_h\sigma_{Xh}\sigma_{X1h} + \beta_h\sigma_{P1h} \\ -\frac{1}{2}\lambda_h\beta_h\sigma_{Xh}\sigma_{p1h} + \frac{1}{2}\lambda_h\beta_h\sigma_{X1h}\sigma_{p1h} - \beta_h\sigma_{Ph} + \frac{1}{2}\lambda_h\beta_h\sigma_{Xh}\sigma_{p1h} \\ -\frac{1}{2}\lambda_h\beta_h\sigma_{Ph}\sigma_{X1h} \end{bmatrix}$$

, where $\lambda_h = \frac{a_h}{a_h\bar{X}_h + b_h}$.

Subtract the population mean from both sides of the equation (35) to obtain

$$\left(t_g - \bar{Y}\right) = \sum_{h=1}^{L} W_h \begin{bmatrix} \bar{Z}_h\lambda_h\sigma_{Xh} + \frac{1}{2}\bar{Z}_h\lambda_h\sigma_{X1h} + \frac{3}{8}\bar{Z}_h\lambda_h^2\sigma_{Xh}^2 - \frac{1}{4}\bar{Z}_h\lambda_h^2\,\sigma_{Xh}\sigma_{X1h} - \frac{1}{8}\,\bar{Z}_h\lambda_h^2\sigma_{X1h}^2 \\ +\sigma_{Zh} - \frac{1}{2}\lambda_h\sigma_{Xh}\sigma_{Zh} + \frac{1}{2}\lambda_h\sigma_{X1h}\sigma_{Zh} + \alpha_h\sigma_{X1h} - \frac{1}{2}\lambda_h\alpha_h\sigma_{Xh}\sigma_{X1h} + \\ \frac{1}{2}\lambda_h\alpha_h\sigma_{X1h}^2 - \alpha_h\sigma_{Xh} + \frac{1}{2}\lambda_h\alpha_h\sigma_{Xh}^2 - \frac{1}{2}\lambda_h\alpha_h\sigma_{Xh}\sigma_{X1h} + \beta_h\sigma_{P1h} \\ -\frac{1}{2}\lambda_h\beta_h\sigma_{Xh}\sigma_{p1h} + \frac{1}{2}\lambda_h\beta_h\sigma_{X1h}\sigma_{p1h} - \beta_h\sigma_{Ph} + \frac{1}{2}\lambda_h\beta_h\sigma_{Xh}\sigma_{p1h} \\ -\frac{1}{2}\lambda_h\beta_h\sigma_{Ph}\sigma_{X1h} \end{bmatrix} \tag{36}$$

Take expectations on both sides of the equation (36) and substitute equations (23) - (34) to obtain an approximation for the bias as

$$Bias\left(t_g\right) \cong \sum_{h=1}^{L} \frac{W_h\lambda_h}{2}\left[\frac{3}{4}\lambda_h\bar{Z}_h(A_h - C_h) + \alpha_h(A_h - C_h) - (E_h - D_h) + \beta_h(J_h - L_h)\right] \tag{37}$$

Square both sides of equation (36) to obtain

$$\left(t_g - \bar{Y}\right)^2 \cong \sum_{h=1}^{L} W_h^3\,\left[\frac{1}{2}\bar{Z}_h\lambda_h\sigma_{X1h} - \frac{1}{2}\bar{Z}_h\lambda_h\sigma_{Xh} + \sigma_{Zh} + \alpha_h\sigma_{X1h} - \alpha_h\sigma_{Xh} + \beta_h\sigma_{P1h} - \alpha_h\sigma_{Ph}\right]^2 \tag{38}$$

Simplify equation (38) while ignoring terms of order greater than two to obtain

$$\left(t_g - \bar{Y}\right)^2 \cong \sum_{h=1}^{L} W_h^3 \begin{bmatrix} \sigma_{Zh}^2 + \frac{1}{4}\lambda_h^2\bar{Z}_h^2\sigma_{Xh}^2 + \frac{1}{4}\lambda_h^2\bar{Z}_h^2\sigma_{X1h}^2 + \alpha_h^2\sigma_{X1h}^2 + \alpha_h^2\sigma_{Xh}^2 + \beta_h^2\sigma_{P1h}^2 + \beta_h^2\sigma_{Ph}^2 \\ -\bar{Z}_h\lambda_h\sigma_{Xh}\sigma_{Zh} + \bar{Z}_h\lambda_h\sigma_{X1h}\sigma_{Zh} + 2\alpha_h\sigma_{Zh}\sigma_{X1h} - 2\alpha_h\sigma_{Zh}\sigma_{Xh} + 2\beta_h\sigma_{P1h}\sigma_{Zh} \\ -2\beta_h\sigma_{Ph}\sigma_{Zh} - \frac{1}{2}\,\lambda_h^2\bar{Z}_h^2\sigma_{X1h}\sigma_{Xh} - \bar{Z}_h\lambda_h\alpha_h\sigma_{Xh}\sigma_{X1h} + \bar{Z}_h\lambda_h\alpha_h\sigma_{Xh}^2 \\ -\bar{Z}_h\lambda_h\beta_h\sigma_{P1h}\sigma_{Xh} + \bar{Z}_h\lambda_h\beta_h\sigma_{Ph}\sigma_{Xh} + \bar{Z}_h\lambda_h\alpha_h\sigma_{X1h}^2 - \bar{Z}_h\lambda_h\alpha_h\sigma_{X1h}\sigma_{Xh} \\ +\bar{Z}_h\lambda_h\beta_h\sigma_{X1h}\sigma_{P1h} - \bar{Z}_h\lambda_h\beta_h\sigma_{X1h}\sigma_{Ph} - 2\alpha_h^2\,\sigma_{X1h}\sigma_{Xh} + 2\beta_h\alpha_h\sigma_{X1h}\sigma_{P1h} \\ -2\beta_h\alpha_h\sigma_{X1h}\sigma_{Ph} - 2\beta_h\alpha_h\sigma_{Xh}\sigma_{P1h} + 2\beta_h\alpha_h\sigma_{Xh}\sigma_{Ph} - 2\beta_h^2\sigma_{P1h}\sigma_{Ph} \end{bmatrix} \tag{39}$$

Take expectations on both sides of equation (39) and substitute equations (24) - (34) to obtain an approximation for the MSE as

$$MSE\left(t_g\right) \cong \sum_{h=1}^{L} W_h^3\,[B_h + \gamma_{1h} + \alpha_h^2\gamma_{2h} + \beta_h^2\gamma_{3h} + \beta_h\gamma_{4h} + \alpha_h\gamma_{5h} + 2\beta_h\,\alpha_h\gamma_{5h}] \tag{40}$$

, where

$$\gamma_{1h} = \frac{1}{4}\bar{Z}_h^2\lambda_h^2(A_h - C_h) - \bar{Z}_h\lambda_h(E_h - D_h)$$

$$\gamma_{2h} = (A_h - C_h)$$

$$\gamma_{3h} = (F_h - G_h)$$

$$\gamma_{4h} = \bar{Z}_h\lambda_h(J_h - L_h) - 2(H_h - I_h)$$

$$\gamma_{5h} = \bar{Z}_h\lambda_h(A_h - C_h) - 2(E_h - D_h)$$

$$\gamma_{6h} = (J_h - L_h)$$

Differentiate equation (40) partially with respect to α_h and β_h, then equate to zero to obtain

$$\alpha_h^{(opt)} = -\frac{(\gamma_{5h}\gamma_{3h} - \gamma_{4h}\gamma_{6h})}{2(\gamma_{2h}\gamma_{3h} - \gamma_{6h}^2)} \qquad (41)$$

$$\beta_h^{(opt)} = \frac{(\gamma_{5h}\gamma_{6h} - \gamma_{4h}\gamma_{2h})}{2(\gamma_{2h}\gamma_{3h} - \gamma_{6h}^2)} \qquad (42)$$

Substitute equation (41) and (42) in (40) to obtain the minimum MSE as

$$MSE(t_g)_{min} \cong \sum_{h=1}^{L} W_h^3 \left[B_h + \gamma_{1h} - \frac{\gamma_{4h}^2}{4\gamma_{3h}} - \frac{(\gamma_{5h}\gamma_{3h} - \gamma_{4h}\gamma_{6h})^2}{(\gamma_{2h}\gamma_{3h} - \gamma_{6h}^2)}\right] \qquad (43)$$

4.2 Members of the family of proposed estimator

Members of the proposed generalized class of estimators can be obtained by making appropriate choices of α_h *and* β_h. Table 1 shows some special cases of the proposed generalized class of estimators. Expressions for the bias and mean squared errors (MSE) for members of the proposed generalized class of estimators are obtained by substituting appropriate values of α_h *and* β_h in equations (37) and (43) respectively.

Table 1: Some Members of the Developed Generalized Class of Estimators.

No.	proposed generalized class of estimators	α_h	β_h
1	$t_0 = \sum_{h=1}^{L} W_h \bar{z}_h$	0	1
2	$t_1 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{(\bar{x}'_h - \bar{x}_h)}{(\bar{x}'_h - \bar{x}_h)}\right)$	1	0
3	$t_2 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{a_h(\bar{x}'_h - \bar{x}_h)}{(\bar{x}'_h - \bar{x}_h) + 2C_{Xh}}\right)$	1	C_{Xh}
4	$t_3 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{C_{Xh}(\bar{x}'_h - \bar{x}_h)}{C_{Xh}(\bar{x}'_h - \bar{x}_h) + 2\rho_{YXh}}\right)$	C_{Xh}	ρ_{YXh}

5	$t_4 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{\beta_{1h}(x)(\bar{x}_h' - \bar{x}_h)}{\beta_{1h}(x)(\bar{x}_h' - \bar{x}_h) + 2\rho_{YXh}}\right)$	$\beta_{1h}(x)$	ρ_{YXh}
6	$t_5 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{\beta_{2h}(x)(\bar{x}_h' - \bar{x}_h)}{\beta_{2h}(x)(\bar{x}_h' - \bar{x}_h) + 2\beta_{1h}(x)}\right)$	$\beta_{2h}(x)$	$\beta_{1h}(x)$
7	$t_6 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{QD_h(x)(\bar{x}_h' - \bar{x}_h)}{QD_h(x)(\bar{x}_h' - \bar{x}_h) + 2TM_h(x)}\right)$	$QD_h(x)$	$TM_h(x)$
8	$t_7 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{QD_h(x)(\bar{x}_h' - \bar{x}_h)}{QD_h(x)(\bar{x}_h' - \bar{x}_h) + 2MR_h(x)}\right)$	$QD_h(x)$	$MR_h(x)$
9	$t_8 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{HL_h(x)(\bar{x}_h' - \bar{x}_h)}{HL_h(x)(\bar{x}_h' - \bar{x}_h) + 2TM_h(x)}\right)$	$HL_h(x)$	$TM_h(x)$
10	$t_9 = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{\rho_{YXh}(\bar{x}_h' - \bar{x}_h)}{\rho_{YXh}(\bar{x}_h' - \bar{x}_h) + 2QD_h(x)}\right)$	ρ_{YXh}	$QD_h(x)$
11	$t_{10} = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{(\bar{x}_h' - \bar{x}_h)}{(\bar{x}_h' - \bar{x}_h) + 2\rho_{YXh}}\right)$	1	ρ_{YXh}
12	$t_{11} = \sum_{h=1}^{L} W_h \xi_h exp\left(\frac{(\bar{x}_h' - \bar{x}_h)}{(\bar{x}_h' - \bar{x}_h) + 2QD_h(x)}\right)$	1	$QD_h(x)$

, where $\xi_h = \bar{z}_h + \alpha_h(\bar{x}_h' - \bar{x}_h) + \beta_h(p_h' - p_h)$

5. Theoretical Efficiency Comparison

The proposed generalized class of estimators performs better than other existing estimators when following conditions are satisfied

(i) From equations (2) and (43), $MSE(t_g)_{min} < Var(t_0)$ if

$$\left[\gamma_{1h} - \frac{\gamma_{4h}^2}{4\gamma_{3h}} - \frac{(\gamma_{5h}\gamma_{3h} - \gamma_{4h}\gamma_{6h})^2}{(\gamma_{2h}\gamma_{3h} - \gamma_{6h}^2)}\right] < 0 \quad (44)$$

(ii) From equations (5) and (43), $MSE(t_g)_{min} < MSE(t_R)$ if

$$\left[\gamma_{1h} - \frac{\gamma_{4h}^2}{4\gamma_{3h}} - \frac{(\gamma_{5h}\gamma_{3h} - \gamma_{4h}\gamma_{6h})^2}{(\gamma_{2h}\gamma_{3h} - \gamma_{6h}^2)} - R_h^2(A_h - C_h) + 2R_h(E_h - D_h)\right] < 0 \quad (45)$$

(iii) From equations (8) and (43), $MSE(t_g)_{min} < MSE(t_{ER})$ if

$$\left[\gamma_{1h} - \frac{\gamma_{4h}^2}{4\gamma_{3h}} - \frac{(\gamma_{5h}\gamma_{3h} - \gamma_{4h}\gamma_{6h})^2}{(\gamma_{2h}\gamma_{3h} - \gamma_{6h}^2)} - \frac{1}{4}R_h^2(A_h - C_h) + R_h(E_h - D_h)\right] < 0 \quad (46)$$

6. Empirical Study

A numerical study is carried out in order to compare the performance of the proposed generalized class of estimators to other existing estimators of the finite population mean. The effects of measurement errors and the three-stage optional RRT model on estimation of the finite population mean are also investigated. Simulated data, Covid-19 global pandemic (www.worldometer.info), and Rosner (2015) data are used in the empirical study. For data simulation and coding, the R programming language is used. Each population unit is subjected to measurement errors, which are normally distributed with mean 2 and variance 5. The efficiency of the proposed estimators is compared to that of other estimators based on the least variance and PRE methods. The PREs of estimators of the finite population mean are calculated using the expression;

$$PRE(t_j) = \frac{Var(t_0)}{MSE(t_j)} \times 100 \quad (47)$$

An estimator with the highest PRE in comparison to the ordinary mean estimator is considered to be more efficient than other estimators. Additionally, the mean square error (MSE) and percent relative efficiency (PRE) are calculated at different sensitivity levels of the survey question. The coefficient of correlation between the sensitive survey variable and auxiliary variable is positive. Furthermore, there is negative correlation between auxiliary attribute, sensitive and auxiliary variable in all the three data sets. The following is a description of the data that is used:

Population I: Simulated data

Stratum 1

$$X_1 = rnorm(100, 450, 15)$$
$$Y_1 = X_1 + rnorm(100, 0, 1)$$
$$X_1 = X_1 + rnorm(100, 2, 5)$$
$$X_1 = rnorm(100, 0, 2)$$
$$Z_1 = Y_1 + S_1$$
$$z_1 = Y_1 + rnorm(100, 2, 5)$$

Auxiliary attributes are values of $Y_1 < mean(Y_1)$

Stratum 2

$$X_2 = rnorm(250, 50, 15)$$
$$Y_2 = X_2 + rnorm(250, 0, 1)$$
$$X_2 = X_2 + rnorm(100, 2, 5)$$
$$X_2 = rnorm(250, 0, 2)$$
$$Z_2 = Y_2 + S_2$$
$$z_2 = Y_2 + rnorm(250, 2, 5)$$

Auxiliary attributes are values of $Y_2 < mean(Y_2)$

Stratum 3

$$X_3 = rnorm(300, 920, 24)$$
$$Y_3 = X_3 + rnorm(300, 0, 1)$$
$$X_3 = X_3 + rnorm(300, 2, 5)$$
$$X_3 = rnorm(300, 0, 2)$$
$$Z_3 = Y_2 + S_2$$
$$z_3 = Y_3 + rnorm(300, 2, 5)$$

Auxiliary attributes are values of $Y_3 < mean(Y_3)$

Stratum 4

$$X_4 = rnorm(350, 500, 8)$$
$$Y_4 = X_4 + rnorm(350, 0, 1)$$
$$X_4 = X_4 + rnorm(350, 2, 5)$$
$$X_4 = rnorm(350, 0, 2)$$

$$Z_4 = Y_4 + S_4$$
$$z_4 = Y_4 + rnorm(350, 2, 5)$$
Auxiliary attributes are values of $Y_4 < mean(Y_4)$

Population II: Covid-19 global pandemic data
The data set is for Covid-19 global pandemic (www.worldometer.info) for the period of January 3rd, 2020 to September 17th, 2021. The data is classified into 6 strata according to World Health organisation (WHO) regions; African Region ($N_1 = 31200$), the American region ($N_2 = 34944$), the Eastern Mediterranean Region ($N_3 = 13728$) the European Region ($N_4 = 38688$), the South-East Asia Region ($N_5 = 6864$), and the Western Pacific Region ($N_6 = 2184$). The number of new cases and deaths in a given day are regarded as the auxiliary and survey variables, respectively. The number of new deaths with a value less than ten is regarded as an auxiliary attribute. For each unit in the data set, a scrambling response with mean 0 and variance 2 was generated and used to calculate the response variable.

Population III: Rosner (2015)
The population consist of two strata of sizes; ($N_1 = 480$) and ($N_2 = 174$) with Y as forced expiratory volume, X as age (in years), and auxiliary attribute as the values of forced expiratory volume less than $\bar{Y}_h$. Furthermore, Smoking (Yes=1, No=0) is taken to be the scrambling variable and is used in generation of response variable

Table 2: Population statistics for the sensitive variable, auxiliary attribute and variable

		$\bar{X}_h$	$\bar{Z}_h$	S^2_{Xh}	S^2_{Zh}	S^2_{Ph}	ρ_{XZh}	ρ_{XPh}	ρ_{ZPh}	S^2_{Th}	S^2_{Vh}
	1	450.2457	450.5318	227.9771	228.2253	0.2516162	0.9914689	-0.7774522	-0.7745486	28.0513	23.65488
	2	49.75684	49.80721	189.9124	190.5358	0.250988	0.9868945	-0.7948887	-0.7876593	20.89942	26.79222
I	3	919.5245	919.5174	558.2454	559.4888	0.250825	0.9957027	-0.8061185	-0.8078304	23.42054	27.01627
	4	500.4659	250.9384	63.43051	18.27174	0.2388948	0.8660414	-0.7593847	-0.7101635	25.25684	27.85556
	1	188.9035	4.543181	1094471	926.4621	0.0631675	0.8171398	-0.4743876	-0.4700577	24.72743	24.8088
	2	2502.012	61.90972	187408859	76639.99	0.1684866	0.7944946	-0.3406924	-0.4228761	24.91892	25.14334
I	3	1120.151	20.51225	8526375	2937.237	0.2190176	0.834325	-0.4254019	-0.4941943	25.03186	25.10566
	4	1757.061	33.79095	24712119	11588.58	0.7043786	0.6559524	-0.4738365	-0.4650428	25.29474	24.9028
	5	6175.008	97.12205	817189958	145353	0.2231698	0.8679977	-0.2992682	-0.3562166	24.74669	25.73482
	6	356.2095	4.833472	3189400	850.8079	0.06043503	0.7237861	-0.6026504	-0.5888455	18.97865	24.87666
	1	8.558333	2.363715	3.604106	0.5254207	0.250087	0.7239923	-0.6367863	-0.7829324	26.04856	22.12586
II	2	13.71839	3.763615	3.301741	0.7556429	0.2487542	0.3619965	-0.203882	-0.706301	20.19661	22.05487

6.1 Results and Discussion

Table 3 shows the values of PREs for population I at different sensitivity levels of the survey question. The proposed generalized class of estimators perform better than other existing mean estimator in both cases for without and with measurement errors. The values of PREs declines as the value of the sensitivity levels of survey question increases for all the estimators with exception of the ratio estimator. Additionally, the values of PREs are high in the absence of measurement errors but decreases when measurement errors are introduced in the survey.

Table 3: PREs of different estimators using the three-stage optional RRT model when $t_h = 0.3$ and $f_h = 0.2$ for population I

Estimator	$\psi_h = 0$		$\psi_h = 0.2$		$\psi_h = 0.8$		$\psi_h = 1.0$	
	without ME	with ME	without ME	with ME	without ME	with ME	without ME	with ME
t_0	100	100	100	100	100	100	100	100
t_R	138.1604	138.1412	138.2786	138.2180	138.5062	138.4569	138.6734	138.6521
t_{ER}	155.1993	155.2676	155.07100	155.2335	154.6867	155.1298	154.5588	155.0948
t_1	167.4635	158.827	167.3225	158.7900	166.9008	158.6784	166.7607	158.6410
t_2	178.3399	167.2935	178.1837	167.2523	177.7169	167.1279	177.5619	167.0863
t_3	176.4973	165.8899	176.3414	165.8476	175.8758	165.7202	175.7213	165.6777
t_4	137.9604	135.1127	137.8586	135.0886	137.5553	135.0167	137.4549	134.9929
t_5	178.4015	167.3417	178.2453	167.3004	177.7784	167.1761	177.6234	167.1345
t_6	175.4756	165.1065	175.3205	165.064	174.8571	164.9364	174.7034	164.8937
t_7	175.5834	165.1911	175.4282	165.1486	174.9646	165.0209	174.8108	164.9783
t_8	178.1468	167.1437	177.9907	167.1024	177.5243	166.978	177.3695	166.9363
t_9	176.2024	165.6364	176.0481	165.5952	175.5871	165.4710	175.4341	165.4295
t_{10}	178.1546	167.1498	177.9985	167.1085	177.5322	166.9841	177.3773	166.9425
t_{11}	176.2278	165.6563	176.0736	165.6151	175.6127	165.491	175.4596	165.4495

Table 4 shows PREs for population II at different values of ψ_h. The values of PREs are observed to decrease with increase in the sensitivity level of the survey question in both cases for without and with measurement errors. However, PREs for the ratio estimator are observed to increase with increase in the sensitivity level of the survey question. For example, in the case for with measurement errors, the value of PRE for t_5 is 152.0858 when $\psi_h = 0$ and decreases to 152.0848 when $\psi_h = 0.8$. The proposed estimators perform better than other existing estimators of the finite population mean in both cases for without and with measurement errors. Generally, the values of PREs for the proposed estimators declines in the presence of measurement errors.

Table 4: PREs of different estimators using the three-stage optional RRT model when $t_h = 0.3$ and $f_h = 0.2$ for population II

Estimator	$\psi_h = 0$		$\psi_h = 0.2$		$\psi_h = 0.8$		$\psi_h = 1.0$	
	without ME	with ME	without ME	with ME	without ME	with ME	without ME	with ME
t_0	100	100	100	100	100	100	100	100
t_R	137.7421	137.6962	137.7424	137.6965	137.7435	137.6976	137.7439	137.6980
t_{ER}	149.0357	148.9713	149.0353	148.9709	149.0342	148.9698	149.0338	148.9694
t_1	152.0858	152.0160	152.0856	152.0157	152.0848	152.015	152.0846	152.0148

t_2	152.0617	151.9918	152.0614	151.9916	152.0607	151.9909	152.0604	151.9906
t_3	152.085	152.0151	152.0847	152.0149	152.0840	152.0142	152.0838	152.0139
t_4	137.3917	137.3464	137.3916	137.3462	137.39110	137.3458	137.3909	137.3456
t_5	152.0858	152.0160	152.0856	152.0157	152.0848	152.0150	152.0846	152.0148
t_6	152.0831	152.0132	152.0828	152.0130	152.0821	152.0123	152.0819	152.0120
t_7	149.7169	149.6512	149.7166	149.651	149.7158	149.6502	149.7155	149.6499
t_8	152.0847	152.0149	152.0845	152.0146	152.0838	152.0139	152.0835	152.0137
t_9	150.6822	150.6149	150.6820	150.6146	150.6812	150.6139	150.6810	150.6137
t_{10}	152.0820	152.0121	152.0817	152.0119	152.0810	152.0112	152.0808	152.0109
t_{11}	150.9761	150.9082	150.9758	150.908	150.9751	150.9072	150.9748	150.9070

Table 5 shows the PREs for population III at different values of sensitivity levels of the survey question. The ratio and exponential ratio-type estimators' underperforms compared to ordinary mean estimator in the presence of measurement errors. However, the ratio estimator performs better than exponential ratio estimator in the case of without measurement errors and vice-versa in the presence of measurement errors. The proposed generalized class of estimators perform better than the ratio and exponential ratio estimators in both cases for without and with measurement errors at different values of ψ_h. The values of PREs decrease with an increase in sensitivity level in the case for without measurement errors. For example, for t_5 the value of PRE is 117.8663 when $\psi_h = 0.2$ and decreases to 117.2759 when and $\psi_h = 0.8$. Furthermore, the values of PRE increases with increase in the values of ψ_h in the presence of measurement errors.

Table 5: PREs of different estimators using the three-stage RRT model when $t_h = 0.3$ and $f_h = 0.2$ for population III

Estimator	$\psi_h = 0$		$\psi_h = 0.2$		$\psi_h = 0.8$		$\psi_h = 1.0$	
	without ME	with ME	without ME	with ME	without ME	with ME	without ME	with ME
t_0	100	100	100	100	100	100	100	100
t_R	121.1715	96.88974	121.0878	96.89571	120.8350	96.91333	120.7506	96.91912
t_{ER}	117.9163	99.42842	117.723	99.43119	117.1757	99.43935	117.0033	99.44202
t_1	118.0769	100.2539	117.8663	100.2553	117.2759	100.2595	117.0916	100.2608
t_2	117.7667	100.2492	117.5608	100.2506	116.9834	100.2547	116.8032	100.256
t_3	114.4735	100.1976	114.3156	100.1987	113.8722	100.2020	113.7337	100.2030
t_4	103.7375	100.0071	103.7210	100.0072	103.6748	100.0075	103.6604	100.0076
t_5	118.0769	100.2539	117.8663	100.2553	117.2759	100.2595	117.0916	100.2608
t_6	114.5483	100.1988	114.3876	100.1999	113.9364	100.2031	113.7955	100.2041
t_7	115.2415	100.2099	115.0694	100.211	114.5869	100.2143	114.4364	100.2154
t_8	116.536	100.2303	116.3481	100.2316	115.8207	100.2353	115.6561	100.2365

t_9	113.0928	100.1751	112.9505	100.1760	112.5512	100.1787	112.4266	100.1796
t_{10}	117.1218	100.2394	116.9258	100.2407	116.3757	100.2446	116.2040	100.2459
t_{11}	114.3631	100.1959	114.2061	100.1969	113.7654	100.2001	113.6277	100.2011

7. Concluding Remarks

In this paper, a generalized class of estimators is proposed using the three stage optional RRT model under measurement errors. The proposed estimators are based on a mixture of auxiliary attribute and variable. Using both simulated and real data, the theoretical properties of biases and MSEs for the proposed estimators are investigated theoretically and numerically. According to the numerical analysis, efficiencies of estimators of the finite population mean decreases as the sensitivity level of the survey question increases in both cases for without and with measurement errors. Furthermore, the use of the three stage optional RRT model reduces impact of the sensitivity level of the survey question on efficiencies of estimators of population mean. The proposed estimator outperforms the ordinary, usual ratio, and exponential ratio-type estimators. As a result, survey practitioners are encouraged to use the proposed estimators when measurement errors are present in a survey.

Acknowledgement
The authors are grateful to the anonymous referees for their helpful comments that improved this paper.

References

1. Chaudhuri, A. & Mukherjee, R. (1988). Randomized response: Theory and Techniques. Marcel Dekker
2. Diwakar, S., Sharad P. & Narendra, S. T. (2012). An Estimator for Mean Estimation in Presence of measurement error. Research and Reviews: A Journal of Statistics, 1, 1-8.
3. Gupta, S.N., Thornton, B., Shabbir, J. & Singhal, S. (2006). A comparison of multiplicative and additive optional RRT models. Journal of Statistical theory and application, 5(3), 226-239.
4. Gupta, S., Shabbir, J. & Sehra S. (2012). Mean and sensitivity estimation in optional randomized response models. Journal of Statistical Planning and Inference, 140(10), 2870-2874
5. Gupta, S. N., and Shabbir, J. (2007). Mean and Sensitivity in a Two-Stage Optional Randomized Response Model on the Estimation of Population. Journal of Indian Society of Agricultural Statistics, 61(2), 164-168
6. Hodges, J. L & Lehmann, E. L. (1963). Estimates of location based on rank tests. Annals of Mathematical Statistics, 34(2), 598-611.
7. Ibrahim, M. A., Amer, I. A. & Emmanuel, J. E. M. S. (2021). Generalized class of mean estimators with known measures outliers' treatment. Computer Systems Science Engineering, 37. DOI:10.32604/csse.2021.015933
8. Khalil, S., Zhang, Q., Gupta, S. (2019). Mean Estimation of Sensitive Variables under Measurement Errors using Optional RRT Models. Communications in Statistics - Simulation and Computation. DOI:10.1080/03610918.2019.1584298
9. Khalil, S., Noor-Ul-Amin, M., and Hanif, M. (2018). Estimation of Population Mean for a Sensitive Variable in the Presence of Measurement Error. Journal of Statistics and Management Systems, 21(1), 81-91.
10. Kumar, M., Singh, R., Singh, K. & Smarandache, F. (2011). Some ratio type estimators under measurement errors. World Applied Science journal, 14, 272-276.
11. Mehta, S., Dass, B. K., Shabbir, J. & Gupta. (2012). A three- stage optional randomized response model},
12. Journal of Statistical Theory and Practice, 6(3), 412-427.
13. Mushtaq, N. & Noor-ul-Amin, M. (2020). Joint influence of double sampling and randomized response technique on estimation method of mean. Applied Mathematics, 10(1), 12-19.
14. Naeem, N. & Shabbir, J. (2018). Use of a scrambled response on two occasion's successive sampling under nonresponse. Hacettepe Journal of Mathematics and Statistics, 47(3), 675-684
15. Neeraj, T. & Prachi, M. (2017). Additive Randomized Response Model with Known Sensitivity Level. International Journal of Computational and Theoretical Statistics, 4(2), 675-684.
16. Onyango, R., Oduor, B. and Odundo, F. (2021). Joint influence of measurement errors and randomized response technique on mean estimation under stratified double sampling. Open Journal of Mathematical Science, 5, 192-199.

17. Pollock, K. & Bek, Y. (1976). A comparison of three randomized response models for quantitative data. Journal of the American Statistical Association, 71(356), 884-886
18. Rosner, B. (2015). Fundamentals of biostatistics. Duxbury Press.
19. Shabbir, J., Shakeel, A., Aamir, S. & Onyango, R. (2021). Measuring Performance of Ratio-Exponential-Log Type General Class of Estimators Using Two Auxiliary Variables. Mathematical Problems in Engineering. https://doi.org/10.1155/2021/5245621.
20. Shalabh, J. (1997). Ratio method of estimation in the presence of measurement errors. Journal of Indian Society of Agricultural Statistics, 50(2), 150-155.
21. Shukla, D., Pathak, S., & Thakur, N. (2012). An estimator for mean estimation in presence of measurement error. Research and Reviews: A Journal of Statistics, 1(1), 1-8.
22. Singh, H. P. & Karpe, N. (2012). Effect of measurement errors on the separate and combined ratio and product estimators in stratified random sampling, Journal of Modern Applied Statistical Methods, 2012}, 9, DOI: 10.22237/jmasm/1288584420.
23. Subzar, M., Maqbool, S., Raja, T. A & Bhat, M. A. (2018). Estimation of finite population mean in stratified random sampling using nonconventional measures of dispersion. Journal of Reliability and Statistical Studies, 11(1), 83-92.
24. Tukey, J. W. (1970). Exploratory Data Analysis. Addison-Welsey Publishing Co., Reading, MA, USA.
25. Vishwakarma, G., Abhishek, S. & Neha, S. (2020). Calibration under measurement errors. Journal of King Saud University Science, 32, 2950-2961.
26. Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60(309), 63-69, doi = 10.1080/01621459.1965.
27. Yadav, D., Sheela, M. & Dipika. (2017). Estimation of population mean using auxiliary information in presence of measurement errors. International Journal of Engineering Sciences and Research Technology, 6(6), DOI: 10.5281/zenodo.817860.
28. Zahid, E. & Shabbir, J. (2019). Estimation of finite population mean for a sensitive variable using dual auxiliary information in the presence of measurement errors. PloS one, 14: e0212111.