SCIENCE FORUM (JOURNAL OF PURE AND APPLIED SCIENCES) 20 (2020) 307 – 315

http://dx.doi.org/10.5455/sf.90563

Science Forum (Journal of Pure and Applied Sciences)

journal homepage: www.atbuscienceforum.com

On the linear combination of sample variance, ratio, and product estimators of finite population variance under two-stage sampling

Olatunji Olawoyin Ishaq[1*], Ahmed Audu[2], Auwalu Ibrahim[3], Hassan S. Abdulkadir[4], Kabiru Tukur[5]

[1]*Department of Statistics, Kano Univ of Science and Technology, Wudil, Nigeria*

[2]*Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria*

ABSTRACT

Ratio estimators are sometime to be found less efficient. There is very little research to estimate the variance of the population, so in this work, an attempt was made to estimate the variance of the finite population under two-stage sampling in which the exponential ratio type estimator existing for estimating the variance of the finite population under two-stage sampling (case II) was modified and new five modified estimators for estimating finite population variance were obtained. The biases, mean square errors (MSEs), minimum MSEs, and PREs of proposed estimators were derived. The empirical study has been conducted with three real life data and the result shows that the proposed estimators have the minimum MSE and higher Percentage Relative Efficiency (PRE), which implies that the proposed estimators are more efficient than the existing estimators considered in the study.

ARTICLE INFO

Article history:
Received 04 March 2020
Received in revised form 16 March 2020
Accepted 17 March 2020
Published 09 April 2020
Available online 09 April 2020

KEYWORDS

Ratio
Variance
Two-stage sampling
Mean square error (MSE)
Efficiency

1. Introduction

Two-stage sampling technique or sub-sampling is a probability sampling technique which offers a variety of possibilities for effective use of auxiliary information. Two-stage sampling is used when the size of the clusters are large, making it difficult or expensive to observed all the units inside them. Unlike in single-stage sampling, where the required information is obtained from all the elements in the selected clusters, but in two-stage sampling, elements in the selected clusters, are again sampled.

In survey sampling, when available population is in the form of clusters, it is useful to use two-stage sampling design in order to reduce cost of the survey (Cochran, 1977; Kalton, 1983; Sarndal et al., 1992).

Mahalanobis (1967) was the first to introduce the two-stage sampling scheme which was extended by Godambe (1951). Saxena et al. (1984) introduced a better approach for multistage design. Most of the research has appeared in two-stage sampling for the estimation of the mean and total of the population, including Yunusa (2010), Saini (2013), and Singh et al. (2013). Many researchers, such as Wolter (1985), Das and Tripathi (1978), Isaki (1983), Upadhyaya et al. (2004), Gupta and Shabbir (2007), Singh and Vishwakarma (2008), Gupta and Shabbir (2010), have worked to estimate the variance of the finite population using simple random sampling and Muhammed et al. (2014) worked in estimating the variance of the finite population under two-stage sampling by adopting an exponential ratio type estimator.

*** Corresponding author** Olatunji Olawoyin Ishaq ✉ babington4u@gmail.com ✉ Department of Statistics, Kano Univ of Science and Technology, Wudil, Nigeria.

Isaki (1983) proposed the ratio type estimator to estimate the population variance of the study variable as

$$t_R = s_y^2\left(\frac{S_x^2}{s_x^2}\right) \tag{1.1}$$

The estimator's bias and mean square error (MSE) are:

$$\text{Bias}(t_R) = \gamma S_y^2\left[\left(\lambda_{40}-1\right)-\left(\lambda_{22}-1\right)\right] \tag{1.2}$$

$$\text{MSE}(t_R) = \gamma S_y^4\left[\left(\lambda_{40}-1\right)+\left(\lambda_{04}-1\right)-2\left(\lambda_{22}-1\right)\right] \tag{1.3}$$

The sample variance estimator of the population variance which is an unbiased estimator of the population variance is defined as

$$t_0 = s_y^2 \tag{1.4}$$

The variance of the estimator t_0 is

$$Var(t_0) = \gamma S_y^4\left(\lambda_{40}-1\right) \tag{1.5}$$

where $\lambda_{rs} = \dfrac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}}$,

$$\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^r\left(X_i-\bar{X}\right)^s \text{ and } \gamma = \left(\frac{1}{n}-\frac{1}{N}\right)$$

Gupta and Shabbir (2008) proposed a new hybrid class of estimators to estimate the population variance based on the two mean estimators of Sahai and Ray (1980) and Kaur (1989) to estimate mean Ȳ. One of the proposed estimators ($\hat{S}_{pr}^{2(0)}$) performs better than the regression estimator; the estimator is more efficient than the Kadilar and Cingi estimators (2006).

$$\hat{S}_{pr}^{2(\alpha)} = \left[d_1 s_y^2 + d_2\left(S_x^2 - s_x^2\right)\right]\left[2-\left(\frac{s_x^2}{S_x^2}\right)^{\alpha}\right], -1\le\alpha\le 1 \tag{1.6}$$

$$\text{Bias}\left(\hat{S}_{pr}^{2(\alpha)}\right) = \left(d_1-1\right)S_y^2 - d_1\alpha\gamma S_y^2\left[\begin{array}{l}\left(\lambda_{22}-1\right)+\\ \dfrac{\left(\alpha-1\right)}{2}\left(\beta_{2(x)}-1\right)\end{array}\right] + d_2\alpha\gamma S_x^2\left(\beta_{2(x)}-1\right) \tag{1.7}$$

$$\text{MSE}\left(\hat{S}_{pr}^{2(\alpha)}\right) = S_y^4 + d_1 S_y^4 A_1^{(\alpha)} + d_2^2 S_y^4 A_2 + 2d_1 d_2 S_y^2 S_x^2 A_3^{(\alpha)} - 2d_1 S_y^4 A_4^{(\alpha)} - 2d_2 S_y^2 S_x^2 A_5^{(\alpha)} \tag{1.8}$$

$$\text{MSE}(\hat{S}_{pr}^{2(\alpha)})_{opt} = S_y^4\left(1-\frac{A_1^{(\alpha)}A_5^{2(\alpha)} + A_2 A_4^{2(\alpha)} - 2A_3^{(\alpha)}A_4^{(\alpha)}A_5^{(\alpha)}}{A_1^{(\alpha)}A_2 - A_3^{2(\alpha)}}\right) \tag{1.9}$$

Singh et al (2011) proposed the exponential ratio estimator for the population variance as

$$t_{\text{Re}} = s_y^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right] \tag{1.10}$$

The bias and MSE up to the first order of approximation, respectively, are

$$\text{Bias}(t_{\text{Re}}) = \gamma S_y^2\left[\frac{3}{8}\left(\lambda_{40}-1\right)-\frac{1}{2}\left(\lambda_{22}-1\right)\right] \tag{1.11}$$

$$\text{MSE}(t_{\text{Re}}) = \gamma S_y^4\left[\left(\lambda_{40}-1\right)+\frac{\left(\lambda_{04}-1\right)}{4}-\left(\lambda_{22}-1\right)\right] \tag{1.12}$$

The usual linear regression estimator for population variance is

$$\hat{S}_{lr}^2 = s_y^2 + b\left(S_x^2 - s_x^2\right) \tag{1.13}$$

The MSE of $\hat{S}_{lr}^2$ to the first order of approximation is

$$\text{MSE}(\hat{S}_{lr}^2) = \gamma S_y^4\left[\left(\lambda_{40}-1\right)-\frac{\left(\lambda_{22}-1\right)^2}{\left(\lambda_{04}-1\right)}\right] \tag{1.14}$$

Yadav and Kadilar (2013) motivated by Upadhyaya et al. (2011) and following them, they proposed the improved an exponential ratio type estimator of the population variance as

$$t_{\text{Re}}^{(\alpha)} = s_y^2 \exp\left[1-\frac{\alpha s_x^2}{S_x^2+(\alpha-1)s_x^2}\right] \tag{1.15}$$

$$\text{Bias}(t_{\text{Re}}^{(\alpha)}) = \gamma S_y^2\frac{\left(\lambda_{04}-1\right)}{2\alpha^2}\left[2\alpha\left(1-\lambda\right)-1\right] \tag{1.16}$$

$$\text{MSE}(t_{\text{Re}}^{(\alpha)}) = \gamma S_y^4\left[\left(\lambda_{40}-1\right)+\frac{\left(\lambda_{04}-1\right)}{\alpha^2}\left(1-2\alpha\lambda\right)\right] \tag{1.17}$$

The MSE($t_{\text{Re}}^{(\alpha)}$) is minimum for $\alpha = \alpha_0 = \dfrac{1}{\lambda}$, substituting $\alpha = \dfrac{1}{\lambda}$ into Eq. (2.1.19), they get the asymptotically optimum estimator in the class of estimators $t_{\text{Re}}^{(\alpha)}$ as

$$t_{\mathrm{Re}}^{(\alpha_0)} = s_y^2 \exp\left[\frac{\lambda\left(S_x^2 - s_x^2\right)}{\lambda S_x^2 + (1-\lambda)s_x^2}\right] \quad (1.18)$$

$$\mathrm{MSE}(t_{\mathrm{Re}}^{\alpha_0}) = \gamma S_y^4\left[(\lambda_{40}-1) - \frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)}\right] \quad (1.19)$$

Olufadi and kadilar (2014) proposed an estimator to estimate population variance S_y^2, as follows

$$t = s_y^2\left(\frac{S_{x1}^2}{s_{x1}^2}\right)^{\alpha_1}\left(\frac{S_{x2}^2}{s_{x2}^2}\right)^{\alpha_2} \quad (1.20)$$

where α_1 and α_2 are real constants to be determined such that the MSE of t is minimum.

The Bias and MSE of the estimator were obtained as follows:

$$\mathrm{Bias}(t) = S_y^2\left(\frac{\alpha_1^2}{2}A_1 + \frac{\alpha_2^2}{2}A_2 + \alpha_1\alpha_2 A_5 - \alpha_1 A_3 - \alpha_2 A_4\right) \quad (1.21)$$

$$\mathrm{Bias}(t) = S_y^2\left(A_2 + \alpha_1^2 A_1 + \alpha_1^2 A_2 - 2\alpha_1 A_3 - 2\alpha_1 A_3 - 2\alpha_2 A_4 + 2\alpha_1\alpha_2 A_5\right) \quad (1.22)$$

Misra and Karan (2014) proposed a generalized double sampling estimator to estimate the variance of the finite population given by

$$d_g = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}f\left(\bar{y}, \bar{x}, \bar{x}^1\right) \quad (1.23)$$

$$\mathrm{Bias}\left(d_g\right) = \frac{-\mu_{20}}{n} - f_1\frac{\mu_{11}}{n} - f_2\frac{\mu_{11}}{n^1} - \bar{Y}f_{01}\frac{\mu_{11}}{n} - \bar{Y}f_{02}\frac{\mu_{11}}{n^1} - \bar{Y}f_{12}\frac{\mu_{02}}{n^1} - \frac{\bar{Y}}{2}f_{11}\frac{\mu_{02}}{n} - \frac{\bar{Y}}{2}f_{22}\frac{\mu_{02}}{n^1} \quad (1.24)$$

The bias and MSE of d_g are

$$\mathrm{MSE}(d_g) = \frac{1}{n}\left(\mu_{40} - \mu_{20}^2\right) + \bar{Y}^2 f_1^2\frac{\mu_{02}}{n} - \bar{Y}^2 f_1^2\frac{\mu_{02}}{n^1} - 2\bar{Y}^2 f_1\frac{\mu_{21}}{n} + 2\bar{Y}^2 f_1\frac{\mu_{21}}{n^1} \quad (1.25)$$

The usual variance estimator for the population variance in the two-stage sampling is given by

$$\hat{S}_{0(2s)}^2 = s_{y(2s)}^2 \quad (1.26)$$

The MSE of the estimator in Eq. (1.26) is

$$\mathrm{MSE}\left(\hat{S}_{0(2s)}^2\right) = S_y^4\left(\gamma\beta_{2(1y)}^* + \sum_{i=1}^{N}\gamma_{2i}\beta_{2(2yi)}^*\right) \quad (1.27)$$

Muhammad et al. (2014) proposed an improved exponential ratio type estimator for the population variance (S_y^2) in a two-stage sampling on the lines of Gupta and Shabbir (2008) as

$$\hat{S}_{p(2s)}^2 = \left[k_1 s_{y(2s)}^2 + k_2(S_x^2 - s_{x(2s)}^2)\right]\exp\left[\frac{S_x^2 - s_{x(2s)}^2}{S_x^2 + s_{x(2s)}^2}\right] \quad (1.28)$$

where k_1 and k_2 are suitably constants. They discussed two cases.

CASE I: When $K_1 + K_2 \neq 1$

At the first order of approximation, the bias and the MSE of estimator in Eq. (1.28) are

$$\mathrm{Bias}(\hat{S}_{p(2s)}^2) = \left[(k_1-1)S_y^2 + k_1 S_y^2\left(\gamma\left(\frac{3}{8}\beta_{2(1x)}^* - \frac{1}{2}\gamma_{22(1)}^*\right) + \sum_{i=1}^{N}\gamma_{2i}\left(\frac{3}{8}\beta_{2(2xi)}^* - \frac{1}{2}\gamma_{22(2i)}^*\right)\right)\right] + \frac{1}{2}k_2 S_x^2\left(\gamma\beta_{2(1x)}^* + \sum_{i=1}^{N}\gamma_{2i}\beta_{2(2xi)}^*\right) \quad (1.29)$$

$$\mathrm{MSE}(\hat{S}_{p(2s)}^2) = S_y^4\left(1 + k_1^2 A_1^* + k_2^2\beta_1^* + 2k_1k_2c_1^* - 2k_1O_1^* - 2k_2E_1^*\right) \quad (1.30)$$

$$\mathrm{MSE}\left(\hat{S}_{p(2s)}^{*2}\right)_{\min} = S_y^4\left(1 - \frac{\left(A_1^*E_1^* + B_1^*D_1^* - 2C_1^*D_1^*E_1^*\right)}{A_1^*B_1^* - C_1^{*2}}\right) \quad (1.31)$$

CASE II: when $k_1 + k_2 = 1$

In case II, the proposed estimator becomes

$$\hat{S}_{p(2s)}^* = \left[k_1 s_{y(2s)}^2 + (1-k_1)\left(S_x^2 - s_{x(2s)}^2\right)\right]\exp\left[\frac{S_x^2 - s_{x(2s)}^2}{S_x^2 + s_{x(2s)}^2}\right] \quad (1.32)$$

By Putting $k_2 = 1-k_1$ in Eq. (2.1.32) and Eq. (2.1.33), they, respectively, have the bias and the MSE of $\hat{S}_{p(2s)}^{*2}$ as

$$\text{Bias}\left(\hat{S}^{*2}_{p(2s)}\right)=\left[\begin{array}{l}(k_1-1)S_y^2+k_1S_y^2\left\{\begin{array}{l}\gamma\dfrac{1}{2}\left(\beta^*_{2(1x)}\left(\dfrac{3}{4}-\varphi\right)-\gamma^*_{22(1)}\right)\\+\displaystyle\sum_{i=1}^{N}\gamma_{2i}\dfrac{1}{2}\left(\begin{array}{l}\beta^*_{2(2xi)}\left(\dfrac{3}{4}-\varphi\right)\\-\gamma^*_{22(2i)}\end{array}\right)\end{array}\right\}\\+\dfrac{1}{2}\varphi\left\{\gamma\beta^*_{2(1x)}+\displaystyle\sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right\}\end{array}\right] \tag{1.33}$$

$$\text{MSE}\left(\hat{S}^{*2}_{p(2s)}\right)=S_y^4\Big[1+k_1^2\left(A_1^*+B_1^*-2C_1^*\right)-2k_1\left(B_1^*-C_1^*+D_1^*-E_1^*\right)+B_1^*-2E_1^*\Big] \tag{1.34}$$

$$\text{MSE}\left(\hat{S}^{*2}_{p(2s)}\right)_{\min}=S_y^4\left(1+B_1^*-2E_1^*-\frac{\left(B_1^*-C_1^*+D_1^*-E_1^*\right)^2}{\left(A_1^*+B_1^*-2C_1^*\right)}\right) \tag{1.35}$$

The estimator proposed $\hat{S}^{*2}_{p(2s)}$ in case II is better than the usual sample variance estimator $\hat{S}^2_{0(2s)}$.

2. Proposed Estimators

Having studying the Muhammad et al. (2014) ratio exponential estimator, the following estimators were proposed

$$T_{m1}=\alpha s^2_{y(2s)}+(1-\alpha)s^2_{y(2s)}\exp\left(\frac{S_x^2-s^2_{x(2s)}}{S_x^2+s^2_{x(2s)}}\right) \tag{2.1}$$

$$T_{m2}=\beta s^2_{y(2s)}+(1-\beta)s^2_{y(2s)}\exp\left(\frac{s^2_{x(2s)}-S_x^2}{s^2_{x(2s)}+S_x^2}\right) \tag{2.2}$$

$$T_{m3}=\lambda s^2_{y(2s)}\exp\left(\frac{S_x^2-s^2_{x(2s)}}{S_x^2+s^2_{x(2s)}}\right)+(1-\lambda)s^2_{y(2s)}\exp\left(\frac{s^2_{x(2s)}-S_x^2}{s^2_{x(2s)}+S_x^2}\right) \tag{2.3}$$

$$T_{m4}=\phi\frac{s^2_{y(2s)}}{s^2_{x(2s)}}S_x^2+(1-\phi)s^2_{y(2s)}\exp\left(\frac{S_x^2-s^2_{x(2s)}}{S_x^2+s^2_{x(2s)}}\right) \tag{2.4}$$

$$T_{m5}=\theta\frac{s^2_{y(2s)}}{S_x^2}s^2_{x(2s)}+(1-\theta)s^2_{y(2s)}\exp\left(\frac{s^2_{x(2s)}-S_x^2}{s^2_{x(2s)}+S_x^2}\right) \tag{2.5}$$

3. Properties of Proposed Estimators (BAIS AND MSE)

To get the bias and the MSE, let's define the following:

$$\left\{\begin{array}{l}
\gamma=\dfrac{1-f}{n},f=\dfrac{n}{N},\gamma_{2i}=\left(\dfrac{M_i}{\bar{M}}\right)^2\cdot\dfrac{1}{Nn}\left(\dfrac{1}{m_i}-\dfrac{1}{M_i}\right),\\
e_o=\dfrac{s^2_{y(2s)}-S^2_y}{S^2_y},\gamma^*_{22(1)}=\gamma_{22(1)}-1,\\
\gamma^*_{22(2i)}=\gamma_{22(2i)}-1,s^2_{y(2s)}=(1+e_0)S^2_y,\\
e_1=\dfrac{s^2_{x(2s)}-S^2_x}{S^2_x},s^2_{x(2s)}=(1+e_1)S^2_x,\\
E(e_0)=E(e_1)=0,\mu_{04(2i)}=\dfrac{\sum_{j=1}^{Mi}\left(x_{ij}-\bar{X}_i\right)^4}{N},\\
E(e_0^2)=\gamma\beta^*_{2(1y)}+\sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2yi)},\\
E(e_0e_1)=\gamma\gamma^*_{22(1)}+\sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)},\\
E(e_1^2)=\gamma\beta^*_{2(1x)}+\sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2yi)},\\
\mu_{rs(1)}=\dfrac{\sum_{i=1}^{N}\left(\dfrac{m_iy_i}{\bar{M}}-\bar{Y}\right)^r\left(\dfrac{m_ix_i}{\bar{M}}-\bar{X}\right)^s}{N},\\
\mu_{40(1)}=\dfrac{\sum_{i=1}^{N}\left(\dfrac{m_iy_i}{\bar{M}}-\bar{Y}\right)^4}{N},\mu_{20(1)}=\dfrac{\sum_{i=1}^{N}\left(\dfrac{m_iy_i}{\bar{M}}-\bar{Y}\right)^2}{N},\\
\mu_{04(1)}=\dfrac{\sum_{i=1}^{N}\left(\dfrac{m_ix_i}{\bar{M}}-\bar{X}\right)^4}{N},\mu_{02(1)}=\dfrac{\sum_{i=1}^{N}\left(\dfrac{m_ix_i}{\bar{M}}-\bar{X}\right)^2}{N},\\
\mu_{22(1)}=\dfrac{\sum_{i=1}^{N}\left(\dfrac{m_iy_i}{\bar{M}}-\bar{Y}\right)^2\left(\dfrac{m_ix_i}{\bar{M}}-\bar{X}\right)^2}{N},\\
\mu_{rs(2i)}=\dfrac{\sum_{j=1}^{Mi}\left(y_{ij}-\bar{Y}_i\right)^r\left(x_{ij}-\bar{X}_i\right)^s}{N},\mu_{40(2i)}=\dfrac{\sum_{j=1}^{Mi}\left(y_{ij}-\bar{Y}_i\right)^4}{N},\\
\mu_{20(2i)}=\dfrac{\sum_{j=1}^{Mi}\left(y_{ij}-\bar{Y}_i\right)^4}{N},\\
\mu_{02(2i)}=\dfrac{\sum_{j=1}^{Mi}\left(x_{ij}-\bar{X}_i\right)^2}{N},\mu_{22(2i)}=\dfrac{\sum_{j=1}^{Mi}\left(y_{ij}-\bar{Y}_i\right)^2\left(x_{ij}-\bar{X}_i\right)^2}{N},\\
\beta_{2(1y)}=\dfrac{\mu_{40(1)}}{\mu^2_{20(1)}},\beta_{2(1x)}=\dfrac{\mu_{04(1)}}{\mu^2_{02(1)}},
\end{array}\right.$$

$$
\begin{cases}
\beta_{2(2xi)}=\dfrac{\mu_{04(2i)}}{\mu_{02(2i)}^{2}},\beta^{*}_{2(1y)}=\beta_{2(1y)}-1,\beta^{*}_{2(1x)} \\
=\beta_{2(1x)}-1,\beta^{*}_{2(2yi)}=\beta_{2(2yi)}-1,\beta^{*}_{2(2xi)}=\beta_{2(2xi)}-1, \\
\gamma_{rs(2i)}=\dfrac{\mu_{rs(2i)}}{\mu_{20}^{r/2}\mu_{02}^{s/2}},\gamma_{22(1)}=\dfrac{\mu_{22(1)}}{\mu_{20(1)}\mu_{02(1)}},\gamma_{22(2i)}=\dfrac{\mu_{22(2i)}}{\mu_{20(2i)}\mu_{02(2i)}}, \\
\gamma_{rs(1)}=\dfrac{\mu_{rs(1)}}{\mu_{20(1)}^{r/2}\mu_{02(1)}^{s/2}},\beta_{2(2yi)}=\dfrac{\mu_{40(2i)}}{\mu_{20(2i)}^{2}},
\end{cases} \tag{3.1}
$$

3..1. Bias and MSE of T_{m1}, T_{m2}, T_{m3}, T_{m4}, and T_{m5}

Expressing Eq. (2.1) up to the first order of approximation, we have

$$T_{m1}=\alpha\left[S_y^2+e_0S_y^2\right]+(1-\alpha)\left[S_y^2+e_0S_y^2\right]$$

$$\exp\left[\frac{S_x^2-\left[S_x^2+e_1S_x^2\right]}{S_x^2+\left[S_x^2+e_1S_x^2\right]}\right] \tag{3.2}$$

Simplifying Eq. (3.3), we obtained,

$$T_{m1}=\alpha S_y^2(1+e_0)+(1-\alpha)S_y^2(1+e_0)\exp\left[\frac{-e_1}{2+e_1}\right] \tag{3.3}$$

$$T_{m1}=\alpha S_y^2(1+e_0)+(1-\alpha)S_y^2(1+e_0)\exp\left[\frac{-e_1}{2}\left(1+\frac{e_1}{2}\right)^{-1}\right] \tag{3.4}$$

$$T_{m1}=\alpha S_y^2(1+e_0)+(1-\alpha)S_y^2(1+e_0)\exp\left[\frac{-e_1}{2}\left(1-\frac{e_1}{2}+\frac{e_1^2}{4}\right)\right] \tag{3.5}$$

$$T_{m1}=S_y^2+e_0S_y^2+(1-\alpha)S_y^2\left[-\frac{e_1}{2}-\frac{e_0e_1}{2}+\frac{3}{8}e_1^2\right] \tag{3.6}$$

Subtracting S_y^2 from Eq. (3.6), take the expectation and apply the results of Eq. (3.1), the bias of T_{m1} is obtained as

$$T_{m1}-S_y^2=S_y^2\left[e_0+(1-\alpha)\left(-\frac{e_1}{2}-\frac{e_0e_1}{2}+\frac{3}{8}e_1^2\right)\right] \tag{3.7}$$

$$\text{Bias}(T_{m1})=S_y^2\left[(1-\alpha)\left(\frac{3}{8}\gamma\beta^{*}_{2(1x)}-\frac{1}{2}\gamma\gamma^{*}_{22(1)}+\frac{3}{8}\sum_{i=1}^{N}\gamma_{2i}\beta^{*}_{2(2xi)}-\frac{1}{2}\sum_{i=1}^{N}\gamma_{2i}\gamma^{*}_{22(2i)}\right)\right] \tag{3.8}$$

To obtain MSE of T_{m1}, Square both sides of Eq. (3.7) and take expectation of both sides

$$E(T_{m1}-S_y^2)^2=S_y^4E\left[e_0+(1-\alpha)\left(-\frac{e_1}{2}-\frac{e_0e_1}{2}+\frac{3}{8}e_1^2\right)\right]^2 \tag{3.9}$$

$$MSE(T_{m1})=S_y^4\left[\begin{array}{l}\left(\gamma\beta^{*}_{2(1y)}+\sum_{i=1}^{N}\gamma_{2i}\beta^{*}_{2(2yi)}\right)\\-(1-\alpha)\left(\gamma\gamma_{2(1)}+\sum_{i=1}^{N}\gamma_{2}\gamma^{*}_{22(2i)}\right)\\+\frac{(1-\alpha)^2}{4}\left(\gamma\beta^{*}_{2(1x)}+\sum_{i=1}^{N}\gamma_{2i}\beta^{*}_{2(2xi)}\right)\end{array}\right] \tag{3.10}$$

To obtain optimality value for α, we differentiate Eq. (3.10) partially with respect to α and equate to zero.

Let $A=\gamma\beta^{*}_{2(1y)}$, $C=\gamma\gamma^{*}_{22(1)}$, $D=\sum_{i=1}^{N}\gamma_{2i}\gamma^{*}_{22(2i)}$,

$E=\gamma\beta^{*}_{2(1x)}$, $F=\sum_{i=1}^{N}\gamma_{2i}\beta^{*}_{2(2xi)}$, $G=\sum_{i=1}^{N}\gamma_{2i}\beta^{*}_{2(2yi)}$

$$\frac{\partial MSE(T_{m1})}{\partial\alpha}=S_y^4\left[\left(C+D\right)-\frac{1}{2}(1-\alpha)(E+F)\right]=0 \tag{3.11}$$

$$\alpha=1-\frac{2(C+D)}{E+F} \tag{3.12}$$

Substituting Eq. (3.12) result into Eq. (3.10), gives minimum MSE of T_{m1}

$$\text{MSE}(T_{m1})_{\min}=S_y^4\left[(A+G)-\frac{(C+D)^2}{E+F}\right] \tag{3.13}$$

Following and adopting the procedures above, the biases and MSEs of T_{m2}, T_{m3}, T_{m4}, and T_{m5}

$$\text{Bias}(T_{m2})=S_y^2(1-\beta)\left[\frac{1}{2}\gamma\gamma^{*}_{22(1)}-\frac{1}{8}\gamma\beta^{*}_{2(1x)}+\frac{1}{2}\sum_{i=1}^{N}\gamma_{2i}\gamma^{*}_{22(2i)}-\frac{1}{8}\sum_{i=1}^{N}\gamma_{2i}\beta^{*}_{2(2xi)}\right] \tag{3.14}$$

$$\mathrm{MSE}(T_{m2}) = S_y^4 \begin{bmatrix} \left(\gamma\beta^*_{2(1y)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2yi)}\right) \\ +\left(1-\beta\right)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)}\right) \\ +\left(\frac{\left(1-\beta\right)^2}{4}\right)\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right) \end{bmatrix} \tag{3.15}$$

$$\beta = 1 + \frac{2(C+D)}{E+F} \tag{3.16}$$

Substituting the Eq. (3.16) into Eq. (3.15), we have minimum MSE of T_{m2}

$$\mathrm{MSE}(T_{m2})_{\min} = S_y^4\left[(A+G) - \frac{(C+D)^2}{E+F}\right] \tag{3.17}$$

$$T_{m3} = \lambda\left(S_y^2 + e_0S_y^2\right)\exp\left[\frac{S_x^2-(S_x^2+e_1S_x^2)}{S_x^2+(S_x^2+e_1S_x^2)}\right] + (1-\lambda)\left(S_y^2+e_0S_y^2\right)\exp\left[\frac{S_x^2+e_1S_x^2-S_x^2}{S_x^2+e_1S_x^2+S_x^2}\right] \tag{3.18}$$

$$\mathrm{Bias}(T_{m3}) = S_y^2\left(\left(\frac{1}{2}-\lambda\right)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)}\right) + \frac{1}{2}\left(\lambda-\frac{1}{4}\right)\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right)\right) \tag{3.19}$$

$$\mathrm{MSE}(T_{m3}) = S_y^4 \begin{bmatrix} \left(\gamma\beta^*_{2(1y)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2yi)}\right) \\ +\left(1-2\lambda\right)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)}\right) \\ +\left(\frac{1}{2}-\lambda\right)^2\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right) \end{bmatrix} \tag{3.20}$$

$$\lambda = \frac{1}{2} + \frac{(C+D)}{(E+F)} \tag{3.21}$$

Substituting Eq. (3.21) into Eq. (3.20), we have minimum MSE of $T_{m3\,\mathrm{w}}$

$$\mathrm{MSE}(T_{m3})_{\min} = S_y^4\left[(A+G) - \frac{(C+D)^2}{E+F}\right] \tag{3.22}$$

$$T_{m4} = \frac{\phi\left(S_y^2+e_0S_y^2\right)}{\left(S_x^2+e_1S_x^2\right)}S_x^2 + (1-\phi)\left(S_y^2+e_0S_y^2\right)\exp\left[\frac{S_x^2-\left(S_x^2+e_1S_x^2\right)}{S_x^2+\left(S_x^2+e_1S_x^2\right)}\right] \tag{3.23}$$

$$\mathrm{Bias}(T_{m4}) = S_y^2\left[-\frac{1}{2}(\phi+1)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)}\right) + \frac{1}{8}(5\phi+3)\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right)\right] \tag{3.24}$$

$$\mathrm{MSE}(T_{m4}) = S_y^4 \begin{bmatrix} \left(\gamma\beta^*_{2(1y)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2yi)}\right) \\ -\left(\phi+1\right)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)}\right) \\ +\frac{1}{4}\left(\phi+1\right)^2\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right) \end{bmatrix} \tag{3.25}$$

$$\phi = \frac{2\left(C+D\right)}{E+F} - 1 \tag{3.26}$$

Substituting Eq. (3.26) into Eq. (3.25) to get the minimum MSE of T_{m4}

$$\mathrm{MSE}(T_{m4})_{\min} = S_y^4\left[(A+G) - \frac{(C+D)^2}{E+F}\right] \tag{3.27}$$

$$T_{m5} = \frac{\theta\left(S_y^2+e_0S_y^2\right)\left(S_x^2+e_1S_x^2\right)}{S_x^2} + (1-\theta)\left(S_y^2+e_0S_y^2\right)\exp\left[\frac{S_x^2+e_1S_x^2-S_x^2}{S_x^2+e_1S_x^2+S_x^2}\right] \tag{3.28}$$

$$\mathrm{Bias}(T_{m5}) = S_y^2\left[\frac{1}{2}(\theta+1)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N}\gamma_{2i}\gamma^*_{22(2i)}\right) + \frac{1}{8}(\theta-1)\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N}\gamma_{2i}\beta^*_{2(2xi)}\right)\right] \tag{3.29}$$

$$\mathrm{MSE}(T_{m5}) = S_y^4 \begin{bmatrix} \left(\gamma\beta^*_{2(1y)} + \sum_{i=1}^{N} \gamma_{2i}\beta^*_{2(2yi)} \right) \\ +(\theta+1)\left(\gamma\gamma^*_{22(1)} + \sum_{i=1}^{N} \gamma_{2i}\gamma^*_{22(2i)} \right) \\ +\frac{1}{4}(\theta+1)^2\left(\gamma\beta^*_{2(1x)} + \sum_{i=1}^{N} \gamma_{2i}\beta^*_{2(2xi)} \right) \end{bmatrix} \quad (3.30)$$

$$\mathrm{MSE}(T_{m5}) = S_y^4\left[(A+G)+(\theta+1)(C+D) + \frac{1}{4}(\theta+1)^2(E+F) \right] \quad (3.31)$$

$$\theta = -\frac{2(C+D)}{E+F} - 1 \quad (3.32)$$

Substituting Eq. (3.32) result into Eq. (3.31), we have minimum MSE of T_{m5}

$$\mathrm{MSE}(T_{m5})_{\min} = S_y^4\left[(A+G) - \frac{(C+D)^2}{E+F} \right] \quad (3.33)$$

4. Empirical Study

We applied Muhammed et al. (2014) estimator (case II) to estimate the variance of the finite population variance under two-stage sampling and the modified estimators on three sets natural data.

The descriptions of the population are given below:

POPULATION I: {Source: Government Technical College, Farufaru: SS1 Classes}

The population I consist of four clusters of unequal sizes and the relationship between study variable (*Y*) and auxiliary variable (*X*) is positive.

Y: Scores of SS1 students in Physics Examination, Third term (2017/2018) Session

X: Scores of SS1 Students in Mathematics Examination, Third term (2017/2018) Session.

POPULATION II: {Source: Government Technical College, Farufaru: SS2 Classes}

The population II consists of four clusters of unequal sizes and the relationship between study variable (*Y*) and auxiliary variable (*X*) is positive.

Y: Scores of SS2 students in Physics Examination, Third term (2017/2018) Session

X: Scores of SS2 Students in Mathematics Examination, Third term (2017/2018) Session

Clusters	M_i	m_i	$\beta^*_{2(2xi)}$	$\beta^*_{2(2yi)}$	γ_{2i}	$\gamma_{22(2i)}$	ρ^{*2}_{2i}	$\bar{X}_i$	$\bar{Y}_i$
	2	1	1.2544	1.1012	1.04436x10^{-5}	0.8698	0.5477	52.3605	42.2558
	1	1	2.2518	1.7589	0	1.8669	0.8800	44.8364	37.3455
	2	1	3.9687	5.0346	1.04436x10^{-5}	4.3661	0.9541	62.2667	66.3238
	2	1	1.2593	1.2811	1.04436x10^{-5}	1.0536	0.6881	42.4359	35.3718

$N = 324,\ n = 193,\ S_x^2 = 343.3438,\ S_y^2 = 501.3859,\ \rho_1^2 = 0.7993,\ \beta^*_{2(1x)} = 0.6411,$
$\beta^*_{2(1y)} = 0.9034, \gamma^*_{22(1)} = 0.6804,\ \bar{X} = 47.5648,\ \bar{Y} = 51.9043,\ \bar{M} = 1.75,$
$\gamma = 0.002095$

Clusters	M_i	m_i	$\beta^*_{2(2xi)}$	$\beta^*_{2(2yi)}$	γ_{2i}	$\gamma_{22(2i)}$	ρ^{*2}_{2i}	$\bar{X}_i$	$\bar{Y}_i$
	2	1	4.2285	4.3883	4.5295x10^{-5}	4.2898	0.9917	26.4151	28.9811
	1	1	8.7854	9.4278	0	9.0865	0.9968	21.5417	24.9583
	2	1	6.1065	5.7068	4.5295x10^{-5}	5.7686	0.9549	24.7647	27.7451
	2	1	3.2003	3.4323	4.5295x10^{-5}	3.1868	0.9273	23.9706	26.6471

$N = 162,\ n = 89,\ S_x^2 = 180.2256,\ S_y^2 = 180.5248,\ \rho_1^2 = 0.8954,\ \beta^*_{2(1x)} = 0.3854,\ \beta^*_{2(1y)} = 0.3654,$
$\gamma^*_{22(1)} = 0.3551,\ \bar{X} = 24.6605,\ \bar{Y} = 27.5062,\ \bar{M} = 1.75\ \gamma = 0.005063$

POPULATION III: {Source: Government Technical College, Farufaru: SS3 Classes}

The population III consists of four clusters of unequal sizes and the relationship between study variable (Y) and auxiliary variable (X) is positive.

Y: Scores of SS3 students in Physics Mock Examination, (2017/2018) Session

X: Scores of SS3 Students in Mathematics Mock Examination, (2017/2018) Session

Table 1 shows MSE and PRE of the usual variance estimator, Muhammed et al. (2014) case II estimator and the proposed estimators. The result of Table 1 shows that the proposed estimators have minMSE and higher PRE. This implies that the proposed estimators are more efficient than the existing estimators considered in the study.

Table 2 shows MSE and PRE of usual variance estimator Muhammed et al. (2014) case II estimator and the proposed estimators. The result of Table 2 shows that the proposed estimators have minMSE and higher PRE. This implies that the proposed estimators are more efficient than the existing estimators considered in the study.

Table 3 shows MSE and PRE of usual variance estimator Muhammed et al. (2014) case II estimator and the proposed estimators. The result of Table 3 shows that the proposed estimators have minMSE and higher PRE. This implies that the proposed estimators are more efficient than the existing estimators considered in the study.

Clusters	M_i	m_i	$\beta^*_{2(2xi)}$	$\beta^*_{2(2yi)}$	γ_{2i}	$\gamma_{22(2i)}$	$\rho^{*^2}_{2i}$	$\bar{X}_i$	$\bar{Y}_i$
	2	1	2.0793	1.3705	2.28791x10^{-5}	0.7682	0.2071	35.6393	45.0000
	1	1	1.1286	1.4267	0	1.0994	0.7523	38.4706	41.7941
	2	1	1.4278	1.7919	2.28791x10^{-5}	1.1262	0.4957	35.9385	38.0462
	2	1	1.8802	2.6914	2.28791x10^{-5}	1.8688	0.6901	34.5397	36.7143

$$N = 223,\ n = 128,\ S_x^2 = 151.5352,\ S_y^2 = 212.1145,\ \rho_1^2 = 0.7597,\ \beta^*_{2(1x)} = 0.5663,\ \beta^*_{2(1y)} = 0.5746,$$
$$\gamma^*_{22(1)} = 0.4972,\ \bar{X} = 35.8475,\ \bar{Y} = 40.1435,\ \bar{M} = 1.75\ \gamma = 0.003328$$

Table 1. MSE and PRE of proposed estimators, Muhammed et al. (2014) case II and Usual variance estimator.

Estimators	MSE	PRE
Usual variance ($\breve{S}^2_{0(2s)}$)	495.2547	100
Muhammed et al (2014) case II ($\breve{S}^{*^2}_{p(2s)}$)	328.6935	150.6737
Proposed estimators (Tm)	99.12242	499.6394

Table 2. MSE and PRE of proposed estimators, Muhammed et al (2014) case II and Usual variance estimator.

Estimators	MSE	PRE
Usual variance ($\breve{S}^2_{0(2s)}$	80.25877	100
Muhammed et al (2014) case II ($\breve{S}^{*^2}_{p(2s)}$)	11.24403	713.7901
Proposed estimators (Tm)	7.191233	1116.064

Table 3. MSE and PRE of proposed estimators, Muhammed et al (2014) case II and Usual variance estimator.

Estimators	MSE	PRE
Usual variance ($\breve{S}^2_{0(2s)}$)	92.06371	100
Muhammed et al (2014) case II ($\breve{S}^{*^2}_{p(2s)}$)	69.07121	133.2881
Proposed estimators (Tm)	24.16119	381.0396

5. Conclusion

From the empirical study, the results show that the modified estimators are more efficient than the existing estimator considered in the study.

References

Cochran WG. Sampling techniques. 3rd edition, Wiley, New York, 1977.

Das AK, Tripathi TP. Use of auxiliary information in estimating the finite population variance. Sankhya C 1978; 40:139–48.

Godambe VP. On two stage sampling. J Roy Stat Soc B 1951; 13:216–8.

Gupta S, Shabbir J. On improvement in variance estimation using auxiliary information. Commun Stat Theory 2007; 36(12):2177–85.

Gupta S, Shabbir J. Variance estimation in Simple random sampling using auxiliary information. Haccet J Math Stat 2008; 37(1):57–67.

Gupta S, Shabbir J. Some estimators of finite population variance of stratified simple mean. Commun Stat Theory 2010; 39(16):3001–8.

Isaki CT. Variance estimation using auxiliary information. J Am Stat Assoc 1983; 78:117–23.

Kadilar C, Cingi H. Improvement in variance estimation using auxiliary information. Haccet J Math Stat 2006; 35(1):111–5.

Kalton G. Introduction to survey sampling. Sage, Thousand Oaks, CA, 1983.

Kaur P. An efficient regression type estimator in survey sampling. Biometrical J 1989; 27(1):107–10.

Mahalanobis PC. The sample census of the area under jute in Bengal in 1940. Sankhya Ser B 1967; 29(2):81–182.

Misra P, Karan RS. Estimation of population variance using a Generalized Double Sampling Estimator. Sri Lankan J Appl Stat 2014; (15-3).

Muhammed J, Shabbir J, Zaheer A, Zainab R. Ratio type exponential estimator for the estimation of finite population variance under two-stage sampling. Res J Appl Sci Eng Technol 2014; 7(19):4095–9.

Olufadi Y, Kadilar C. A study on the chain ratio estimator for finite population variance. J Probab Stat 2014; 2(5).

Sahai A, Ray SK. Efficient families of ratio and product type estimators. Biometrika 1980; 67(1):211–5.

Saini M. A class of predictive estimators in two-stage sampling when auxiliary character is estimated at SSU level. Int J Pure Appl Math 2013; 85(2):285–95.

Sarndal CE, Swensson B, Wretman J. Model assisted survey sampling. Springer-Verlag, New York, p 652, 1992.

Saxena BC, Narain P, Srivastava AK. Multiple frame surveys in two stage sampling. Indian J Stat 1984; 46(1):75–82.

Singh HP, Vishwakarma GK. A family of estimators of population mean using auxiliary information in stratified sampling. Commun Stat Theory 2008; 37(7):1038–50.

Singh R, Chauhan P, Sawan N, Smarandache F. Improvement exponential estimator for population variance using two auxiliary variables. Italian J Pure Appl Math 2011; 28:1–108.

Singh R, Vishwakarma GK, Gupta PC, Pareek S. An alternative approach to estimation of population mean in to stage sampling. Math Theory Model 2013; 3(13):48–54.

Upadhyaya LN, Singh HP, Singh S. A class of estimators for estimating the variance of the ratio estimator. J Jpn Stat Soc 2004; 34(1):47–63.Upadhyaya L, Singh H, Chatterjee S, Yadav R. Improved ratio and product exponential type estimators. J Stat Theory Pract 2011; 5(2):285–302.Wolter KM. Introduction to variance estimation. Springer-Verlag, New York, 1985.

Yadav SK, Kadilar C. Improved exponential type ratio estimator of population variance. Revista Colombiana de Estadistica 2013; 36(1):145–52.

Yunusa O. On the estimation of ratio-cum-product estimators using two-stage sampling. Stat Trans New Ser 2010; 11(2):253–65.