REVSTAT – Statistical Journal
Volume 19, Number 2, April 2021, 275–290

VARIANCE ESTIMATION IN THE PRESENCE OF MEASUREMENT ERRORS UNDER STRATIFIED RANDOM SAMPLING

Authors: Neha Singh
– Department of Mathematics & Computing, Indian Institute of Technology (ISM), Dhanbad-826004, Jharkhand, India
nehaismstats@gmail.com

Gajendra K. Vishwakarma
– Department of Mathematics & Computing, Indian Institute of Technology (ISM), Dhanbad-826004, Jharkhand, India
vishwagk@rediffmail.com

Raj K. Gangele
– Department of Mathematics and Statistics, Dr. Harisingh Gour University (Central University),
Sagar-470003, M.P., India
rkgangele23@gmail.com

Received: August 2018 Revised: April 2019 Accepted: April 2019

Abstract:

- This study focuses on the estimation of population variance of study variable in stratified random sampling using auxiliary information when the observations are contaminated by measurement errors. Three classes of estimators of variance under measurement error are proposed by using the approach of Srivastava and Jhajj [18] for the study variable. The properties of the estimator viz. bias and mean square error of the proposed classes of estimators are provided. The conditions for which proposed estimators are more efficient compared to usual estimators are discussed. It is shown that the proposed classes of estimators include a large number of estimators of the population variance of stratified random sampling and their bias and mean square error can be easily derived.

Keywords:

- *variance; auxiliary variable; mean square error; measurement error; stratified random sampling.*

AMS Subject Classification:

- 49A05, 78B26.

1. INTRODUCTION

In survey sampling, the auxiliary information is mainly used in order to gain efficiency for the estimation. The literature on estimating the population variance by using auxiliary variable is substantial and widely discussed. Some authors including, Das and Tripathi [4], Srivastava and Jhajj [18], Isaki [5], Upadhyay and Singh [19, 20], Singh *et al.* [13], Prasad and Singh [8], Biradar and Singh [2], Singh and Biradar [11] have paid their attention towards the estimation of population variance of study variable using auxiliary information in simple random sampling. While dealing with planning surveys, in case of heterogeneous population, stratified random sampling has more importance in precise estimates over the simple random sampling. Singh and Vishwakarma [14] discussed a general method for the estimation of the variance of the stratified random sample mean by using auxiliary information.

The theories of survey sampling assume that the observations recorded during data collection are always free from measurement error. However, this assumption does not meet in many real-life situations and the data is contaminated with errors. The mean square error and other properties of the estimator obtained with significant measurement error may lead to serious fallacious results. Cochran [3], has discussed the source of measurement errors in survey data. Many authors such as Shalabh [9], Srivastava and Shalabh [16], Maneesha and Singh [7], Allen *et al.* [1], Shalabh and Tsai [10], Singh and Vishwakarma [15] have studied the impacts of measurement errors in the ratio, product and regression methods of estimation under simple random sampling.

Let us consider a finite heterogeneous population of size N, partitioned into L non-overlapping strata of sizes N_h, $h = 1, 2, ..., L$, where $\sum_{h=1}^{L} N_h = N$. Let (y_{hj}, x_{hj}) be the pair of observed values instead of true pair values (Y_{hj}, X_{hj}) of the study character y and the auxiliary character x respectively of the j-th unit $(j = 1, 2, ..., N_h)$ in the h-th stratum. Also, let (y_{hj}, x_{hj}) be the pair of values on (y, x) drawn from the h-th stratum $(j = 1, 2, ..., n_h$; $h = 1, 2, ..., L)$. It is familiar that in stratified random sampling an unbiased estimator of the population mean $(\mu_Y = \sum_{h=1}^{L} W_h \mu_{Yh};\ W_h = \frac{N_h}{N})$ of the variable y is given by

$$\bar{y}_{\text{st}} = \sum_{h=1}^{L} W_h \bar{y}_h, \tag{1.1}$$

where $\bar{y}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} y_{hj}$ is the sample mean of h-th stratum and $\mu_{Yh} = \frac{1}{N_h} \sum_{j=1}^{N_h} y_{hj}$ is the population mean of h-th stratum. Let the observational errors be

$$u_{hj} = y_{hj} - Y_{hj}, \quad v_{hj} = x_{hj} - X_{hj}, \tag{1.2}$$

which are normally distributed with mean zero and variances σ^2_{uh} and σ^2_{vh} respectively. Also let ρ_h be the population correlation coefficient between Y and X in h-th stratum. For simplicity in exposition, it is assumed that the variables u_{hj} and v_{hj} are uncorrelated although (Y_{hj}, X_{hj}) are correlated.

To obtain the bias and mean square error we define

$$\hat{\sigma}^2_{Yh} = \sigma^2_{Yh}(1 + \varepsilon_{0h}), \quad \hat{\sigma}^2_{Xh} = \sigma^2_{Xh}(1 + \varepsilon_{1h}), \quad \bar{x}_h = \mu_{Xh}(1 + \varepsilon_{2h}),$$

such that $E(\varepsilon_{ih}) = 0$, $\forall i = 0, 1, 2$;

$$E(\varepsilon_{0h}^2) = \frac{A_{Yh}}{n_h}, \quad E(\varepsilon_{1h}^2) = \frac{A_{Xh}}{n_h}, \quad E(\varepsilon_{2h}^2) = \frac{C_{Xh}^2}{n_h \theta_{Xh}}, \quad E(\varepsilon_{0h}\varepsilon_{1h}) = \frac{1}{n_h}(\delta_{22h} - 1),$$

$$E(\varepsilon_{1h}\varepsilon_{2h}) = \frac{1}{n_h}(\delta_{03h} C_{Xh}), \quad E(\varepsilon_{0h}\varepsilon_{2h}) = \frac{1}{n_h}(\delta_{21h} C_{Xh}),$$

where

$$A_{Yh} = \gamma_{2Yh} + \gamma_{2Uh}\frac{\sigma_{Uh}^4}{\sigma_{Yh}^4} + \frac{2}{\theta_{Yh}^2}, \qquad \beta_{2h}(Y) = \delta_{40h} = \frac{\mu_{40h}}{\mu_{20h}^2}, \qquad C_{Xh} = \frac{\sigma_{Xh}}{\mu_{Xh}},$$

$$A_{Xh} = \gamma_{2Xh} + \gamma_{2Vh}\frac{\sigma_{Vh}^4}{\sigma_{Xh}^4} + \frac{2}{\theta_{Xh}^2}, \qquad \beta_{2h}(X) = \delta_{04h} = \frac{\mu_{04h}}{\mu_{02h}^2},$$

$$\gamma_{2Yh} = \beta_{2h}(Y) - 3, \qquad \gamma_{2Xh} = \beta_{2h}(X) - 3, \qquad \gamma_{2Uh} = \beta_{2h}(U) - 3,$$

$$\gamma_{2Vh} = \beta_{2h}(V) - 3, \qquad \theta_{Yh} = \frac{\sigma_{Yh}^2}{\sigma_{Yh}^2 + \sigma_{Uh}^2}, \qquad \theta_{Xh} = \frac{\sigma_{Xh}^2}{\sigma_{Xh}^2 + \sigma_{Vh}^2},$$

$$\delta_{rsh} = \frac{\mu_{rsh}}{(\mu_{20h}^r \mu_{02h}^s)^{\frac{1}{2}}}, \qquad \mu_{rsh} = \frac{1}{N_h}\sum_{j=1}^{N_h}(y_{hj} - \mu_{Yh})^r (x_{hj} - \mu_{Xh})^s.$$

(r, s) are positive integers, μ_{Yh} and μ_{Xh} are the h-th stratum population mean of study and auxiliary variable respectively. C_{Xh} is the coefficient of variation of h-th stratum, θ_{Yh} and θ_{Xh} are the reliability ratio of h-th stratum of study and auxiliary variable respectively and lying between zero and one.

The variance of the stratified random sample mean is given by

$$\text{(1.3)} \qquad \mathrm{V}(\bar{y}_{\mathrm{st}}) = \sum_{h=1}^{L} W_h^2 \frac{\sigma_{Yh}^2}{n_h} = \sigma_{\mathrm{st}}^2,$$

where $\sigma_{Yh}^2 = \frac{1}{N_h}\sum_{j=1}^{N_h}(y_{ij} - \bar{\mu}_{Yh})^2$ is the population variance of h-th stratum.

The unbiased estimator of σ_{st}^2, i.e. $\mathrm{V}(\bar{y}_{\mathrm{st}})$, is given by

$$\text{(1.4)} \qquad \hat{\sigma}_{\mathrm{st}}^2 = \sum_{h=1}^{L} W_h^2 \frac{s_{yh}^2}{n_h},$$

where $s_{yh}^2 = \frac{1}{(n_h-1)}\sum_{j=1}^{n_h}(y_{hj} - \bar{y}_h)^2$ is an unbiased estimator of σ_{st}^2. But in the presence of measurement error s_{yh}^2 is not an unbiased estimator for σ_{st}^2. In the measurement error case the unbiased estimator of σ_{st}^2 is given by $\hat{\sigma}_{\mathrm{st}}^2 = \sum_{h=1}^{L} W_h^2 \frac{\hat{\sigma}_{Yh}^2}{n_h}$, where $\hat{\sigma}_{Yh}^2 = (s_{yh}^2 - \sigma_{uh}^2)$.

The variance of $\hat{\sigma}_{\mathrm{st}}^2$ in the presence of measurement error is given by

$$\text{(1.5)} \qquad \mathrm{V}(\hat{\sigma}_{\mathrm{st}}^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3}[A_{Yh}] = \mathrm{MSE}(\hat{\sigma}_{\mathrm{st}}^2).$$

Singh and Karpe [12] have studied the impact of measurement error on separate ratio and product also combined ratio as well as product estimators for the population mean under stratified random sampling. We have considered the problem of estimating population variance using information on the auxiliary variable by adopting Srivastava and Jhajj [18] method in stratified random sampling in the presence of measurement error. Three classes of estimators for the estimation of population variance are proposed under stratified random sampling when both the study and auxiliary variables are commingled with measurement errors as:

i) Estimator of variance σ^2_{st} when the mean μ_{Xh} of the auxiliary variable x in the h-th stratum of the population is known.

ii) Estimation of variance σ^2_{st} when the variance σ^2_{Xh} of the auxiliary variable x in the h-th stratum of the population is known.

iii) Estimation of variance σ^2_{st} when the mean μ_{Xh} and the variance σ^2_{Xh} of the auxiliary variable x in the h-th stratum of the population are known.

The crux of this study is to exhibit the effect of measurement errors on the estimates of the variance of the stratified random sample mean while using auxiliary information.

2. THE PROPOSED CLASSES OF ESTIMATORS

2.1. Estimation of population variance σ^2_{st} of the stratified simple random sample mean when mean μ_{Xh} of h-th stratum of the auxiliary variable x in the population is known

By using information on population mean μ_{Xh} of the h-th stratum of auxiliary variable, a class of estimators of population variance σ^2_{st} for the study variable is proposed as

$$\hat{\sigma}^2_a = \sum_{h=1}^{L} \left(\frac{W_h^2}{n_h}\right) \hat{\sigma}^2_{Yh} a_h(l_h), \tag{2.1}$$

where $l_h = \bar{x}_h/\mu_{Xh}$ and $a_h(\cdot)$ is a function of l_h such that $a_h = 1$. It satisfies conditions given by Srivastava [17] viz. function are continuous and bounded also the first as well as second order partial derivatives of the function exist. Expanding the function about the point 'unity' in a second order Taylor's series, we have

$$a_h(l_h) = a_h(1) + (l_h - 1)a_{1h}(1) + \frac{1}{2}(l_h - 1)^2 a_{2h}(1), \tag{2.2}$$

where a_{1h}, a_{2h} are first order and second order derivative with respect to l_h about point unity.

$$\begin{aligned} \hat{\sigma}^2_a &= \sum_{h=1}^{L} \left(\frac{W_h^2}{n_h}\right) \sigma^2_{Yh}(1 + \varepsilon_{0h}) \left[1 + (l_h - 1)a_{1h}(1) + \frac{1}{2}(l_h - 1)^2 a_{2h}(1)\right], \\ \hat{\sigma}^2_a &= \sum_{h=1}^{L} \left(\frac{W_h^2}{n_h}\right) \sigma^2_{Yh}(1 + \varepsilon_{0h}) \left[1 + \varepsilon_{2h} a_{1h}(1) + \frac{1}{2}\varepsilon^2_{2h} a_{2h}(1)\right], \end{aligned} \tag{2.3}$$

$$
\begin{aligned}
(\hat{\sigma}_a^2 - \sigma_{\mathrm{st}}^2) = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)(\sigma_{Yh}^2)\Big[&\varepsilon_{0h} + \varepsilon_{2h}a_{1h}(1) + \varepsilon_{0h}\varepsilon_{2h}a_{1h}(1) \\
&+ \frac{1}{2}\varepsilon_{2h}^2 a_{2h}(1) + \frac{1}{2}\varepsilon_{0h}\varepsilon_{2h}^2 a_{2h}(1)\Big].
\end{aligned} \tag{2.4}
$$

Taking expectation on both sides of (2.4) we get

$$
\mathrm{Bias}(\hat{\sigma}_a^2) = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h^2}\right)\sigma_{Yh}^2\left[\delta_{21h}C_{Xh}a_{1h}(1) + \frac{1}{2}\frac{C_{Xh}^2}{\theta_{Xh}}a_{2h}(1)\right]. \tag{2.5}
$$

For the mean square error we have

$$
(\hat{\sigma}_a^2 - \sigma_{\mathrm{st}}^2)^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\sigma_{Yh}^4\Big\{\varepsilon_{0h} + \varepsilon_{2h}a_{1h}(1)\Big\}^2, \tag{2.6}
$$

$$
(\hat{\sigma}_a^2 - \sigma_{\mathrm{st}}^2)^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)^2\sigma_{Yh}^4\Big\{\varepsilon_{0h}^2 + \varepsilon_{2h}^2 a_{1h}^2(1) + 2\varepsilon_{0h}\varepsilon_{2h}a_{1h}(1)\Big\}. \tag{2.7}
$$

Taking expectation up to terms of order n_h^{-3}, we get the mean square error of $\hat{\sigma}_a^2$ as

$$
\mathrm{MSE}(\hat{\sigma}_a^2) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[A_{Yh} + \frac{C_{Xh}^2}{\theta_{Xh}}a_{1h}^2(1) + 2\delta_{21h}C_{Xh}a_{1h}(1)\right]. \tag{2.8}
$$

The MSE in (2.8) is minimized for

$$
a_{1h}(1) = -\left(\frac{\delta_{21h}\theta_{Xh}}{C_{Xh}}\right). \tag{2.9}
$$

Thus, the resultant minimum MSE of $\hat{\sigma}_a^2$ is given by

$$
\mathrm{min.MSE}(\hat{\sigma}_a^2) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\Big[A_{Yh} - \delta_{21h}^2\theta_{Xh}\Big]. \tag{2.10}
$$

Hence, a theorem can be established as follows.

Theorem 2.1. *Up to terms of the order* n_h^{-3},

$$
\mathrm{min.MSE}(\hat{\sigma}_a^2) \geq \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\Big[A_{Yh} - \delta_{21h}^2\theta_{Xh}\Big],
$$

with equality holding if $a_{1h}(1) = -\left(\frac{\delta_{21h}\theta_{Xh}}{C_{Xh}}\right)$.

The following estimators

$$\hat{\sigma}^2_{a1} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh} l_h^{\alpha 1h}, \qquad \hat{\sigma}^2_{a2} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\Big[2 - l_h^{\alpha_{1h}}\Big],$$

$$\hat{\sigma}^2_{a3} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\left[\frac{\alpha_{1h}+l_h}{1+\alpha_{1h}l_h}\right], \qquad \hat{\sigma}^2_{a4} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\Big[\alpha_{1h} + (1-\alpha_{1h})l_h\Big],$$

$$\hat{\sigma}^2_{a5} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\Big[\alpha_{1h} + (1-\alpha_{1h})l_h^{-1}\Big],$$

$$\hat{\sigma}^2_{a6} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\Big[\alpha_{1h} + (1-\alpha_{1h})l_h^{-\alpha_{2h}}\Big],$$

$$\hat{\sigma}^2_{a7} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\Big[\alpha_{1h} + (1-\alpha_{1h})l_h\Big]^{-1},$$

are some of the members of the proposed class of estimators $\hat{\sigma}^2_a$. The optimum values of the scalars α_{1h} and α_{2h} can be derived from the right-hand side (2.9) of and the minimum mean square error of the listed estimators can be derived from (2.8). The lower bound of the MSE of estimators $\hat{\sigma}^2_{ai}, (i = 1 \text{ to } 7)$ is the same as given by (2.10).

Following by [17] and Srivastava and Jhajj [18] we have proposed a wider class of estimators of σ^2_{st} as

$$\hat{\sigma}^2_D = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)D_h(\hat{\sigma}^2_{Yh}, l_h), \tag{2.11}$$

where function $D_h(\cdot,\cdot)$ satisfies

$$D_h(\sigma^2_{Yh}, 1) = \sigma^2_{Yh} \quad \Rightarrow \quad D_{1h}(\sigma^2_{Yh}, 1) = \frac{\partial D_h(\cdot)}{\partial \hat{\sigma}^2_{Yh}}|_{(\sigma^2_{Yh},1)} = 1.$$

It can be shown that the minimum MSE of $\hat{\sigma}^2_D$ and the minimum MSE of $\hat{\sigma}^2_a$ are equal. We can state that the difference type estimator

$$\hat{\sigma}^2_{\text{std}_1} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\Big\{\hat{\sigma}^2_{Yh} + d_{1h}(l_h - 1)\Big\}, \tag{2.12}$$

is a member of class $\hat{\sigma}^2_D$ where d_{1h} is a suitably chosen constant.

2.2. Estimation of population variance σ^2_{st} of the stratified simple random sample mean when variance σ^2_{Xh} of h-th stratum of the auxiliary variable x in the population is known

A class of estimators of the variance σ^2_{st} of the stratified simple random sample mean when the variance σ^2_{Xh} of the auxiliary variable x of the h-th stratum in the population is known, is defined as

$$\hat{\sigma}^2_b = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh} b_h(m_h), \tag{2.13}$$

where $m_h = \frac{\hat{\sigma}^2_{Xh}}{\sigma^2_{Xh}}$, and $b_h(m_h)$ is a function of m_h such that $b_h(1) = 1$. The function is continuous and bounded in $\mathbb{R}$ and its first as well as the second order partial derivatives exist. Now expanding the function at point 'unity' in a second order Taylor's series, we can write

$$b_h(m_h) = b_h(1) + (m_h - 1)b_{1h}(1) + \frac{1}{2}(m_h - 1)^2 b_{2h}(1), \tag{2.14}$$

where $b_{1h}(1)$ and $b_{2h}(1)$ are the first order and second order derivative with respect to m_h of the function $b_h(m_h)$ about the point 'unity'.

$$\hat{\sigma}^2_b = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}\left[b_h(1) + (m_h - 1)b_{1h}(1) + \frac{1}{2}(m_h - 1)^2 b_{2h}(1)\right], \tag{2.15}$$

$$\hat{\sigma}^2_b = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\sigma^2_{Yh}(1 + \varepsilon_{0h})\left[1 + \varepsilon_{1h}b_{1h}(1) + \frac{1}{2}\varepsilon^2_{1h}b_{2h}(1)\right]. \tag{2.16}$$

To calculate the bias and the MSE of the estimator we can write

$$\begin{aligned}(\hat{\sigma}^2_b - \sigma^2_{\text{st}}) = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\sigma^2_{Yh}\Big[&\varepsilon_{0h} + \varepsilon_{1h}b_{1h}(1) + \varepsilon_{0h}\varepsilon_{1h}b_{1h}(1)\\ &+ \frac{1}{2}\varepsilon^2_{1h}b_{2h}(1) + \frac{1}{2}\varepsilon_{0h}\varepsilon^2_{1h}b_{2h}(1)\Big].\end{aligned} \tag{2.17}$$

Taking expectation on both sides of (2.17) we get the bias of $\hat{\sigma}^2_b$ as

$$\text{Bias}(\hat{\sigma}^2_b) = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h^2}\right)\sigma^2_{Yh}\left[(\delta_{22h} - 1)b_{1h}(1) + \frac{1}{2}A_{Xh}b_{2h}(1)\right]. \tag{2.18}$$

For the mean square error we have

$$(\hat{\sigma}^2_b - \sigma^2_{\text{st}})^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)^2\sigma^4_{Yh}\left[\varepsilon_{0h} + \varepsilon_{1h}b_{1h}(1)\right]^2, \tag{2.19}$$

$$(\hat{\sigma}^2_b - \sigma^2_{\text{st}})^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)^2\sigma^4_{Yh}\left[\varepsilon^2_{0h} + \varepsilon^2_{1h}b^2_{1h}(1) + 2\varepsilon_{0h}\varepsilon_{1h}b_{1h}(1)\right]. \tag{2.20}$$

Taking expectation up to terms of order n_h^{-3}, we get the mean square error of $\hat{\sigma}^2_b$ as

$$\text{MSE}(\hat{\sigma}^2_b) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[A_{Yh} + A_{Xh}b^2_{1h}(1) + 2(\delta_{22h} - 1)b_{1h}(1)\right], \tag{2.21}$$

which is minimized for

$$b_{1h}(1) = -\left(\frac{\delta_{22h} - 1}{A_{Xh}}\right). \tag{2.22}$$

Thus, the resultant minimum MSE of $\hat{\sigma}^2_b$ can be written as:

$$\text{min.MSE}(\hat{\sigma}^2_b) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[A_{Yh} - \frac{(\delta_{22h} - 1)^2}{A_{Xh}}\right]. \tag{2.23}$$

Hence, a theorem can be established as follows.

Theorem 2.2. *Up to terms of order* n_h^{-3},

$$\min.\text{MSE}(\hat{\sigma}_b^2) \geq \sum_{h=1}^{L} \frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[A_{Yh} - \frac{(\delta_{22h}-1)^2}{A_{Xh}}\right],$$

with equality holding if $b_{1h}(1) = -\left(\frac{\delta_{22h}-1}{A_{Xh}}\right)$.

The listed estimators

$$\hat{\sigma}_{b1}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2 m_h^{\eta_{1h}}, \qquad \hat{\sigma}_{b2}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[2 - m_h^{\eta_{1h}}\right],$$

$$\hat{\sigma}_{b3}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\frac{\eta_{1h}+m_h}{1+\eta_{1h}m_h}\right], \quad \hat{\sigma}_{b4}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\eta_{1h} + (1-\eta_{1h})m_h\right],$$

$$\hat{\sigma}_{b5}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\eta_{1h} + (1-\eta_{1h})\ m_h^{-1}\right],$$

$$\hat{\sigma}_{b6}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\eta_{1h} + (1-\eta_{1h})m_h^{\eta_{2h}}\right],$$

$$\hat{\sigma}_{b7}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\eta_{1h} + (1-\eta_{1h})m_h\right]^{-1},$$

are some of the members of the proposed class of estimators $\hat{\sigma}_b^2$. The optimum values of the scalars η_{1h} and η_{2h} from $\hat{\sigma}_{b1}^2$ to $\hat{\sigma}_{b7}^2$ can be derived from (2.22) and the minimum mean square errors of each of the listed estimators can be derived from (2.21). The lower bound of the MSE of the estimators $\hat{\sigma}_{bi}^2$ $(i = 1 \text{ to } 7)$ is given by (2.23).

A wider class of estimators of σ_{st}^2 than $\hat{\sigma}_b^2$ is

$$\hat{\sigma}_e^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)e_h(\hat{\sigma}_{Yh}^2, m_h), \tag{2.24}$$

where $e_h(\hat{\sigma}_{Yh}^2, m_h)$ is a function of $(\hat{\sigma}_{Yh}^2, m_h)$ and

$$e_h(\sigma_{Yh}^2, 1) = \sigma_{Yh}^2 \ \Rightarrow\ e_{1h}(\sigma_{Yh}^2) = 1 \qquad \text{with} \quad e_{1h}(\sigma_{Yh}^2, 1) = \frac{\partial e_h(\cdot)}{\partial\hat{\sigma}_{Yh}^2}\Big|_{(\sigma_{Yh}^2,1)}.$$

It can be exhibited that up-to third order, the optimum mean square error of $\hat{\sigma}_e^2$ and $\hat{\sigma}_b^2$ is the same as given by (2.23). It can also be shown that the difference-type estimator

$$\hat{\sigma}_{\text{std}_2}^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\left\{\hat{\sigma}_{Yh}^2 + d_{2h}(m_h - 1)\right\} \tag{2.25}$$

is a specific member of the class of estimator $\hat{\sigma}_e^2$ but not of the $\hat{\sigma}_b^2$ class, where d_{2h} is an appropriately chosen constant.

2.3. Estimation of population variance σ^2_{st} of the stratified simple random sample mean when the population mean μ_{Xh} and the variance σ^2_{Xh} of h-th stratum of the auxiliary variable x in the population are known

We define a class of estimators $\hat{\sigma}^2_c$, for the estimation of variance of the stratified simple random sample mean when the mean μ_{Xh} and the variance σ^2_{Xh} of the auxiliary variable x of the h-th stratum in the population are known, as

$$\hat{\sigma}^2_c = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}^2_{Yh}c_h(l_h, m_h), \tag{2.26}$$

where $c_h(l_h, m_h)$ is a function of $l_h = \bar{x}_h/\mu_{Xh}$ and $m_h = \hat{\sigma}^2_{Xh}/\sigma^2_{Xh}$, such that $c_h(1,1) = 1$ and it also satisfies similar conditions as mentioned in [18]:

$$\begin{aligned} c_h(l_h, m_h) = \Big[&c_h(1,1) + (l_h-1)c_{1h}(1,1) + (m_h-1)c_{2h}(1,1) \\ &+ \frac{1}{2}\Big\{(l_h-1)^2c_{11h}(1,1) + 2(l_h-1)(m_h-1)c_{12h}(1,1) + (m_h-1)^2c_{22h}(1,1)\Big\} \\ &+ \frac{1}{6}\Big\{(l_h-1)^3c_{111h}(l_h^*, m_h^*) + 3(l_h-1)^2)(m_h-1)c_{112h}(l_h^*, m_h^*) \\ &+ 3(l_h-1)(m_h-1)^2c_{122h}(l_h^*, m_h^*) + (m_h-1)^3c_{222h}(l_h^*, m_h^*)\Big\}\Big], \end{aligned} \tag{2.27}$$

where

$$\begin{gathered} l_h^* = 1 + \phi(l_h - 1), \quad m_h^* = 1 + \phi(m_h - 1), \qquad 0 < \phi < 1; \\ \Big\{c_{1h}(1,1), c_{2h}(1,1)\}, \{c_{11h}(1,1), c_{12h}(1,1), c_{22h}(1,1)\Big\}, \\ \Big\{c_{111h}(l_h^*, m_h^*), c_{112h}(l_h^*, m_h^*), c_{122h}(l_h^*, m_h^*), c_{222h}(l_h^*, m_h^*)\Big\}, \end{gathered}$$

respectively denote the first, second and third order partial derivatives of the function $c_h(l_h, m_h)$. Expressing (2.27) in terms of ε_{0h}, ε_{1h} and ε_{2h} using $c_h(1,1) = 1$ we have

$$\begin{aligned} \hat{\sigma}^2_c = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\sigma^2_{Yh}(1+\varepsilon_{0h})\Big[&\Big\{1 + \varepsilon_{2h}c_{1h}(1,1) + \varepsilon_{1h}c_{2h}(1,1)\Big\} \\ &+ \frac{1}{2}\Big\{\varepsilon^2_{2h}c_{11h}(1,1) + 2\varepsilon_{1h}\varepsilon_{2h}c_{12h}(1,1) + \varepsilon^2_{1h}c_{22h}(1,1)\Big\} \\ &+ \frac{1}{6}\Big\{\varepsilon^3_{2h}c_{111h}(l_h^* m_h^*) + 3\varepsilon_{1h}\varepsilon^2_{2h}c_{112h}l_h^* m_h^* \\ &+ 3\varepsilon^2_{1h}\varepsilon_{2h}c_{122h}l_h^* m_h^* + \varepsilon^3_{1h}c_{222h}(l_h^* m_h^*)\Big\}\Big]. \end{aligned} \tag{2.28}$$

To calculate the bias and the MSE of the estimator we can write (2.28) as

$$\begin{aligned} \hat{\sigma}^2_c - \sigma^2_{\text{st}} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\sigma^2_{Yh}\Big[&\Big\{\varepsilon_{0h} + \varepsilon_{2h}c_{1h}(1,1) + \varepsilon_{1h}c_{2h}(1,1) \\ &+ \varepsilon_{0h}\varepsilon_{2h}c_{1h}(1,1) + \varepsilon_{0h}\varepsilon_{1h}c_{2h}(1,1)\Big\} \\ &+ \frac{1}{2}\Big\{\varepsilon^2_{2h}c_{11h}(1,1) + 2\varepsilon_{1h}\varepsilon_{2h}c_{12h}(1,1) + \varepsilon^2_{1h}c_{22h}(1,1)\Big\}\Big]. \end{aligned} \tag{2.29}$$

Taking expectation of both sides of (2.29) we get the bias of the estimator $\hat{\sigma}_c^2$:

$$\begin{aligned}\text{Bias}(\hat{\sigma}_c^2) = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h^2}\right)\sigma_{Yh}^2\Bigg[&\delta_{21h}C_{Xh}c_{1h}(1,1) + (\delta_{22h}-1)c_{2h}(1,1)\\ &+\frac{1}{2}\left(\frac{C_{Xh}^2}{\theta_{Xh}}c_{11h}(1,1) + 2\delta_{03h}C_{Xh}c_{12h}(1,1) + A_{Xh}c_{22h(1,1)}\right)\Bigg]. \end{aligned} \tag{2.30}$$

For the mean square error we have

$$(\hat{\sigma}_c^2 - \sigma_{\text{st}}^2)^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)^2\sigma_{Yh}^4\Big[\varepsilon_{0h} + \varepsilon_{2h}c_{1h}(1,1) + \varepsilon_{1h}c_{2h}(1,1)\Big]^2, \tag{2.31}$$

$$\begin{aligned}(\hat{\sigma}_c^2 - \sigma_{\text{st}}^2)^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)^2\sigma_{Yh}^4\Big[&\varepsilon_{0h}^2 + \varepsilon_{2h}^2c_{1h}^2(1,1) + \varepsilon_{1h}^2c_{2h}^2(1,1)\\ &+2\varepsilon_{0h}\varepsilon_{2h}c_{1h}(1,1) + 2\varepsilon_{1h}\varepsilon_{2h}c_{1h}(1,1)c_{2h}(1,1) + 2\varepsilon_{0h}\varepsilon_{1h}c_{2h}(1,1)\Big]. \end{aligned} \tag{2.32}$$

Taking expectation up-to order n_h^{-3}, we get the mean square error of $\hat{\sigma}_c^2$ as

$$\begin{aligned}\text{MSE}(\hat{\sigma}_c^2) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\Bigg[&A_{Yh} + \frac{C_{Xh}^2}{\theta_{Xh}}c_{1h}^2(1,1) + A_{Xh}c_{2h}^2(1,1)\\ &+2\delta_{21h}C_{Xh}c_{1h}(1,1) + 2\delta_{03h}C_{Xh}c_{1h}(1,1)c_{2h}(1,1) + 2(\delta_{22h}-1)c_{2h}(1,1)\Bigg], \end{aligned} \tag{2.33}$$

where $c_{1h}(1,1)$ and $c_{2h}(1,1)$ denote the first order partial derivatives of $c_h(l_h, m_h)$ with respect to l_h and m_h respectively about the point $(1,1)$:

$$\begin{bmatrix} C_{Xh}^2\theta_{Xh} & \delta_{03h}C_{Xh}\\ \delta_{021h}C_{Xh} & A_{Xh}\end{bmatrix}\begin{bmatrix} c_{1h}(1,1)\\ c_{2h}(1,1)\end{bmatrix} = -\begin{bmatrix}\delta_{21h}C_{Xh}\\ \delta_{22h}-1\end{bmatrix}. \tag{2.34}$$

By solving (2.34) we can determine the minimum values of $c_{1h}(1,1)$ and $c_{2h}(1,1)$ respectively as

$$\begin{aligned} c_{1h}(1,1) &= \frac{\Big[\delta_{03h}(\delta_{22h}-1) - \delta_{21h}A_{Xh}\Big]}{\Big[C_{Xh}(A_{Xh}/\theta_{Xh}) - \delta_{03h}^2\Big]},\\ c_{2h}(1,1) &= \frac{\Big[\delta_{03h}\delta_{21h} - (\delta_{22h}-1)/\theta_{Xh}\Big]}{\Big[(A_{Xh}/\theta_{Xh}) - \delta_{03h}^2\Big]}. \end{aligned} \tag{2.35}$$

Putting (2.35) in (2.33) we obtain minimum MSE of $\hat{\sigma}_c^2$ as

(2.36)

$$\text{min.MSE}(\hat{\sigma}_c^2) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[A_{Yh} - \frac{(\delta_{22h}-1)^2 + \delta_{21h}^2\theta_{Xh}A_{Xh} - 2\theta_{Xh}\delta_{21h}\delta_{03h}(\delta_{22h}-1)}{(A_{Xh} - \delta_{03h}^2\theta_{Xh})}\right],$$

$$\text{min.MSE}(\hat{\sigma}_c^2) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[A_{Yh} - \delta_{21h}^2\theta_{Xh} - \frac{\{\delta_{21h}\theta_{Xh}\delta_{03h} - (\delta_{22h}-1)\}^2}{(A_{Xh} - \delta_{03h}^2\theta_{Xh})}\right]. \tag{2.37}$$

Hence, a theorem can be established as follows.

Theorem 2.3. *Up to terms of order* n_h^{-3},

$$\min.\text{MSE}(\hat{\sigma}_c^2) \geq \sum_{h=1}^{L}\left(\frac{(W_h\sigma_{Yh})^4}{n_h^3}\right)\left[A_{Yh} - \delta_{21h}^2\theta_{Xh} - \frac{\left\{\delta_{21h}\theta_{Xh}\delta_{03h} - (\delta_{22h}-1)\right\}^2}{(A_{Xh} - \delta_{03h}^2\theta_{Xh})}\right],$$

with equality holding if

$$(2.38)\qquad \begin{aligned} c_{1h}(1,1) &= \frac{\left[\delta_{03h}(\delta_{22h}-1) - \delta_{21h}A_{Xh}\right]}{\left[C_{Xh}(A_{Xh}/\theta_{Xh}) - \delta_{03h}^2\right]},\\ c_{2h}(1,1) &= \frac{\left[\delta_{03h}\delta_{21h} - (\delta_{22h}-1)/\theta_{Xh}\right]}{\left[(A_{Xh}/\theta_{Xh}) - \delta_{03h}^2\right]}. \end{aligned}$$

The present estimators

$$\begin{aligned}
\hat{\sigma}_{c1}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2 l_h{}^{\alpha_{1h}} m_h^{\alpha_{2h}},\\
\hat{\sigma}_{c2}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\alpha_{1h}l_{1h} + (1-\alpha_{1h}\ )m_{1h}^{\alpha_{2h}}\right],\\
\hat{\sigma}_{c3}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\alpha_{3h}l_{1h}^{\alpha_{1h}} + (1-\alpha_{3h})m_{2h}^{\alpha_{2h}}\right],\\
\hat{\sigma}_{c4}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[\exp\left\{\alpha_{1h}(l_h-1) + \alpha_{2h}(m_h-1)\right\}\right],\\
\hat{\sigma}_{c5}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\frac{\hat{\sigma}_{Yh}^2}{\left[1 + \alpha_{1h}\left\{l_{1h}{}^{\alpha_{2h}} m_{2h}^{\alpha_{3h}} - 1\right\}\right]},\\
\hat{\sigma}_{c6}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[1 - a_{1h}(l_h-1) + \alpha_{2h}(m_h-1)\right],\\
\hat{\sigma}_{c7}^2 &= \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\hat{\sigma}_{Yh}^2\left[1 - \alpha_{1h}(l_h-1) + \alpha_{2h}(m_h-1)\right]^{-1},
\end{aligned}$$

are some particular members of the suggested class of estimator $\hat{\sigma}_c^2$. The mean square error of these estimators can be obtained from (2.33) by choosing the suitable value for the constants. The lower bound of the MSE of the estimators $\hat{\sigma}_{ci}^2$ ($i = 1$ to 7) is the same as given by (2.37).

A class of estimators for σ_{st}^2 wider than $\hat{\sigma}_c^2$ is proposed as

$$(2.39)\qquad \hat{\sigma}_f^2 = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)f_h(\hat{\sigma}_{Yh}^2, l_h, m_h),$$

where $f_h(\hat{\sigma}_{Yh}^2, l_h, m_h)$ is a function of $(\hat{\sigma}_{Yh}^2, l_h, m_h)$ such that

$$f_h(\sigma_{Yh}^2, 1, 1) = \sigma_{Yh}^2 \quad\Rightarrow\quad f_{1h}(\sigma_{Yh}^2, 1, 1) = \frac{\partial f_h(\sigma_{Yh}^2, l_h, m_h)}{\partial\hat{\sigma}_{Yh}^2}\ |_{(\sigma_{Yh}^2,1,1)} = 1.$$

It can unveil that up-to order n_h^{-3} the optimum MSE of $\hat{\sigma}_f^2$ is same as the optimum MSE of $\hat{\sigma}_c^2$ at (2.36) or (2.37) and is not reduced. The difference-type estimator

$$\hat{\sigma}^2_{\mathrm{st}d_3} = \sum_{h=1}^{L}\left(\frac{W_h^2}{n_h}\right)\left\{\hat{\sigma}^2_{Yh} + d_{3h}(l_h - 1) + d_{4h}(m_h - 1)\right\} \tag{2.40}$$

is a specific member of the class (2.39) but not (2.26), where d_{3h} and d_{4h} are acceptable constants.

3. EFFICIENCY COMPARISONS

To obtain the conditions for which the proposed classes of estimators $\hat{\sigma}_a^2$, $\hat{\sigma}_b^2$ and $\hat{\sigma}_c^2$ perform better than usual unbiased estimator $\hat{\sigma}^2_{\mathrm{st}}$, from (1.5), (2.10), (2.23) and (2.37) we can write

$$\mathrm{MSE}(\hat{\sigma}^2_{\mathrm{st}}) - \mathrm{min.MSE}(\hat{\sigma}^2_a) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\delta^2_{21h}\theta_{Xh} \geq 0, \tag{3.1}$$

$$\mathrm{MSE}(\hat{\sigma}^2_{\mathrm{st}}) - \mathrm{min.MSE}(\hat{\sigma}^2_b) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\frac{(\delta_{22h}-1)^2}{A_{Xh}} \geq 0, \tag{3.2}$$

(3.3)
$$\mathrm{MSE}(\hat{\sigma}^2_{\mathrm{st}}) - \mathrm{min.MSE}(\hat{\sigma}^2_c) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[\delta^2_{21h}\theta_{Xh} + \frac{\left\{\theta_{Xh}\delta_{03h}\delta_{21h} - (\delta_{22h}-1)\right\}^2}{(A_{Xh} - \delta^2_{h03}\theta_{Xh})}\right] \geq 0.$$

Remarks: To exhibit the impact of measurement error on MSE of the estimators, let the observation for both the study variable and auxiliary variable be recorded without error. Now the MSE of the proposed class of estimators $\hat{\sigma}_a^2$, to the third degree of approximation is given as

$$\mathrm{MSE}^*(\hat{\sigma}^2_a) = \sum_{h=1}^{L}\frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[(\delta_{40h}-1) + C^2_{Xh}a^2_{1h}(1) + 2\delta_{21h}C_{Xh}a_{1h}(1)\right], \tag{3.4}$$

which is the same as the obtained by Singh and Vishwakarma [14].

From (2.8) and (3.4) we have

$$\begin{aligned}\mathrm{MSE}(\hat{\sigma}^2_a) - \mathrm{MSE}^*(\hat{\sigma}^2_a) = \sum_{h=1}^{L}\frac{(W_h\sigma^4_{Yh}}{n_h^3}\Bigg[&(\gamma_{2Uh}+2)\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)^2 + 4\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)\\ &+ C^2_{Xh}(1-\theta_{Xh})a^2_{1h}(1)\Bigg].\end{aligned}$$

The difference is always positive in nature, thus we can infer that the presence of measurement error incorporates larger mean square error than the absence of measurement error.

To obtain the optimum value of the constant differentiating partially (3.4) with respect to a_{1h} and equate to zero we get

$$a_{1h}(1) = -\left(\frac{\delta_{21h}}{C_{Xh}}\right).$$

Thus, the resultant minimum mean square error is

$$\text{min.MSE}^*(\hat{\sigma}_a^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \Big[(\delta_{40h} - 1) - \delta_{21h}^2\Big]. \tag{3.5}$$

Now the impact of measurement error can be obtained (2.10) and (3.5) from and as

$$\begin{aligned} \text{min.MSE}(\hat{\sigma}_a^2) - \text{min.MSE}^*(\hat{\sigma}_a^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \Bigg[&(\gamma_{2Uh} + 2)\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)^2 \\ &+ 4\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right) + \delta_{21h}^2(1-\theta_{Xh})\Bigg]. \end{aligned} \tag{3.6}$$

The MSE of another proposed class of estimators $\hat{\sigma}_b^2$ for the estimation of σ_{st}^2 in the absence of measurement error is given in (3.7) and is similar to Singh and Vishwakarma [14]:

$$\text{MSE}^*(\hat{\sigma}_b^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \Big[\delta_{40h} - 1) + (\delta_{04h} - 1)b_{1h}^2(1) + 2(\delta_{22h} - 1)b_{1h}(1)\Big]. \tag{3.7}$$

From equation (2.21) and (3.7) we can write

$$\begin{aligned} \text{MSE}(\hat{\sigma}_b^2) - \text{MSE}^*(\hat{\sigma}_b^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \Bigg[&(\gamma_{2Uh} + 2)\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)^2 + 4\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right) \\ &+ (\gamma_{2Vh} + 2)\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right)^2 + 4\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right) b_{1h}^2(1)\Bigg]. \end{aligned} \tag{3.8}$$

The right-hand side of (3.8) is always positive in nature, thus we can infer that mean square error of the proposed estimator is always larger when observation is recorded with error. $\text{MSE}^*(\hat{\sigma}_b^2)$ is minimized for

$$b_{1h}(1) = -\left(\frac{\delta_{22h} - 1}{\delta_{04} - 1}\right). \tag{3.9}$$

Thus, the resultant minimum mean square error is

$$\text{min.MSE}^*(\hat{\sigma}_b^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \left[(\delta_{40h} - 1) - \frac{(\delta_{22h} - 1)^2}{(\delta_{04h} - 1)}\right]. \tag{3.10}$$

From (2.23) and (3.10) we can derive the impact of measurement error as

$$\begin{aligned} \text{min.MSE}(\hat{\sigma}_b^2) - \text{min.MSE}^*(\hat{\sigma}_b^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \Bigg[&(\gamma_{2Uh} + 2)\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)^2 + 4\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right) \\ &+ \frac{(\delta_{22h} - 1)^2}{A_{Xh}(\delta_{40h} - 1)}(\gamma_{2Vh} + 2)\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right)^2 + 4\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right)\Bigg]. \end{aligned} \tag{3.11}$$

The mean square error of the third proposed class of estimators $\hat{\sigma}_c^2$ for the estimation of $\hat{\sigma}_{\text{st}}^2$ in the absence of measurement error is given by Singh and Vishwakarma [14] as

$$\begin{aligned} \text{MSE}^*(\hat{\sigma}_c^2) = \sum_{h=1}^{L} \frac{(W_h \sigma_{Yh})^4}{n_h^3} \Big[&(\delta_{40h} - 1) + C_{Xh}^2 c_{1h}^2(1,1) + (\delta_{04h} - 1)c_{2h}^2(1,1) \\ &+ 2\delta_{21h} C_{Xh} c_{1h}(1,1) + 2\delta_{03h} C_{Xh} c_{1h}(1,1) c_{2h}(1,1) \\ &+ 2(\delta_{22h} - 1)c_{2h}(1,1)\Big]. \end{aligned} \tag{3.12}$$

From we can write as

$$\begin{aligned}\text{MSE}(\hat{\sigma}_c^2) - \text{MSE}^*(\hat{\sigma}_c^2) = \sum_{h=1}^{L} \frac{(W_h\sigma_{Yh})^4}{n_h^3} \Bigg[& (\gamma_{2Uh}+2)\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)^2 + 4\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right) \\ & + C_{Xh}^2\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right)c_{1h}^2(1) + (\gamma_{2Vh}+2)\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right)^2 \\ & + 4\left(\frac{1-\theta_{Xh}}{\theta_{Xh}}\right)c_{2h}^2\Bigg]\end{aligned} \tag{3.13}$$

and $\text{MSE}^*(\hat{\sigma}_c^2)$ is minimum for

$$\begin{aligned} c_{1h}(1,1) &= \frac{\Big[\delta_{03h}(\delta_{22h}-1) - \delta_{21h}(\delta_{04h}-1)\Big]}{C_{Xh}\Big[\delta_{04h} - \delta_{03h}^2 - 1\Big]}, \\ c_{2h}(1,1) &= \frac{\Big[\delta_{03h}\delta_{21h} - (\delta_{22h}-1)\Big]}{\Big[\delta_{04h} - \delta_{03h}^2 - 1\Big]}. \end{aligned} \tag{3.14}$$

Thus we can write the resultant minimum MSE as

$$\text{min.MSE}^*(\hat{\sigma}_c^2) = \sum_{h=1}^{L} \frac{(W_h\sigma_{Yh})^4}{n_h^3}\left[(\delta_{40h}-1) - \delta_{21h}^2 - \frac{\left\{\delta_{03h}\delta_{21h} - (\delta_{22h}-1)\right\}^2}{(\delta_{04h} - \delta_{03h}^2 - 1)}\right], \tag{3.15}$$

which can be easily obtained from (2.37) by putting $\sigma_{Uh}^2 = \sigma_{Vh}^2 = 0$.
Hence, we can derive the impact of measurement error on the mean square error of the estimator $(\hat{\sigma}_c^2)$ as

$$\begin{aligned} & \text{min.MSE}(\hat{\sigma}_c^2) - \text{min.MSE}^*(\hat{\sigma}_c^2) = \\ & = \sum_{h=1}^{L} \frac{(W_h\sigma_{Yh})^4}{n_h^3}\Bigg[(\gamma_{2Uh}+2)\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right)^2 + 4\left(\frac{1-\theta_{Yh}}{\theta_{Yh}}\right) + (1-\theta_{Xh})\delta_{21h}^2 \\ & \quad - \left\{\frac{A_1(\delta_{04h} - \delta_{03h}^2 - 1) - B_1(A_{Xh} - \delta_{h03}^2\theta_{Xh})}{(\delta_{04h} - \delta_{03h}^2 - 1)(A_{Xh} - \delta_{h03}^2\theta_{Xh})}\right\}\Bigg], \end{aligned} \tag{3.16}$$

where

$$\begin{aligned} A_1 &= \Big\{\delta_{03h}\delta_{21h}\theta_{Xh} - (\delta_{22h}-1)\Big\}^2, \\ B_1 &= \Big\{\delta_{03h}\delta_{21h} - (\delta_{22h}-1)\Big\}^2. \end{aligned}$$

The right-hand side of (3.16) is the effect of measurement error in the mean square error of the estimator which is always positive in nature. Thus, the proposed classes of estimators have larger MSE in the presence of measurement errors in both study and auxiliary variables than in the absence of measurement errors. When the measurement error is insignificant, the inference based on these data may remain valid. Nevertheless, when the amount of error is more significant in observed data, the inference may be invalid and inaccurate and often may lead to unexpected and undesirable consequences.

4. DISCUSSION AND CONCLUSION

The data available for statistical analysis are always contaminated with measurement error and may lead to fallacious inference results. When data contains heterogeneity among units in terms of value, survey users are advised to form several homogeneous groups, and the sampling design is well known as stratified sampling. To the best of our knowledge, the study of measurement error for the estimation of variance in stratified random sampling has not been addressed yet. Singh and Karpe [12] have studied the effect of measurement error on estimation of population mean in stratified random sampling. Estimation of variance has vital importance as it has practical uses in real-life. It is discussed by Lee [6], Srivastava and Jhajj [18], Wu [21] and Singh and Vishwakarma [14] without the measurement error framework.

The present study deals with the problem of estimation of variance by using auxiliary information under the stratified sampling framework when observations are contaminated by measurement errors. Three wider classes of estimators have been proposed. The theoretical comparisons show that the proposed classes of estimators ($\hat{\sigma}_a^2$, $\hat{\sigma}_b^2$ and $\hat{\sigma}_c^2$) in the presence of measurement error are more efficient than usual unbiased estimators. Since the proposed estimators are defined as a class, a large number of estimators become the members of this class. So the impact of measurement error on the bias and the mean square error of these estimators can be obtained easily. We can also conclude that the MSE in the presence of measurement error is larger than in the absence of it. Thus, the present study for the estimation of variance under measurement error for the stratified random sampling is useful and may attract others to carry out some work of practical use in this direction.

ACKNOWLEDGMENTS

Authors are deeply indebted to the editor-in-chief Prof. Maria Ivette Gomes and learned referee for their valuable suggestions leading to improvement the quality of contents and presentation of the original manuscript. Authors are also thankful to Dr. (Md.) Mojibur Rahman, Associate Professor of English Language, Department of Humanities and Social Sciences, Indian Institute of Technology (ISM) Dhanbad, India, for his kind support in improving the grammar of this manuscript.

REFERENCES

[1] Allen, J.; Singh, H.P. and Smarandache, F. (2003). A family of estimators of population mean using multi-auxiliary information in presence of measurement errors, *International Journal of Social Economics*, **30**(7), 837–849.

[2] Biradar, R.S. and Singh, H.P. (1994). An alternative to ratio estimator for population variance, *Assam Statistical Review*, **8**(2), 18–33.

[3] Cochran, W.G. (1968). Errors of measurement in statistics, *Technometrics*, **10**(4), 637–666.

[4] Das, A.K. and Tripathi, T.P. (1978). Use of auxiliary information in estimating the finite population variance, *Sankhya*, **40**(C), 139–148.

[5] Isaki, C.T. (1983). Variance estimation using auxiliary information, *Journal of the American Statistical Association*, **78**(381), 117–123.

[6] Lee, K.H. (1973). Variance estimation in stratified sampling, *Journal of the American Statistical Association*, **68**(342), 336–342.

[7] Maneesha and Singh, R.K. (2002). Role of regression estimator involving measurement errors, *Brazilian Journal of Probability and Statistics*, **16**(1), 39–46.

[8] Prasad, B. and Singh, H.P. (1990). Some improved ratio-type estimators of population variance using auxiliary information in sample surveys, *Communication in Statistics – Theory and Methods*, **21**(5), 1367–1376.

[9] Shalabh (1997). Ratio method of estimation in the presence of measurement errors, *Journal of the Indian Society of Agricultural Statistics*, **50**(2), 150–155.

[10] Shalabh and Tsai, J.R. (2017). Ratio and product methods of estimation of population mean in the presence of correlated measurement errors, *Communications in Statistics – Simulation and Computation*, **46**(7), 5566–5593.

[11] Singh, H.P. and Biradar, R.S. (1994). Estimation of finite population variance using auxiliary information, *Journal of the Indian Society of Statistics and Operation Research*, **15**(1), 47–63.

[12] Singh, H.P. and Karpe, N. (2010). Effect of measurement errors on the separate and combined ratio and product estimators in stratified random sampling, *Journal of Modern Applied Statistical Methods*, **9**(2), 388–402.

[13] Singh, H.P.; Upadhyaya, L.N. and Namjoshi, U.D. (1988). Estimation of finite population variance, *Current Science*, **57**(27), 1331–1334.

[14] Singh, H.P. and Vishwakarma, G.K. (2008). Some families of estimators of variance of stratified random sample mean using auxiliary information, *Journal of Statistical Theory and Practice*, **2**(1), 21–43.

[15] Singh, N. and Vishwakarma, G.K. (2019). A generalised class of estimator of population mean with the combined effect of measurement errors and non-response in sample survey, *Revista Investigacion Operacional*, **40**(2), 275–285.

[16] Srivastava, A.K. and Shalabh (2001). Effect of measurement errors on the regression method of estimation in survey sampling, *Journal of Statistical Research*, **35**(2), 35–44.

[17] Srivastava, S.K. (1971). A generalized estimator for the mean of a finite population using multiauxiliary information, *Journal of the American Statistical Association*, **66**, 404–407.

[18] Srivastava, S.K. and Jhajj, H.S. (1980). A class of estimators using auxiliary information for estimating finite population variance, *Sankhya*, **42**(C), 87–96.

[19] Upadhyaya, L.N. and Singh, H.P. (1983). Use of auxiliary information in the estimation of population variance, *Mathematical Forum*, **6**(2), 33–36.

[20] Upadhyaya, L.N. and Singh, H.P. (1986). On a dual to ratio estimator for estimating finite population variance, *Nepalese Mathematical Scientific Report*, **11**(1), 37–42.

[21] Wu, C.F.J. (1985). Variance estimation for the combined ratio and combined regression estimators, *Journal of the Royal Statistical Society, Series B (Methodological)*, **47**(1), 147–154.