Statistics and Applications {ISSN 2452-7395 (online)}
Volume 15 Nos. 1&2, 2017 (New Series), pp 63-78

Improved Ratio and Regression Estimators of the Mean of a Sensitive Variable in Stratified Sampling

Geeta Kalucha[1], Rita Sousa[2], Sat Gupta[3] and Javid Shabbir[4]

[1]*Department of Mathematics, University of Delhi, India*
[2]*New University of Lisbon and Banco de Portugal*
[3]*Department of Mathematics and Statistics, The University of North Carolina at Greensboro, North Carolina, USA*
[4]*Department of Statistics, Quaid-I-Azam University, Islamabad, Pakistan*

Received: September 23, 2016; Revised February 07, 2017; Accepted: February 10, 2017

Abstract

Gupta et al. (2014) introduced ratio and regression estimators for the mean of a sensitive variable using optional additive RRT models which perform better than the Sousa et al. (2010) and Gupta et al. (2012) ratio and regression estimators that were based on non-optional additive RRT model. In the present study we extend Gupta et al. (2014) estimators to the stratified sampling setting and compare them with the existing non optional estimators in the stratified sampling setting proposed by Sousa et al. (2014). The performance of the proposed estimators is also compared with the corresponding estimators in simple random sampling.

Key words: Mean square error; Optional randomized response technique; Combined ratio estimator; Combined regression estimator; Stratified random sampling.

1. Introduction and Terminology

The main goal of this paper is to extend the ratio and regression mean estimator results of Gupta et al. (2014) to the case of stratified sampling. It is assumed that the study variable is sensitive and a non-sensitive auxiliary variable is available which is positively correlated with the study variable.

Many authors have presented traditional ratio and regression estimators for the population mean in simple random sampling when both the study variable Y and the auxiliary variable X are directly observable. These include Ray and Singh (1981), Kadilar and Cingi (2004, 2005), Kadilar et al. (2007), Shabbir and Gupta (2007, 2010) and Nangsue (2009). Gupta and Shabbir (2008) have suggested a general class of ratio estimators when the population parameters of the auxiliary variable are known. Kadilar and Cingi (2003, 2005), Singh and Vishwakarma (2008), Koyuncu and Kadilar (2008,2009, 2010) have proposed a family of combined-type estimators in stratified random sampling. These estimators are identified as members of the recently proposed class of estimators by Singh and Solanki (2013).Some studies on estimation of the mean have been done with other sampling schemes such as Singh and Solanki (2012) for a systematic sampling design and Singh and Vishwakarma (2007) in double sampling.

Corresponding author: Geeta Kalucha
E-mail: *geetakalucha@gmail.com*

Gupta et al. (2014) suggested ratio and regression estimators for the sensitive variable *Y* using a non-sensitive variable *X* improving the estimators of Sousa et al. (2010) and Gupta et al. (2012) in simple random sampling without replacement *(SRSWOR).* The improvement was seen as a result of using an optional additive RRT model (introduced in Gupta et al. (2002)) as compared to the non-optional additive RRT model used by Sousa et al. (2010) and Gupta et al. (2012). The introduction of optionality led to the estimation of *W* (the sensitivity level) along with the estimation of population mean. It may be noted that the sensitivity level *W* is the proportion of respondents in the population who consider the question sensitive enough to not feel comfortable answering the question in a face-to-face survey. Recently Sousa et al. (2014) extended the estimators of Sousa et al. (2010) and Gupta et al. (2012) to the stratified sampling setting. Motivated by Sousa et al. (2014), we extend the estimators of Gupta et al. (2014) to the stratified sampling setting.

This paper suggests a combined ratio estimator and a combined regression estimator for the population mean of a sensitive variable using non-sensitive auxiliary variable and an optional RRT methodology in stratified sampling. The *Bias* and the *Mean Square Error (MSE)* of the suggested estimators are derived and they are compared theoretically and empirically with the non-optional combined ratio and combined regression estimators of Sousa et al. (2014). It is shown that among the proposed estimators the combined regression estimator is always most efficient.

We denote the finite population by $U=\{U_1,U_2,...,U_N\}$. The study population is divided into *L* strata with strata sizes N_h such that $\sum_{h=1}^{L} N_h = N \ (h=1,...,L)$. Let *Y* be the sensitive study variable which cannot be observed directly. Let *X* be a non-sensitive auxiliary variable which is positively correlated with *Y*. Let T be a scrambling random variable independent of *Y* and *X*. We assume that $\mu_T = E(T)=0$. Let *W* be the sensitivity level of the underlying sensitive question. Each respondent in the sample is asked to report an additively scrambled response for *Y* if he/she considers the question sensitive and a true response otherwise. Thus a scrambled response on *Y* is received with probability *W* and a true response is received with probability (1-*W*). The respondent always provides a correct response for the auxiliary variable *X*. The reported response *Z* for the study variable can thus be written

$$Z=\begin{cases} Y+T & \text{, with probability } W \\ Y & \text{, with probability } (1-W) \end{cases}$$

We draw a sample of size n_h from each stratum by using simple random sampling without replacement (SRSWOR) such that $\sum_{h=1}^{L} n_h = n$. Let y_{hi} and x_{hi} respectively be the values of the i-th study variable *Y* and the auxiliary variable *X* in the h^{th} stratum with $i=1,2,...,n_h$. Let $\bar{y}_{st}=\sum_{h=1}^{L}\delta_h\bar{y}_h$, $\bar{x}_{st}=\sum_{h=1}^{L}\delta_h\bar{x}_h$ and $\bar{z}_{st}=\sum_{h=1}^{L}\delta_h\bar{z}_h$ be the stratified sample means, where $\bar{y}_h=\frac{1}{n_h}\sum_{i=1}^{n_h}y_{hi}$, $\bar{x}_h=\frac{1}{n_h}\sum_{i=1}^{n_h}x_{hi}$ and $\bar{z}_h=\frac{1}{n_h}\sum_{i=1}^{n_h}z_{hi}$ are the stratum sample means corresponding to population stratum means $\bar{Y}_h=\mu_{Y_h}=E(Y_h)$, $\bar{X}_h=\mu_{X_h}=E(X_h)$, $\bar{Z}_h=\mu_{Z_h}=E(Z_h)$, and $\delta_h = N_h/N$ are the known stratum weights.

To estimate $\overline{Y} = \mu_Y = \sum_{h=1}^{N_h} \delta_h \mu_{Y_h}$ we assume that $\overline{X} = \mu_X = \sum_{h=1}^{N_h} \delta_h \mu_{X_h}$ is known. Let $\overline{Z} = \mu_Z = \sum_{h=1}^{N_h} \delta_h \mu_{Z_h}$ be the mean for the scrambled variable Z.

To discuss the properties of different estimators, we define the following error terms. Let $e_{0st} = \frac{\bar{z}_{st} - \mu_Z}{\mu_Z}$, $e_{1st} = \frac{\bar{x}_{st} - \mu_X}{\mu_X}$, $e_{2st} = \frac{s_{xst}^2 - S_{xst}^2}{S_{xst}^2}$ and $e_{3st} = \frac{s_{zxst} - S_{zxst}}{S_{zxst}}$ such that $E(e_{ist}) = 0$ $(i = 0,1,2,3)$.

Below we list some existing mean estimators in the case of simple random sample.

(i) Gupta et al. (2014) Mean and Sensitivity Estimators:

$$\hat{\mu}_{YW} = \bar{z} \tag{1.1}$$

$$\hat{W} = \frac{\frac{1}{n}\sum_{i=1}^{n} z_i^2 - \left\{\frac{1}{n}\sum_{i=1}^{n} z_i + \left(\frac{1}{n}\sum_{i=1}^{n} z_i\right)^2\right\}}{\delta}, \text{ when } Y \text{ has Poisson distribution} \tag{1.2}$$

and
$$\hat{W} = \frac{\hat{\sigma}_z^2 - (C_x \bar{z})^2}{E(T^2)}, \text{ when } C_x = C_y \tag{1.3}$$

$$MSE(\hat{\mu}_{YW}) = Var(\hat{\mu}_{YW}) = \left(\frac{1-f}{n}\right)\left(S_y^2 + WS_T^2\right) \tag{1.4}$$

where $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_Y)^2$ and. $S_T^2 = \frac{1}{N-1}\sum_{i=1}^{N}(s_i - \mu_S)^2$

(ii) Gupta et al. (2014) Ratio Estimator:

$$\hat{\mu}_{RW} = \bar{z}\left(\frac{\mu_X}{\bar{x}}\right) \tag{1.5}$$

The Bias and *MSE* of $\hat{\mu}_{RW}$ to first degree of approximation are given by

$$Bias(\hat{\mu}_{RW}) \cong \frac{1-f}{n}\mu_Y\left(C_x^2 - \rho_{zx}C_zC_x\right) \tag{1.6}$$

and
$$MSE(\hat{\mu}_{RW}) \cong \frac{1-f}{n}\mu_Y^2\left(C_z^2 + C_x^2 - 2\rho_{zx}C_zC_x\right) \tag{1.7}$$

where $C_z^2 = C_y^2 + W\frac{S_T^2}{\mu_Y^2}$, $\rho_{zx} = \frac{\rho_{yx}}{\sqrt{1 + W\frac{S_T^2}{S_y^2}}}$ and C_z, C_y and C_x are the coefficients of variation of *Z*, *Y* and *X* respectively and $0 < W < 1$.

(iii) Gupta et al. (2014) Regression Estimator:

$$\hat{\mu}_{\mathrm{Reg}W} = \bar{z} + \hat{\beta}_{zx}(\mu_X - \bar{x}) \quad (1.8)$$

where $\hat{\beta}_{zx}$ is the sample regression coefficient between Z and X.

The Bias and *MSE* of $\hat{\mu}_{\mathrm{Reg}W}$ to first degree of approximation, are given by

$$Bias(\hat{\mu}_{\mathrm{Reg}W}) \cong -\beta_{zx}\left(\frac{1-f}{n}\right)\left\{\frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}}\right\} \quad (1.9)$$

and $$MSE(\hat{\mu}_{\mathrm{Reg}W}) \cong \left(\frac{1-f}{n}\right)\mu_Y^2 C_z^2 (1-\rho_{zx})^2 = \left(\frac{1-f}{n}\right) S_y^2 \left\{\left(1 + W\frac{S_T^2}{S_y^2}\right) - \rho_{yx}^2\right\} \quad (1.10)$$

where $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(z_i - \mu_Z)^r (x_i - \mu_X)^s$

For a stratified random sample the usual combined sample mean estimator, ignoring the auxiliary information, is given by

$$\hat{\mu}_{YWst} = \bar{z}_{st}, \quad (1.11)$$

which is the unbiased estimator of population mean μ_Y.

The *MSE* of $\hat{\mu}_{YWst}$ is given by

$$MSE(\hat{\mu}_{YWst}) = \sum_{h=1}^{L} \delta_h^2 \gamma_h \left(S_{yh}^2 + W S_{Th}^2\right) \quad (1.12)$$

where $\gamma_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$, $S_{yh}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}\left(y_{hi} - \mu_{Y_h}\right)^2$ and $S_{Th}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}\left(t_{hi} - \mu_{T_h}\right)^2$

and $0 < W < 1$.

It is interesting to note that for $W = 1$, we have

$$MSE(\hat{\mu}_{YWst}) = MSE(\hat{\mu}_{Yst}) \quad (1.13)$$

where $\hat{\mu}_{Yst} = \bar{z}_{st}$, the combined sample mean for a stratified random sample of non-optional additively scrambled responses, is the unbiased estimator of population mean μ_Y as given by Sousa et al. (2014).

Also we observe from (1.4) and (1.12), that

$MSE(\hat{\mu}_{YWst}) < MSE(\hat{\mu}_{YW})$ if

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h \left(S_{yh}^2 + W S_{Th}^2\right) < \left(\frac{1-f}{n}\right)\left(S_y^2 + W S_T^2\right) \quad (1.14)$$

a condition that can be ensured by a suitable stratification for all values of $W: 0 < W < 1$.

In Section 2, we now introduce a combined ratio estimator and compare it to the ordinary mean estimator and the ratio estimator (Gupta et al., 2014), and the combined ratio estimator of Sousa et al. (2014). In Section 3, we propose a combined regression estimator and compare it with the regression estimator proposed by Gupta et al. (2014) and the combined regression estimator proposed by Sousa et al. (2014). In Section 4, we present a comparative simulation study.

2. Proposed Combined Ratio Estimator

We propose the following combined ratio estimator

$$\hat{\mu}_{RWst} = \bar{z}_{st}\left(\frac{\bar{X}}{\bar{x}_{st}}\right) \tag{2.1}$$

Using Taylor's approximation and retaining terms of order up to 2, (2.1) can be rewritten as

$$\hat{\mu}_{RWst} - \mu_Z \cong \mu_Z \left\{e_{0st} - e_{1st} + e_{1st}^2 - e_{0st}\, e_{1st}\right\} \tag{2.2}$$

Under the assumption of bivariate normality (see Sukhatme and Sukhatme, 1970), we can get the expressions for the *Bias* and *MSE* for $\hat{\mu}_{RWst}$, correct to first order of approximation, as given by

$$Bias(\hat{\mu}_{RWst}) \cong \mu_Y \sum_{h=1}^{L} \delta_h^2 \gamma_h \left\{C_{xh}^2 - C_{zxh}\right\} \tag{2.3}$$

and $$MSE(\hat{\mu}_{RWst}) \cong \mu_Y^2 \sum_{h=1}^{L} \delta_h^2 \gamma_h \left(C_{zh}^2 + C_{xh}^2 - 2C_{zxh}\right) \tag{2.4}$$

where $C_{zxh} = \rho_{zxh} C_{zh} C_{xh}$, $C_{zh}^2 = C_{yh}^2 + W\dfrac{S_{Th}^2}{\mu_Y^2}$

and $\rho_{zxh} = \dfrac{\rho_{yxh}}{\sqrt{1 + W\dfrac{S_{Th}^2}{s_{yh}^2}}}$ and $0 < W < 1$.

The *Bias* and *MSE* of the non-optional combined ratio estimator $\hat{\mu}_{Rst}$, as proposed by *Sousa et al. (2014)*, correct to first order of approximation, is given by

$$Bias(\hat{\mu}_{Rst}) \cong \mu_Y \sum_{h=1}^{L} \delta_h^2 \gamma_h \left\{C_{xh}^2 - C_{zxh}\right\} \tag{2.5}$$

and $$MSE(\hat{\mu}_{Rst}) \cong \mu_Y^2 \sum_{h=1}^{L} W_h^2 \gamma_h \left(C_{zh}^2 + C_{xh}^2 - 2C_{zxh}\right) \tag{2.6}$$

where $C_{zxh} = \rho_{zxh} C_{zh} C_{xh}$, $C_{zh}^2 = C_{yh}^2 + \dfrac{S_{Th}^2}{\mu_Y^2}$ and. $\rho_{zxh} = \dfrac{\rho_{yxh}}{\sqrt{1 + \dfrac{S_{Th}^2}{S_{yh}^2}}}$.

As expected, our results for the *Bias* and *MSE* coincide with those of Sousa et al. (2014) for $W = 1$ and hence for $W = 1$,

$$MSE(\hat{\mu}_{RWst}) = MSE(\hat{\mu}_{Rst}) \tag{2.7}$$

Below we compare the proposed combined ratio estimator $(\hat{\mu}_{RWst})$ with the ordinary sample mean estimator $(\hat{\mu}_{YWst})$, Sousa et al. (2014) combined ratio estimator $(\hat{\mu}_{Rst})$ and the Gupta et al. (2014) ratio estimator $(\hat{\mu}_{RW})$.

It can be verified easily that

(i) From equations (2.4) and (1.12),
$MSE(\hat{\mu}_{RWst}) < MSE(\hat{\mu}_{YWst})$,
if

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zxh} - \frac{1}{2} \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{xh}^2 > 0$$

$$\Rightarrow \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zh} C_{xh} \left(\rho_{zxh} - \frac{1}{2} \frac{C_{xh}}{C_{zh}} \right) > 0 \Rightarrow \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{yh} C_{xh} \left(\rho_{yxh} - \frac{1}{2} \frac{C_{xh}}{C_{yh}} \right) > 0$$

Hence we can conclude

$$MSE(\hat{\mu}_{RWst}) < MSE(\hat{\mu}_{YWst}) \text{ if} \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{yh} C_{xh} \left(\rho_{yxh} - \frac{1}{2} \frac{C_{xh}}{C_{yh}} \right) > 0. \tag{2.8}$$

(ii) From equations (2.4) and (2.6), $MSE(\hat{\mu}_{RWst}) < MSE(\hat{\mu}_{Rst})$ if

$$(W-1) \sum_{h=1}^{L} \delta_h^2 \gamma_h \frac{S_{Th}^2}{\mu_Y^2} < 0 \tag{2.9}$$

which holds true if $(W-1) < 0$, a condition which always holds true as $0 < W < 1$.

(iii) From equations (2.4) and (1.7), $MSE(\hat{\mu}_{RWst}) < MSE(\hat{\mu}_{RW})$ if,

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h \left\{ C_{zh}^2 + C_{xh}^2 - 2\rho_{zxh} C_{zh} C_{xh} \right\} < \frac{1-f}{n} \left\{ C_z^2 + C_x^2 - 2\rho_{zx} C_z C_x \right\} \tag{2.10}$$

a condition that can be ensured by a suitable stratification, for all $0 < W < 1$.

3. Proposed Combined Regression Estimator

We propose the following combined regression estimator for the population mean μ_Y

$$\hat{\mu}_{\mathrm{Re}gWst} = \bar{z}_{st} + \hat{\beta}_c \left(\bar{X} - \bar{x}_{st} \right), \tag{3.1}$$

where $\hat{\beta}_c = \dfrac{\sum_{h=1}^{L} \delta_h^2 \gamma_h s_{zxh}}{\sum_{h=1}^{L} \delta_h^2 \gamma_h s_{xh}^2}$ is the sample regression coefficient between *Z* and *X*, and

$Z = \begin{cases} Y+T, & \text{with probability } W, \\ Y, & \text{with probability } (1-W) \end{cases}$ is the optional additive scrambled response on *Y*.

Using Taylor's approximation and retaining terms of order up to 2, (3.1) can be rewritten as

$$\hat{\mu}_{\mathrm{Re}gWst} - \mu_Z \cong \mu_Z e_{0st} - \beta_c \mu_X \left[e_{1st} + e_{1st} e_{3st} - e_{1st} e_{2st} \right], \tag{3.2}$$

where $\beta_c = \dfrac{\sum_{h=1}^{L} \delta_h^2 \gamma_h s_{zxh}}{\sum_{h=1}^{L} \delta_h^2 \gamma_h s_{xh}^2}$ is the population regression coefficient between *Z* and *X*.

Considering a simple random sample selected from each population stratum we can deduce, using Mukhopadhyay (1998, p.123), as in Sousa et al. (2014), that the *Bias* and *MSE* of $\hat{\mu}_{\mathrm{Re}gWst}$ to first degree of approximation are given by

$$Bias\left(\hat{\mu}_{\mathrm{Re}gWst}\right) \cong -\sum_{h=1}^{L} \delta_h^2 \gamma_h \beta_c \left\{ \frac{\mu_{12h}}{\mu_{11h}} - \frac{\mu_{03h}}{\mu_{02h}} \right\} \tag{3.3}$$

and
$$MSE\left(\hat{\mu}_{\mathrm{Re}gWst}\right) \cong \mu_Y^2 \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zh}^2 \left(1 - \rho_c^2\right) \tag{3.4}$$

where $\rho_c = \dfrac{\sum_{h=1}^{L} \delta_h^2 \gamma_h S_{zxh}}{\sqrt{\sum_{h=1}^{L} \delta_h^2 \gamma_h S_{zh}^2} \sqrt{\sum_{h=1}^{L} \delta_h^2 \gamma_h S_{xh}^2}}$, $\mu_{rsh} = \dfrac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(z_{h_i} - \mu_{Z_h}\right)^r \left(x_{h_i} - \mu_{X_h}\right)^s$

$C_{zh}^2 = C_{yh}^2 + W \dfrac{S_{Th}^2}{\mu_Y^2}$ and $0 < W < 1$.

The *Bias* and *MSE* of the non-optional combined regression estimator $\hat{\mu}_{\mathrm{Re}gst}$, as proposed by *Sousa et al. (2014)*, correct to first order of approximation, is given by

$$Bias\left(\hat{\mu}_{\mathrm{Re}gst}\right) \cong -\sum_{h=1}^{L} \delta_h^2 \gamma_h \beta_c \left\{ \frac{\mu_{12h}}{\mu_{11h}} - \frac{\mu_{03h}}{\mu_{02h}} \right\} \tag{3.5}$$

and
$$MSE\left(\hat{\mu}_{\mathrm{Re}gst}\right) \cong \mu_Y^2 \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zh}^2 \left(1 - \rho_c^2\right). \tag{3.6}$$

We again see that as expected, for $W = 1$, our results for the *Bias* and *MSE* coincide with those of Sousa et al. (2014). Hence for $W = 1$, we again have

$$MSE\left(\hat{\mu}_{\mathrm{Re}gWst}\right) = MSE\left(\hat{\mu}_{\mathrm{Re}gst}\right) . \tag{3.7}$$

We compare below the proposed combined regression estimator $(\hat{\mu}_{\mathrm{Re}gWst})$ with the ordinary sample mean $(\hat{\mu}_{YWst})$, the proposed combined ratio estimator $(\hat{\mu}_{RWst})$, Sousa et al. (2014) combined regression estimator $(\hat{\mu}_{\mathrm{Re}gst})$ and the Gupta et al. (2014) regression estimator $(\hat{\mu}_{\mathrm{Re}gW})$.

It can be easily verified that

(i) From equations (3.4) and (1.12), $MSE(\hat{\mu}_{\mathrm{Re}gWst}) < MSE(\hat{\mu}_{YWst})$ if

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zh}^2 \rho_c^2 > 0 \tag{3.8}$$

(ii) From equations (3.4) and (2.4), $MSE(\hat{\mu}_{\mathrm{Re}gWst}) < MSE(\hat{\mu}_{RWst})$ if

$$\left(\sqrt{\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{xh}^2} - \frac{\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zxh}}{\sqrt{\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{xh}^2}} \right)^2 > 0 \tag{3.9}$$

(iii) From equations (3.4) and (3.6), $MSE(\hat{\mu}_{\mathrm{Re}gWst}) < MSE(\hat{\mu}_{\mathrm{Re}gst})$ if

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h \left\{ S_{yh}^2 + W S_{Th}^2 + \beta_c^{\,2} S_{xh}^2 - 2\beta_c S_{yxh} \right\} - \sum_{h=1}^{L} \delta_h^2 \gamma_h \left\{ S_{yh}^2 + S_{Th}^2 + \beta_c^{\,2} S_{xh}^2 - 2\beta_c S_{yxh} \right\} < 0$$

which amounts to $$(W-1)\sum_{h=1}^{L} \delta_h^2 \gamma_h S_{Th}^2 < 0 \tag{3.10}$$

The above holds true if $(W-1) < 0$, a condition that will always be true.

(iv) From equations (3.4) and (1.10), $MSE(\hat{\mu}_{\mathrm{Re}gWst}) < MSE(\hat{\mu}_{\mathrm{Re}gW})$ if,

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{zh}^2 (1-\rho_c^2) < \frac{1-f}{n} C_z^2 (1-\rho_{zx}^2) \tag{3.11}$$

a condition that can be ensured by a suitable stratification for all values of $W : 0 < W < 1$.

Also conditions (i), (ii) and (iii) will hold true for all values of $W : 0 < W < 1$, indicating that, up to first order of approximation, the regression estimator performs better than ordinary mean and ratio estimators in the optional setting in stratified sampling also, as it did in the case of simple random sampling in Gupta et al. (2014).

4. A Simulation Study

In this section, we present a simulation study with particular focus on comparing the performance of the proposed optional estimators $\hat{\mu}_{RWst}$ and $\hat{\mu}_{\mathrm{Re}gWst}$ to the ordinary mean estimator $\hat{\mu}_{YWst}$, to the corresponding estimators in simple random sampling (Gupta et al., 2014) and to the non-optional estimators in stratified sampling (Sousa et al., 2014). For this comparison we rely on *Bias* and *MSE*, correct up to first order of approximation.

We consider abivariate normal population with aspecified mean vector and covariance matrix to represent the distribution of (Y, X). The scrambling variable $T$is taken to be a normal distribution with mean equal to zero and standard deviation equal to 3. The reported response for the study variable Y is given by $Z = Y+TV$,where V is a binomial random variable with parameters n and $p=W$, where W is the probability of a respondent considering the question sensitive.

We simulate two bivariate populations each of size N = 5000. These populations have theoretical means of $\mu = [6 \;\; 4]$ and covariance matrices given by:

$$\Sigma = \begin{bmatrix} 9 & 4.8 \\ 4.8 & 4 \end{bmatrix} \text{and} \, \Sigma = \begin{bmatrix} 9 & 2.9 \\ 2.9 & 4 \end{bmatrix}.$$

For this choice of covariance matrices we have the corresponding coefficients of correlation: $\rho_{XY} = 0.8$ and $\rho_{XY} = 0.4833$. However, after we select 5000 observations from these populations for the purpose of further simulations, the correlations are 0.8020 and 0.4926 respectively for the first and the second populations,.

The data consist of 5000 observations which are divided into two strata according to the auxiliary variable X. We consider a total sample of size n = 200. The sample size from each stratum is based on the *Neyman* allocation. We present below more detailed information on the two finite populations used in the simulation study.

Bivariate Population I: $\rho_{YX} = .8020$

Stratum 1: $X \in [0,6]$

$N_1 = 4096$, $n_1 = 150$, $\rho_{XY1} = 0.6964$, $S_{YX1} = 2.5253$
$\bar{Y}_1 = 5.4463$, $S_{Y1} = 2.4941$, $C_{Y1} = 0.4579$
$\bar{X}_1 = 3.5678$, $S_{X1} = 1.4540$, $C_{X1} = 0.4075$

Stratum 2: $X > 6$

$N_2 = 904$, $n_2 = 50$, $\rho_{YX2} = 0.8740$, $S_{YX2} = 9.0265$
$\bar{Y}_2 = 8.5167$, $S_{Y2} = 3.7689$, $C_{Y2} = 0.4425$
$\bar{X}_2 = 6.0424$, $S_{X2} = 2.7401$, $C_{X2} = 0.4534$

Bivariate Population II: $\rho_{YX} = .4926$
Stratum 1: $X \in [0,5]$

$N_1 = 3366$, $n_1 = 29$, $\rho_{XY1} = 0.3542$, $S_{YX1} = 1.2349$
$\bar{Y}_1 = 5.3508$, $S_{Y1} = 2.7954$, $C_{Y1} = 0.5224$
$\bar{X}_1 = 3.1025$, $S_{X1} = 1.2472$, $C_{X1} = 0.4020$

Stratum 2: $X > 5$
$N_2 = 1634$, $n_2 = 67$, $\rho_{YX2} = 0.4660$, $S_{YX2} = 2.5590$
$\bar{Y}_2 = 7.3864$, $S_{Y2} = 2.8949$, $C_{Y2} = 0.3919$
$\bar{X}_2 = 5.9132$, $S_{X2} = 1.8968$, $C_{X2} = 0.3208$

We estimate the empirical *Bias* and *MSE* using 5000 samples of various sizes from the study populations. The absolute relative bias (ARB), used in the tables below, is given by

$$\left|\frac{Bias(\hat{\mu}_\lambda)}{\bar{Y}}\right|, \text{ where } \lambda = YWst, RWst, \mathrm{Re}\, gWst.$$

The empirical and the theoretical results for the two estimators under study are presented in Table 1 and Table 2, for the higher and lower correlation respectively. From these tables we can observe that the proposed estimators show similar Bias as compared to the RRT mean.

Table 1: Empirical and theoretical absolute relative bias for the ratio estimator (underlined) and for the regression estimator (bold) relative to the RRT mean estimator

Population	n	W	Simple Random Sample (SRS) Empirical ARB	Simple Random Sample (SRS) Theoretical ARB	Stratified Random Sample (Str) Empirical ARB	Stratified Random Sample (Str) Theoretical ARB
$N = 5000$, $\rho_{XY} = 0.8020$	200	0.0	0.000055	0.000000	0.000477	0.000000
			<u>0.000322</u>	<u>0.000232</u>	<u>0.000152</u>	<u>0.000174</u>
			0.000055	**0.000045**	**0.000303**	**0.000032**
		0.1	0.001747	0.000000	0.002430	0.000000
			<u>0.001598</u>	<u>0.000232</u>	<u>0.001678</u>	<u>0.000177</u>
			0.001450	**0.000042**	**0.001391**	**0.000060**
		0.2	0.000473	0.000000	0.000344	0.000000
			<u>0.000619</u>	<u>0.000232</u>	<u>0.001092</u>	<u>0.000180</u>
			0.000770	**0.000066**	**0.001434**	**0.000066**
		0.3	0.001221	0.000000	0.002448	0.000000
			<u>0.001083</u>	<u>0.000232</u>	<u>0.003188</u>	<u>0.000181</u>
			0.000924	**0.000060**	**0.003586**	**0.000065**
		0.4	0.004778	0.000000	0.003135	0.000000
			<u>0.004638</u>	<u>0.000232</u>	<u>0.003873</u>	<u>0.000182</u>
			0.004481	**0.000068**	**0.004318**	**0.000051**
		0.5	0.002804	0.000000	0.001307	0.000000
			<u>0.002664</u>	<u>0.000232</u>	<u>0.002050</u>	<u>0.000183</u>
			0.002507	**0.000061**	**0.002534**	**0.000030**
		0.6	0.000857	0.000000	0.006586	0.000000
			<u>0.000736</u>	<u>0.000232</u>	<u>0.007353</u>	<u>0.000184</u>
			0.000560	**0.000068**	**0.007854**	**0.000002**
		0.7	0.004414	0.000000	0.007273	0.000000
			<u>0.004290</u>	<u>0.000232</u>	<u>0.008038</u>	<u>0.000184</u>
			0.004117	**0.000076**	**0.008581**	**0.000020**
		0.8	0.006108	0.000000	0.009377	0.000000
			<u>0.005992</u>	<u>0.000232</u>	<u>0.010134</u>	<u>0.000184</u>
			0.005811	**0.000070**	**0.010724**	**0.000025**
		0.9	0.003888	0.000000	0.012151	0.000000
			<u>0.003775</u>	<u>0.000232</u>	<u>0.012904</u>	<u>0.000185</u>
			0.003590	**0.000095**	**0.013536**	**0.000020**
		1.0	0.005241	0.000000	0.010205	0.000000
			<u>0.005626</u>	<u>0.000232</u>	<u>0.009906</u>	<u>0.000185</u>
			0.005351	**0.000091**	**0.009954**	**0.000013**

Table 2: Empirical and theoretical absolute relative bias for the ratio estimator (underlined) and for the regression estimator (bold) relative to the RRT mean estimator for a lower correlation

Population	n	W	Simple Random Sample (SRS) Empirical ARB	Simple Random Sample (SRS) Theoretical ARB	Stratified Random Sample (Str) Empirical ARB	Stratified Random Sample (Str) Theoretical ARB
			0.000048	0.000000	0.000208	0.000000
		0.0	0.000732	0.000595	0.000158	0.000353
			0.000032	**0.000030**	**0.000631**	**0.000037**
			0.001729	0.000000	0.002116	0.000000
		0.1	0.001901	0.000595	0.001845	0.000354
			0.001524	**0.000018**	**0.001343**	**0.000076**
			0.000492	0.000000	0.000012	0.000000
		0.2	0.000317	0.000595	0.000285	0.000354
			0.000698	**0.000042**	**0.000823**	**0.000062**
			0.001203	0.000000	0.001882	0.000000
		0.3	0.001383	0.000595	0.001609	0.000354
			0.000997	**0.000036**	**0.001034**	**0.000061**
			0.004760	0.000000	0.005493	0.000000
		0.4	0.004938	0.000595	0.005224	0.000355
			0.004555	**0.000040**	**0.004610**	**0.000006**
$N = 5000$			0.002787	0.000000	0.003373	0.000000
	200	0.5	0.002962	0.000595	0.003093	0.000355
$\rho_{XY} = 0.4926$			**0.002581**	**0.000032**	**0.002456**	**0.000037**
			0.000839	0.000000	0.001724	0.000000
		0.6	0.001041	0.000595	0.001445	0.000355
			0.000633	**0.000063**	**0.000774**	**0.000007**
			0.004396	0.000000	0.005336	0.000000
		0.7	0.004596	0.000595	0.005061	0.000355
			0.004191	**0.000066**	**0.004354**	**0.000053**
			0.006091	0.000000	0.007229	0.000000
		0.8	0.006295	0.000595	0.006954	0.000356
			0.005885	**0.000060**	**0.006217**	**0.000058**
			0.003870	0.000000	0.005102	0.000000
		0.9	0.004078	0.000595	0.004825	0.000356
			0.003664	**0.000086**	**0.004060**	**0.000077**
			0.005250	0.000000	0.005551	0.000000
		1.0	0.006040	0.000595	0.005186	0.000356
			0.005329	**0.000073**	**0.004324**	**0.000034**

Tables 3 and 4 below give the empirical and theoretical *MSE*'s for the proposed combined estimators based on 1st order approximation. We use the following expressions to find the percent relative efficiency (*PRE*) of proposed estimators as compared to the ordinary sample mean in both designs:

$$PRE_{RS} = \frac{MSE(\hat{\mu}_{YW})}{MSE(\hat{\mu}_{\alpha})} \times 100 \text{, where } \alpha = RW, \mathrm{Re}\,gW \qquad \text{(simple random sampling)}$$

and
$$PRE_{St} = \frac{MSE(\hat{\mu}_{YWst})}{MSE(\hat{\mu}_{\beta})} \times 100 \text{, where } \beta = RWst, \mathrm{Re}\,gWst \qquad \text{(stratified sampling)}$$

These measures are calculated using first degree of approximation for *MSE*. We estimate the empirical *MSE* using 5000 samples of size *n* from the simulated bivariate population.

We obtain the Optionality Effect by calculating the ratio of non-optional RRT *MSE* values (W=1) relative to the *MSE* of the corresponding Optional RRT estimator and multiplying it by 100.

We also calculate the Design Effect (*Deff*) comparing the *MSE's* of the proposed estimators in stratified sampling (Str) relative to the ordinary sample mean in simple random sample (SRS):

$$Deff = \frac{MSE(\hat{\mu}_{YW})}{MSE(\hat{\mu}_{\lambda})} \times 100, \text{ where } \lambda = YWst, RWst, \mathrm{Re}\, gWst.$$

Table 3: Empirical and Theoretical MSE, *PRE* for the ratio estimator (underlined) and for the regression estimator (bold) relative to the RRT mean estimator, *PRE* for the simple random sample (SRS) relative to the stratified sample (Str)

			Simple Random Sample (SRS)			Optionality	Stratified Random Sample (Str)			Optionality	
Population	n	W	Empirical *MSE*	Theoretical *MSE*	PRE_{RS}	Effect	Empirical *MSE*	Theoretical *MSE*	PRE_{St}	Effect	*Deff*
			0.0422	0.0428	100.00	202.95	0.0367	0.0356	100.00	226.01	120.39
		0.0	0.0101	0.0098	437.37	550.27	0.0165	0.0191	186.15	338.51	224.11
			0.0094	**0.0092**	**464.16**	**577.86**	**0.0173**	**0.0169**	**210.66**	**365.67**	**253.62**
			0.0483	0.0473	100.00	184.01	0.0406	0.0401	100.00	200.72	117.93
		0.1	0.0142	0.0142	332.62	379.43	0.0229	0.0238	168.20	271.64	198.36
			0.0137	**0.0136**	**346.41**	**391.01**	**0.0247**	**0.0214**	**187.05**	**288.35**	**220.58**
			0.0533	0.0517	100.00	168.30	0.0449	0.0446	100.00	180.52	115.96
		0.2	0.0195	0.0186	277.51	289.53	0.0276	0.0285	156.55	227.37	181.53
			0.0190	**0.0181**	**286.20**	**295.47**	**0.0307**	**0.0259**	**171.81**	**238.21**	**199.23**
			0.0574	0.0561	100.00	155.06	0.0486	0.0490	100.00	164.01	114.35
		0.3	0.0242	0.0230	243.51	234.08	0.0319	0.0331	148.36	195.77	169.64
			0.0236	**0.0225**	**249.64**	**237.45**	**0.0366**	**0.0304**	**161.16**	**203.00**	**184.28**
			0.0620	0.0605	100.00	143.75	0.0538	0.0535	100.00	150.27	113.01
		0.4	0.0284	0.0274	220.44	196.45	0.0372	0.0376	142.28	172.02	160.79
			0.0278	**0.0269**	**225.08**	**198.48**	**0.0433**	**0.0349**	**153.28**	**176.90**	**173.22**
N = 5000			0.0650	0.0649	100.00	133.98	0.0578	0.0580	100.00	138.65	111.88
ρ_{XY} = 0.8020	200	0.5	0.0316	0.0318	203.76	169.24	0.0409	0.0422	137.58	153.49	153.93
			0.0310	**0.0313**	**207.45**	**170.50**	**0.0483**	**0.0394**	**147.21**	**156.77**	**164.70**
			0.0693	0.0693	100.00	125.46	0.0648	0.0625	100.00	128.70	110.91
		0.6	0.0371	0.0363	191.14	148.66	0.0464	0.0467	133.85	138.60	148.45
			0.0364	**0.0357**	**194.17**	**149.43**	**0.0550**	**0.0439**	**142.40**	**140.76**	**157.93**
			0.0741	0.0737	100.00	117.95	0.0695	0.0670	100.00	120.09	110.07
		0.7	0.0416	0.0407	181.26	132.53	0.0512	0.0512	130.80	126.38	143.97
			0.0408	**0.0401**	**183.81**	**132.99**	**0.0614**	**0.0484**	**138.48**	**127.72**	**152.42**
			0.0785	0.0781	100.00	111.29	0.0752	0.0715	100.00	112.55	109.33
		0.8	0.0465	0.0451	173.31	119.57	0.0576	0.0557	128.26	116.15	140.23
			0.0456	**0.0445**	**175.51**	**119.82**	**0.0698**	**0.0528**	**135.23**	**116.89**	**147.85**
			0.0835	0.0825	100.00	105.34	0.0816	0.0759	100.00	105.90	108.69
		0.9	0.0519	0.0495	166.77	108.91	0.0643	0.0602	126.12	107.47	137.07
			0.0510	**0.0489**	**168.70**	**109.01**	**0.0785**	**0.0573**	**132.48**	**107.76**	**143.99**
			0.0883	0.0869	100.00	100.00	0.0858	0.0804	100.00	100.00	108.11
		1.0	0.0564	0.0539	161.31	100.00	0.0644	0.0647	124.29	100.00	134.36
			0.0555	**0.0533**	**163.02**	**100.00**	**0.0787**	**0.0618**	**130.20**	**100.00**	**140.76**

Table 4: Empirical and Theoretical MSE, *PRE* for the ratio estimator (underlined) and for the regression estimator (bold) relative to the RRT mean estimator, *PRE* for the simple random sample (SRS) relative to the stratified sample (Str) for a lower correlation

Population	n	W	Simple Random Sample (SRS)			Optionality	Stratified Random Sample (Str)			Optionality	
			Empirical *MSE*	Theoretical *MSE*	PRE_{RS}	Effect	Empirical *MSE*	Theoretical *MSE*	PRE_{St}	Effect	*Deff*
		0.0	0.0422	0.0428	100.00	203.13	0.0386	0.0384	100.00	214.92	111.41
			0.0445	0.0433	98.78	201.88	0.0384	0.0416	92.37	206.64	102.91
			0.0325	**0.0324**	**132.04**	**236.17**	**0.0445**	**0.0345**	**111.28**	**228.41**	**123.98**
		0.1	0.0483	0.0472	100.00	184.14	0.0444	0.0428	100.00	192.77	110.23
			0.0484	0.0477	98.90	183.21	0.0441	0.0460	93.04	186.69	102.56
			0.0376	**0.0368**	**128.20**	**207.87**	**0.0526**	**0.0388**	**110.43**	**203.30**	**121.72**
		0.2	0.0533	0.0516	100.00	168.40	0.0505	0.0472	100.00	174.76	109.27
			0.0536	0.0521	98.99	167.71	0.0501	0.0504	93.60	170.26	102.28
			0.0428	**0.0412**	**125.18**	**185.62**	**0.0612**	**0.0432**	**109.22**	**182.29**	**119.35**
		0.3	0.0574	0.0560	100.00	155.13	0.0554	0.0516	100.00	159.82	108.48
			0.0581	0.0565	99.07	154.62	0.0550	0.0549	94.07	156.49	102.05
			0.0470	**0.0456**	**122.75**	**167.67**	**0.0685**	**0.0477**	**108.26**	**165.25**	**117.44**
		0.4	0.0620	0.0604	100.00	143.81	0.0605	0.0560	100.00	147.24	107.81
			0.0625	0.0609	99.14	143.43	0.0604	0.0593	94.48	144.79	101.86
			0.0514	**0.0500**	**120.74**	**152.89**	**0.0764**	**0.0521**	**107.48**	**151.14**	**115.87**
		0.5	0.0649	0.0648	100.00	134.02	0.0640	0.0604	100.00	136.49	107.24
$N = 5000$ $\rho_{XY} = 0.4926$	200		0.0652	0.0653	99.19	133.75	0.0631	0.0637	94.83	134.73	101.69
			0.0543	**0.0544**	**119.06**	**140.51**	**0.0815**	**0.0566**	**106.83**	**139.26**	**114.57**
		0.6	0.0694	0.0692	100.00	125.48	0.0633	0.0649	100.00	127.21	106.74
			0.0717	0.0698	99.25	125.29	0.0624	0.0682	95.14	125.97	101.56
			0.0599	**0.0589**	**117.63**	**129.98**	**0.0837**	**0.0610**	**106.29**	**129.13**	**113.46**
		0.7	0.0742	0.0736	100.00	117.97	0.0682	0.0693	100.00	119.11	106.31
			0.0764	0.0742	99.29	117.84	0.0677	0.0726	95.41	118.29	101.44
			0.0644	**0.0633**	**116.40**	**120.92**	**0.0916**	**0.0655**	**105.82**	**120.37**	**112.50**
		0.8	0.0786	0.0780	100.00	111.30	0.0742	0.0737	100.00	111.97	105.93
			0.0811	0.0786	99.33	111.23	0.0737	0.0770	95.66	111.49	101.33
			0.0689	**0.0677**	**115.33**	**113.03**	**0.1003**	**0.0699**	**105.42**	**112.74**	**111.67**
		0.9	0.0836	0.0825	100.00	105.35	0.0795	0.0781	100.00	105.65	105.60
			0.0864	0.0830	99.37	105.31	0.0788	0.0814	95.87	105.43	101.24
			0.0742	**0.0721**	**114.40**	**106.12**	**0.1083**	**0.0743**	**105.07**	**106.01**	**110.95**
		1.0	0.0883	0.0869	100.00	100.00	0.0819	0.0825	100.00	100.00	105.30
			0.0916	0.0874	99.40	100.00	0.0815	0.0859	96.07	100.00	101.16
			0.0788	**0.0765**	**113.57**	**100.00**	**0.1138**	**0.0788**	**104.71**	**100.00**	**110.25**

According to the results in the two tables above, the Design Effect (*Deff*) shows an increase in efficiency by using a stratified sample, more so when the correlation between the auxiliary and study variable is high as seen in Table 3. All the PRE_{St} values in Table 3 are greater than100, indicating that the proposed combined estimators are more efficient than the mean estimator. Also all the PRE_{RS} in Table 3 are greater than 100, indicating that the estimators $\hat{\mu}_{RW}$ and $\hat{\mu}_{\mathrm{Reg}W}$ perform better than the ordinary mean estimator $\hat{\mu}_{YW}$ .This result agrees with Gupta et al. (2014) findings for simple random sampling when the correlation is high. We see in Table 3, that the proposed combined ratio estimator $\hat{\mu}_{RWst}$ and the proposed combined regression estimator $\hat{\mu}_{\mathrm{Reg}Wst}$ are both efficient than the mean estimator $\hat{\mu}_{YWst}$ .It can be seen below that the theoretically obtained sufficient (but not necessary) condition in (2.8) above given by

$$MSE(\hat{\mu}_{RWst}) < MSE(\hat{\mu}_{YWst}) \text{if} \sum_{h=1}^{L} \delta_h^2 \gamma_h C_{yh} C_{xh} \left(\rho_{yxh} - \frac{1}{2} \frac{C_{xh}}{C_{yh}} \right) > 0$$

holds true for the stratum statistics for population I with

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{yh} C_{xh} \left(\rho_{yxh} - \frac{1}{2} \frac{C_{xh}}{C_{yh}} \right) = 0.000247029 > 0$$

However the above condition does not hold for the stratum statistics for Population II, as

$$\sum_{h=1}^{L} \delta_h^2 \gamma_h C_{yh} C_{xh} \left(\rho_{yxh} - \frac{1}{2} \frac{C_{xh}}{C_{yh}} \right) = -0.0000885389 < 0$$

Consequently we observe in Table 4, that the combined ratio estimator $\hat{\mu}_{RWst}$ is no longer more efficient than the mean estimator $\hat{\mu}_{YWst}$. However the proposed combined regression estimator $\hat{\mu}_{\mathrm{Re}gWst}$ remains more efficient than both the proposed combined ratio estimator $\hat{\mu}_{RWst}$ and the mean estimator $\hat{\mu}_{YWst}$, in Table 3 and in Table 4. This is justified by the theoretically obtained conditions (3.8) and (3.9) for the combined regression estimator $\hat{\mu}_{\mathrm{Re}gWst}$ which always hold true.

We obtained theoretically in (2.10) and (3.10) that the proposed optional combined ratio $\hat{\mu}_{RWst}$ and combined regression estimator $\hat{\mu}_{\mathrm{Re}gWst}$ will always be more efficient than the corresponding non-optional combined ratio and combined regression estimator given by $\hat{\mu}_{RWst}$ and $\hat{\mu}_{regWst}$ for $(W = 1)$. The same can be observed in both the Tables 3 and 4. Also it can be verified easily that the stratification in both the populations is such that the condition (1.14) holds true for both the populations and consequently we observe that $MSE(\hat{\mu}_{YWst}) < MSE(\hat{\mu}_{YW})$ in both Tables 3 and 4.

The PRE_{St} of all the proposed estimators $\hat{\mu}_{YWst}$, $\hat{\mu}_{RWst}$ and $\hat{\mu}_{\mathrm{Re}gWst}$ are greater than PRE_{St}*'s* of the corresponding non-optional estimators $(W = 1)$ of Sousa et al. (2014) showing that the use of optional RRT model has an advantage over the non-optional RRT model in the presence of auxiliary information in the context of stratified sampling also. As expected, we see that the optionality effect dissipates as the sensitivity level W increases.

5. Conclusions

From the discussions in Sections 2, 3, 4 and higher values of the optionality effect in Table 1 and Table 2, we infer that the proposed optional combined RRT estimators are more efficient than the corresponding non-optional combined estimators in Sousa et al. (2014). Also higher values of the Design Effect (*Deff*) show that the proposed combined estimators are more efficient than the estimators of Gupta et al. (2014) derived in simple random sampling. Although both the ratio and regression estimators perform better than the ordinary RRT mean estimator, the improvement is much larger with the regression estimator in both simple random sampling and stratified sampling and is most efficient for all values of ρ_{XY}. Clearly both Design and Optionality effects are smaller with smaller correlation value. The study also confirms that the estimation of the mean of a sensitive variable can be improved by using a non-sensitive auxiliary variable. The main conclusion of this study is that the advantage of using optional RRT model over the non-optional RRT model in the presence of auxiliary information still holds in the context of stratified sampling.

References

Cochran, W.G. (1997). *Sampling Techniques*, *3rd edition*, Wiley Eastern Ltd., New York.

Gupta, S.N., Gupta, B.C. and Singh, S. (2002). Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and inference,* **100**, 239-247.

Gupta, S., Kalucha, G., Shabbir, J. and Dass, B.K. (2014). Estimation of finite population mean using optional RRT models in the presence of non-sensitive auxiliary information. *American Journal of Mathematical and Management Sciences,* **33(2)**, 147-159.

Gupta, S. and Shabbir, J. (2008). On improvement in estimating the population mean in simple random sampling. *Journal of Applied Statistics*, **35(5)**, 559-566

Gupta, S., Shabbir, J., Sousa, R. and Real, P.C. (2012). Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics-Theory and Methods*, **41**, 2394-2404

Kadilar, C., Candan, M. and Cingi, H. (2007). Ratio estimators using robust regression. *Hacettepe Journal of Mathematics and Statistics,* **36(2)**, 81-188.

Kadilar, C. and Cingi, H. (2003). Ratio estimator in stratified sampling. *Biometrical Journal,* **45**, 218-225

Kadilar, C. and Cingi, H. (2004). Ratio Estimators in Simple Random Sampling. *Applied Mathematics and Computation*, **151**, 893-902

Kadilar, C. and Cingi, H. (2005). A new estimator in stratified random sampling. *Communication in Statistics-Theory and Methods,* **34**, 597-602

Koyuncu, N. and Kadilar, C. (2008). Ratio and product estimators in stratified random sampling. *Journal of Statistical planning and Inference*, **139(8)**, 2552-2558

Koyuncu, N. and Kadilar, C. (2009). Family of estimators of population mean using two auxiliary variables in stratified random sampling. *Communication in Statistics-Theory and Methods*, **38(14)**, 2398-2417.

Koyuncu, N. and Kadilar, C. (2010). On the family of estimators of population mean in stratified random sampling. *Pakistan Journal of Statistics*, **26(2)**, 427-443.

Mukhopadhyay, P. (1998). *Theory and Methods of Survey Sampling*, Prentice hall of India, New Delhi.

Nangsue, N. (2009). Adjusted ratio and regression type estimators for estimation of population mean when some observations are missing. *World Academy of Science, Engineering and Technology*, **53**, 787-790

Ray, S.K. and Singh, R.K. (1981). Difference-cum-ratio type estimators. *Journal of Indian Statistical Association,* **19**, 147-151

Shabbir, J. and Gupta, S. (2007). On improvement in variance estimation using auxiliary information. *Communication in Statistics–Theory and Methods,* **36(12)**, 2177-2185.

Shabbir, J. and Gupta, S. (2010). Estimation of the finite population mean in two-phase sampling when auxiliary variables are attribute. *Hacettepe Journal of Mathematics and Statistics,* **39(1)**, 121-129

Singh, H.P. and Solanki, R.S. (2012). An efficient class of estimators for the population mean using auxiliary information in systematic sampling, *Journal of Statistical Theory and Practice*, **6(2)**, 274-285

Singh, H.P. and Solanki, R.S. (2013). Efficient Ratio and product Estimators in Stratified Random Sampling. *Communications in Statistics–Theory and Methods,* **42(6)**, 1008-1023

Singh, H.P. and Vishwakarma, G.K. (2008). A family of estimators of population mean using auxiliary information in stratified sampling. *Communication in Statistics–Theory and Methods*, **37(7)**, 1038-1050

Singh, H.P. and Vishwakarma, G.K. (2007). Modified exponential ratio and product estimators for finite population mean in double sampling. *Austrian Journal of Statistics*, **36(3)**, 217-225

Sousa, R., Gupta, S., Shabbir, J. and Real, P.C. (2014). Improved mean estimation of a sensitive variable using auxiliary information in stratified sampling. *Journal of Statistics and Management Systems*, **17(5-6)**, 503-518

Sousa, R., Shabbir, J., Real, P.C. and Gupta, S. (2010). Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information. *Journal of Statistical Theory and Practice*, **4(3)**, 495-507.