# Improvement in estimation of population variance utilising known auxiliary parameters for a decision-making model

Dinesh K. Sharma

Department of Business, Management and Accounting,
University of Maryland Eastern Shore,
Princess Anne, MD, USA
Email: profdksharma@gmail.com

S.K. Yadav*

Department of Statistics,
Babasaheb Bhimrao Ambedkar University,
Lucknow, India
Email: drskystats@gmail.com
*Corresponding author

Hari Sharma

Department of Accounting and Finance,
Virginia State University,
VA, USA
Email: hsharma@vsu.edu

**Abstract:** There is variation in similar things, whether natural or artificial. It is therefore in our best interest to estimate this variation. In this article, we suggest a Searls ratio type estimator for the main variable using the available information on the tri-mean and the third quartile of the auxiliary variable for an enhanced population variance estimation. The bias and mean squared error (MSE) of the proposed estimator are derived up to the first-degree approximation. The optimal value of the characterising scalar is obtained and, for this optimal value, the least MSE is achieved. The suggested estimator is compared with the competing estimators based on their MSEs, both theoretically and empirically. The calculation of biases and MSEs of suggested and competing estimators are accomplished by using R programming. The study's outcome is evidenced in the least MSE of the proposed model compared to competing estimators used in the study for business decision making.

**Keywords:** population variance; Searls type estimator; auxiliary variable; bias; mean squared error; MSE; percentage relative efficiency; PRE.

**Reference** to this paper should be made as follows: Sharma, D.K., Yadav, S.K. and Sharma, H. (2022) 'Improvement in estimation of population variance utilising known auxiliary parameters for a decision-making model', *Int. J. Mathematical Modelling and Numerical Optimisation*, Vol. 12, No. 1, pp.15–28.

**Biographical notes:** Dinesh K. Sharma is a Professor of Quantitative Methods and Computer Applications in the Department of Business, Management and Accounting at the University of Maryland Eastern Shore. He earned his MS in Mathematics, MS in Computer Science, PhD in Operations Research, and a second PhD in Management. He has over 26 years of teaching experience and served on several committees to supervise PhD students. He has published over 200 refereed journal articles and conference proceedings and has won numerous best paper awards. He is an accomplished grant writer and he has collaborated on a number of funded research/education grants. He is the Editor-in-Chief of the *Journal of Global Information Technology* and *Review of Business and Technology Research*.

S.K. Yadav is a Faculty in the Department of Statistics at Babasaheb Bhimrao Ambedkar University Lucknow, U.P., India. He earned his MSc and PhD in Statistics from Lucknow University and qualified the National Eligibility Test. He has published more than 50 research papers in national and international journals of repute indexed in Scopus/Web of Science and two books from an international publisher. He is a referee for 20 reputed international journals. He has been awarded Young Scientist Award in 2016 for the contribution in the field survey sampling by Venus International Research Foundation, Chennai, India and best paper award in 2018 by MTMI, USA. He has presented papers in more than 20 national and international conferences and also delivered 22 invited talks in several conferences.

Hari Sharma is a Professor of Accounting and Finance and Chair of the Department of Accounting and Finance. He is an experienced professional with a unique combination of more than 35 years of academic and professional experience in the area of quantitative finance including designing of accounting and management information systems for senior management reporting. He has numerous refereed articles, proceedings, and presentations. He has also developed many curricula, written grants, spoken at many conferences, and received many awards for his outstanding contributions for teaching, research and administrative assignments.

The paper is a revised and expanded version of a paper 'Application of auxiliary variable parameters for enhanced estimation of population variance for business decisions', presented at the 2019 Annual Meeting of the Decision Sciences Institute, New Orleans, Louisiana, 23–25 November 2019.

## 1 Introduction

Population variance is one of the most important dispersion measures and plays a significant role in our day-to-day business decisions. The variation is natural and can be easily seen. The precise estimation of the parameters is also well established in the literature. It is beneficial for large populations to minimise errors and ultimately save time and money in planning and decision-making. Estimation is the need for rapid policy decision-making. Teplicka (2015) suggested that the proper use of variance analysis could be very beneficial to the management decisions of the company. In this work, Teplicka (2015) has shown that the analysis of variances is an excellent tool for assessing price variations for production factors in manufacturing products. Conine and McDonald (2017) worked on the application of variance analysis in financial planning and analysis to organisations through their survey evidence study and made some recommendations to enhance organisations. Cohen and Pant (2018) suggested that the effective use of variance analysis is a powerful tool for any organisation to achieve its long-term objectives. Suppose there is a significant variance in the accounting system of the organisation. In that case, it needs to know the impact of the variation on the financial results and how this variance can be managed to achieve its objectives smoothly.

The variance is primarily estimated by using the sample variance as a representative of the desirable properties of a good estimator. One of the significant drawbacks of this method is that the sampling variance could be fairly large. It is therefore necessary to search for an estimator with a sampling distribution that is closely dispersed around the variance of the population. Thus, the auxiliary information is required to fulfil this purpose. Auxiliary information is obtained from the auxiliary variable ($X$) having a high degree of correlation with the main variable ($Y$). The ratio estimators are being used for the elevated population variance ($S_y^2$) estimation when $X$ and $Y$ are having a high degree of positive correlation, and the regression line passes by origin. Product type estimators are used for enhanced estimation of $S_y^2$ when $X$ and $Y$ have a high degree of negative correlation, and regression line passes by origin. In either case, the regression type estimators are applied for an improved $S_y^2$ estimation of $Y$ using the known $X$.

Singh and Singh (2001) articulated a ratio-type estimator for an elevated $S_y^2$ estimation using $X$. Later, Singh and Singh (2003) proposed an enhanced estimation of $S_y^2$ through a regression approach in a double sampling regime. In addition, under the sub-sampling system, Jhajj et al. (2005) proposed an important family of chain estimators for the elevated estimation of $S_y^2$. Additionally, using auxiliary parameters, Shabbir and Gupta (2007) focused on improving variance estimation. Then, under the simple random sampling scheme, Kadilar and Cingi (2007) suggested some progress in variance estimation. Singh et al. (2008) suggested an almost unbiased ratio and a finite $S_y^2$ estimator of the product type using kurtosis knowledge of $X$ in the sample surveys.

Grover (2010) described the correction note for the improved $S_y^2$ estimation using auxiliary parameters. A novel technique for $S_y^2$ estimation in simple random sampling using $X$ was also given by Singh and Solanki (2012). On the other hand, a two-parameter elevated variance estimator using auxiliary parameters was proposed by Yadav and Kadilar (2014). Singh and Pal (2016) suggested an improved class of estimators of $S_y^2$ using quartiles of $X$. Yadav et al. (2017) proposed an enhanced estimator of $S_y^2$ using the known tri-mean and interquartile range of the $X$. Recently, Yadav et al. (2019) proposed an elevated population variance estimator that utilises the known tri-mean and third quartile of the auxiliary variable. Gulzar et al. (2020) worked on an elevated population variance estimate using non-traditional auxiliary variable measures. Naz et al. (2020) used non-conventional dispersion measures of $X$, while having a high correlation with $Y$ under consideration, and suggested ratio-type estimators of $S_y^2$ when outliers were present. Oghenekevwe et al. (2020) demonstrated the distribution effect of various groups of population variance estimators on efficiency using a known auxiliary variable under a simple random sampling scheme. Shahzad et al. (2021) worked on a new class of calibration variance estimators based on L moments.

This study is motivated by Searls (1964) and Yadav et al. (2019). In this study, we suggest a Searls type estimator for $S_y^2$, and apply a known population tri-mean and third quartile of *X* for improving $S_y^2$ estimation of *Y*. The sampling properties are bias and the mean squared error (MSE) is studied until the order one is approximated. The remainder of the paper was presented in different sections. Section 2 presents a review of the population variance estimators of the study variable using known auxiliary variable parameters. Section 3 describes the suggested estimators and their sampling properties up to the first order approximation. Section 4 explains the efficiency comparison of the introduced estimator with the competing estimators, along with the efficiency conditions of the proposed estimator over the competing estimators. The empirical analysis in which the biases and MSEs for the real natural population were measured is represented in Section 5. The results obtained from the numerical analysis findings are discussed in Section 6. Section 7 provides the conclusion concerning the study's findings.

## 2 Literature review

In Table 1, with their biases and MSEs, some of the important estimators in the literature are shown. We define the following notations used throughout the manuscript, before giving various estimators of $S_y^2$ in Table 1.

| | |
|---|---|
| $N$ | population size |
| $n$ | sample size |
| $f = n / N$ | fraction of sampling |
| $Y$ | main variable |
| $X$ | secondary variable |
| $\bar{Y}$ | population mean of $Y$ |
| $\bar{X}$ | population mean of $X$ |
| $S_y$ | population standard deviation of $Y$ |
| $S_x$ | population standard deviation of $X$ |
| $\rho$ | population correlation coefficient between $Y$ and $X$ |
| $S_{yx}$ | population covariance of $Y$ and $X$ |
| $C_y$ | population coefficient of variation of $Y$ |
| $C_x$ | population coefficient of variation of $X$ |
| $Q_i$ | population $i^{\text{th}}$ quartile of $X$ |
| $M_d$ | population median of $X$ |
| $\beta_1$ | population coefficient of skewness of $X$ |
| $\beta_2$ | population coefficient of kurtosis of $X$ |
| $Qr = Q_3 - Q_1$ | population quartile range of $X$ |

$Q_d = \dfrac{Q_3 - Q_1}{2}$ population quartile deviation of $X$

$Q_a = \dfrac{Q_3 + Q_1}{2}$ population quartile average of $X$

$TM = \dfrac{Q_1 + 2Q_2 + Q_3}{4}$ population tri-mean of $X$.

**Table 1** Various estimators of $S_y^2$, their biases and MSEs

| *S.N.* | *Estimator* | *Bias* | *MSE* |
|---|---|---|---|
| 1 | $t_0 = s_y^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2$ Sample variance | --- | $\gamma S_y^4(\lambda_{40} - 1)$ |
| 2 | $t_r = s_y^2\left[\dfrac{S_x^2}{s_x^2}\right]$ Isaki (1983) | $\gamma S_y^2[(\lambda_{40} - 1) - (\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + (\lambda_{04} - 1)$ $-2(\lambda_{22} - 1)]$ |
| 3 | $t_1 = s_y^2\left[\dfrac{S_x^2 + \beta_2}{s_x^2 + \beta_2}\right]$ Upadhyaya and Singh (1999) | $\gamma S_y^2 R_1[R_1(\lambda_{40} - 1) - (\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_1^2(\lambda_{04} - 1)$ $-2R_1(\lambda_{22} - 1)]$ |
| 4 | $t_2 = s_y^2\left[\dfrac{S_x^2 + C_x}{s_x^2 + C_x}\right]$ Kadilar and Cingi (2006) | $\gamma S_y^2 R_2[R_2(\lambda_{40} - 1)$ $-(\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_2^2(\lambda_{04} - 1)$ $-2R_2(\lambda_{22} - 1)]$ |
| 5 | $t_3 = s_y^2\left[\dfrac{S_x^2\beta_2 + C_x}{s_x^2\beta_2 + C_x}\right]$ Kadilar and Cingi (2006) | $\gamma S_y^2 R_3[R_3(\lambda_{40} - 1)$ $-(\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_3^2(\lambda_{04} - 1)$ $-2R_3(\lambda_{22} - 1)]$ |
| 6 | $t_4 = s_y^2\left[\dfrac{S_x^2 C_x + \beta_2}{s_x^2 C_x + \beta_2}\right]$ Kadilar and Cingi (2006) | $\gamma S_y^2 R_4[R_4(\lambda_{40} - 1)$ $-(\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_4^2(\lambda_{04} - 1)$ $-2R_4(\lambda_{22} - 1)]$ |
| 7 | $t_5 = s_y^2\left[\dfrac{S_x^2 + M_d}{s_x^2 + M_d}\right]$ Subramani and Kumarpandiyan (2012a) | $\gamma S_y^2 R_5[R_5(\lambda_{40} - 1)$ $-(\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_5^2(\lambda_{04} - 1)$ $-2R_5(\lambda_{22} - 1)]$ |
| 8 | $t_6 = s_y^2\left[\dfrac{S_x^2 + Q_1}{s_x^2 + Q_1}\right]$ Subramani and Kumarpandiyan (2012b) | $\gamma S_y^2 R_6[R_6(\lambda_{40} - 1)$ $-(\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_6^2(\lambda_{04} - 1)$ $-2R_6(\lambda_{22} - 1)]$ |
| 9 | $t_7 = s_y^2\left[\dfrac{S_x^2 + Q_3}{s_x^2 + Q_3}\right]$ Subramani and Kumarpandiyan (2012b) | $\gamma S_y^2 R_7[R_7(\lambda_{40} - 1)$ $-(\lambda_{22} - 1)]$ | $\gamma S_y^4[(\lambda_{40} - 1) + R_7^2(\lambda_{04} - 1)$ $-2R_7(\lambda_{22} - 1)]$ |

**Table 1** Various estimators of $S_y^2$, their biases and MSEs (continued)

| *S.N.* | *Estimator* | *Bias* | *MSE* |
|---|---|---|---|
| 10 | $t_8 = s_y^2\left[\dfrac{S_x^2 + Q_r}{s_x^2 + Q_r}\right]$ Subramani and Kumarpandiyan (2012b) | $\gamma S_y^2 R_8[R_8(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_8^2(\lambda_{04}-1)$ $-2R_8(\lambda_{22}-1)]$ |
| 11 | $t_9 = s_y^2\left[\dfrac{S_x^2 + Q_d}{s_x^2 + Q_d}\right]$ Subramani and Kumarpandiyan (2012b) | $\gamma S_y^2 R_9[R_9(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_9^2(\lambda_{04}-1)$ $-2R_9(\lambda_{22}-1)]$ |
| 12 | $t_{10} = s_y^2\left[\dfrac{S_x^2 + Q_a}{s_x^2 + Q_a}\right]$ Subramani and Kumarpandiyan (2012b) | $\gamma S_y^2 R_{10}[R_{10}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{10}^2(\lambda_{04}-1)$ $-2R_{10}(\lambda_{22}-1)]$ |
| 13 | $t_{11} = s_y^2\left[\dfrac{S_x^2 C_x + M_d}{s_x^2 C_x + M_d}\right]$ Subramani and Kumarpandiyan (2013) | $\gamma S_y^2 R_{11}[R_{11}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{11}^2(\lambda_{04}-1)$ $-2R_{11}(\lambda_{22}-1)]$ |
| 14 | $t_{12} = s_y^2\left[\dfrac{S_x^2 \rho + Q_3}{s_x^2 \rho + Q_3}\right]$ Khan and Shabbir (2013) | $\gamma S_y^2 R_{12}[R_{12}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{12}^2(\lambda_{04}-1)$ $-2R_{12}(\lambda_{22}-1)]$ |
| 15 | $t_{13} = s_y^2\left[\dfrac{S_x^2 + (TM + Q_a)}{s_x^2 + (TM + Q_a)}\right]$ Maqbool and Javaid (2017) | $\gamma S_y^2 R_{13}[R_{13}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{13}^2(\lambda_{04}-1)$ $-2R_{13}(\lambda_{22}-1)]$ |
| 16 | $t_{14} = s_y^2\left[\dfrac{S_x^2 + C_x S_x}{s_x^2 + C_x S_x}\right]$ Khalil et al. (2018) | $\gamma S_y^2 R_{14}[R_{14}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{14}^2(\lambda_{04}-1)$ $-2R_{14}(\lambda_{22}-1)]$ |
| 17 | $t_{15} = s_y^2\left[\dfrac{S_x^2 + C_x \bar{X}}{s_x^2 + C_x \bar{X}}\right]$ Khalil et al. (2018) | $\gamma S_y^2 R_{15}[R_{15}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{15}^2(\lambda_{04}-1)$ $-2R_{15}(\lambda_{22}-1)]$ |
| 18 | $t_{16} = s_y^2\left[\dfrac{S_x^2 + C_x M_d}{s_x^2 + C_x M_d}\right]$ Khalil et al. (2018) | $\gamma S_y^2 R_{16}[R_{16}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{16}^2(\lambda_{04}-1)$ $-2R_{16}(\lambda_{22}-1)]$ |
| 19 | $t_{17} = s_y^2\left[\dfrac{S_x^2 + (TM + Q_3)}{s_x^2 + (TM + Q_3)}\right]$ Yadav et al. (2019) | $\gamma S_y^2 R_{17}[R_{17}(\lambda_{40}-1)$ $-(\lambda_{22}-1)]$ | $\gamma S_y^4[(\lambda_{40}-1)+R_{17}^2(\lambda_{04}-1)$ $-2R_{17}(\lambda_{22}-1)]$ |

In the general sense, the bias and the MSE of the various estimators in Table 1 may be expressed as

$$B(t_i) = \gamma S_y^2 R_i\left[R_i(\lambda_{40}-1) - (\lambda_{22}-1)\right], \quad i = 1, 2, ..., 15$$

$$MSE(t_i) = \gamma\left[(\lambda_{40} - 1) + R_i^2(\lambda_{04} - 1) - 2R_i(\lambda_{22} - 1)\right] \quad i = 1, 2, ..., 15 \tag{1}$$

where

$$\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}}, \quad \mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^r\left(X_i - \bar{X}\right)^s, \quad \gamma = \frac{(1-f)}{n}, \quad f = \frac{n}{N},$$

$$R_1 = \frac{S_x^2}{S_x^2 + \beta_2}, \quad R_2 = \frac{S_x^2}{S_x^2 + C_x}, \quad R_3 = \frac{S_x^2\beta_2}{S_x^2\beta_2 + C_x}, \quad R_4 = \frac{S_x^2 C_x}{S_x^2 C_x + \beta_2},$$

$$R_5 = \frac{S_x^2}{S_x^2 + M_d}, \quad R_6 = \frac{S_x^2}{S_x^2 + Q_1}, \quad R_7 = \frac{S_x^2}{S_x^2 + Q_3}, \quad R_8 = \frac{S_x^2}{S_x^2 + Q_r}, \quad R_9 = \frac{S_x^2}{S_x^2 + Q_d},$$

$$R_{10} = \frac{S_x^2}{S_x^2 + Q_a}, \quad R_{11} = \frac{S_x^2 C_x}{S_x^2 C_x + M_d}, \quad R_{12} = \frac{S_x^2\rho}{S_x^2\rho + Q_3}, \quad R_{13} = \frac{S_x^2}{S_x^2 + (TM + Q_a)},$$

$$R_{14} = \frac{S_x^2}{S_x^2 + C_x S_x}, \quad R_{15} = \frac{S_x^2}{S_x^2 + C_x\bar{X}}, \quad R_{16} = \frac{S_x^2}{S_x^2 + C_x M_d}, \quad R_{17} = \frac{S_x^2}{S_x^2 + (TM + Q_3)}.$$

## 3 Proposed estimator

This study is motivated by Searls (1964) who has shown that an improved estimator may be obtained by taking some constant multiple of sample mean as an estimator of $\bar{Y}$ and this constant is obtained by minimising the MSE of the suggested estimator and Yadav et al. (2019), and to improve estimation, we suggest the following estimator as,

$$t_p = \kappa\, s_y^2\left[\frac{S_x^2 + (TM + Q_3)}{s_x^2 + (TM + Q_3)}\right] \tag{2}$$

where $\kappa$ is a characterising scalar, which is so obtained that the MSE of $t_p$ is a least.

To study the bias and MSE of the suggested estimator, we have some assumptions given as, $s_y^2 = S_y^2(1 + \varepsilon_0)$ and $s_x^2 = S_x^2(1 + \varepsilon_1)$ such that $E(\varepsilon_i) = 0$ for ($i$ = 0, 1) and $E(\varepsilon_0^2) = \gamma(\lambda_{40} - 1), E(\varepsilon_1^2) = \gamma(\lambda_{04} - 1),\ E(\varepsilon_0\varepsilon_1) = \gamma(\lambda_{22} - 1)$.

The proposed estimator in (2) may be expressed in terms of $\varepsilon_i's$ as,

$$t_p = \kappa S_y^2(1 + e_0)(1 + R_{17}e_1)^{-1}$$

Expanding the term in above equation, simplifying and having the terms till the approximation of order one, we have,

$$t_p = \kappa S_y^2\left(1 + e_0 - R_{17}e_1 - R_{17}e_0e_1 + R_{17}^2e_1^2\right)$$

Subtracting $S_y^2$ on both the sides of the above equation, we have,

$$t_p - S_y^2 = \kappa S_y^2\left(1 + e_0 - R_{17}e_1 - R_{17}e_0e_1 + R_{17}^2e_1^2\right) - S_y^2 \tag{3}$$

We take the expectation of (3) for the bias of the proposed estimator and obtain the bias by placing different expectation values as,

$$B(t_p) = \gamma\kappa S_y^2 \left[ R_{17}^2 (\lambda_{04} - 1) - R_{17} (\lambda_{22} - 1) \right] + S_y^2 (\kappa - 1) \tag{4}$$

For the MSE of $t_p$, we square equation (3), simplifying and get the MSE by taking expectation and putting values of various expectations as,

$$MSE(t_p) = S_y^4 \begin{bmatrix} \left[ \kappa^2 \gamma (\lambda_{40} - 1) + (3\kappa^2 - 2\kappa) R_{17}^2 \gamma (\lambda_{04} - 1) \right. \\ \left. -2(2\kappa^2 - \kappa) R_{17} \gamma (\lambda_{22} - 1) + (\kappa - 1)^2 \right] \end{bmatrix} \tag{5}$$

The MSE of the suggested estimator is obtained for the optimum value of $\kappa$ as,

$$\kappa = \frac{A}{B} \tag{6}$$

where

$$A = 1 + R_{17}^2 \gamma (\lambda_{04} - 1) - R_{17} \gamma (\lambda_{22} - 1)$$

and

$$B = 1 + \gamma (\lambda_{40} - 1) + 3R_{17}^2 \gamma (\lambda_{04} - 1) - 4R_{17} \gamma (\lambda_{22} - 1)$$

The least value of the MSE of $t_p$, for the optimal value of $\kappa$ in (6), is:

$$MSE_{\min}(t_p) = S_y^4 \left[ 1 - \frac{A^2}{B} \right] \tag{7}$$

## 4 Efficiency comparison

Under this section, $t_p$ is compared theoretically with the existing estimators in competition and the efficiency conditions over these competing estimators are obtained.

The estimator $t_p$ performs better than the sample variance if,

$$V(t_0) - MSE_{\min}(t_p) = S_y^2 \left[ 1 - \frac{A^2}{B} - \gamma (\lambda_{40} - 1) \right] > 0 \text{ or } \frac{A^2}{B} + \gamma (\lambda_{40} - 1) < 1 \tag{8}$$

The proposed estimator performs better than Isaki (1983) estimator if,

$$\begin{aligned} & MSE(t_R) - MSE_{\min}(t_p) = S_y^2 \left[ 1 - \frac{A^2}{B} - \gamma \left\{ (\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1) \right\} \right] > 0 \\ & \text{or } \frac{A^2}{B} + \gamma \left\{ (\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1) \right\} < 1 \end{aligned} \tag{9}$$

The suggested estimator is better than the estimators in Table 1 under the conditions if,

$$MSE(t_i) - MSE_{\min}(t_p) = S_y^2 \left[ 1 - \frac{A^2}{B} - \gamma \left\{ (\lambda_{40} - 1) + R_i^2 (\lambda_{04} - 1) - 2R_i (\lambda_{22} - 1) \right\} \right] > 0, \tag{10}$$

$(i = 1, 2, \ldots, 17)$

## 5 Numerical study

Under this section, the efficiency conditions of $t_p$ over other competing estimators are verified. For this reason, we considered the population given in Yadav et al. (2019). This data set from Murthy (1967) on page 228 was taken by Yadav et al. (2019) with fixed capital as $X$ and the output of 80 factories as the $Y$. The biases and the MSEs of the introduced and the competing estimators have been numerically computed. Table 2 presents the parameters of this population.

**Table 2** Parameters of the population in Yadav et al. (2019)

$N = 80,\ n = 20,\ \bar{Y} = 51.8264,\ \bar{X} = 11.2646,\ \rho = 0.9413,\ S_y = 18.3549,\ C_y = 0.3542,$
$S_x = 8.4563,\ C_x = 0.7507,\ \lambda_{04} = 2.8664,\ \lambda_{40} = 2.2667,\ \lambda_{22} = 2.2209,\ Q_1 = 5.1500,$
$Q_3 = 16.975,\ Q_r = 11.825,\ Q_d = 5.9125,\ Q_a = 11.0625,\ T_m = 9.318,\ M_d = 7.575$

**Table 3** Bias, MSE of various estimators with the PRE of over the estimators in competition

| *Estimator* | *Bias* | *MSE* | *PRE* |
|---|---|---|---|
| Sample variance $t_0$ | 0 | 5,393.89 | 271.57 |
| Isaki (1983) estimator $t_r$ | 10.87 | 3,925.16 | 197.62 |
| Upadhyaya and Singh (1999) estimator $t_1$ | 9.29 | 3,658.41 | 184.19 |
| Kadilar and Cingi (2006) estimator $t_2$ | 10.44 | 3,850.16 | 193.84 |
| Kadilar and Cingi (2006) estimator $t_3$ | 10.72 | 3,898.56 | 196.28 |
| Kadilar and Cingi (2006) estimator $t_4$ | 8.81 | 3,580.83 | 180.28 |
| Subramani and Kumarpandiyan (2012a) estimator $t_5$ | –1.76 | 4,157.95 | 209.34 |
| Subramani and Kumarpandiyan (2012b) estimator $t_6$ | 8.17 | 3,480.55 | 175.23 |
| Subramani and Kumarpandiyan (2012b) estimator $t_7$ | 3.91 | 2,908.65 | 146.44 |
| Subramani and Kumarpandiyan (2012b) estimator $t_8$ | 5.50 | 3,098.41 | 156.00 |
| Subramani and Kumarpandiyan (2012b) estimator $t_9$ | 7.82 | 3,427.19 | 172.55 |
| Subramani and Kumarpandiyan (2012b) estimator $t_{10}$ | 5.77 | 3,133.33 | 157.75 |
| Subramani and Kumarpandiyan (2013) estimator $t_{11}$ | –0.94 | 2,467.88 | 124.25 |
| Khan and Shabbir (2013) estimator $t_{12}$ | 3.62 | 2,878.56 | 144.93 |
| Maqbool and Javaid (2017) estimator $t_{13}$ | 3.03 | 2,820.06 | 141.98 |
| Khalil et al. (2018) estimator $t_{14}$ | –0.67 | 2,547.21 | 128.24 |
| Khalil et al. (2018) estimator $t_{15}$ | –0.99 | 2,450.18 | 123.36 |
| Khalil et al. (2018) estimator $t_{16}$ | –0.56 | 2,580.75 | 129.93 |
| Yadav et al. (2019) estimator $t_{17}$ | 1.33 | 2,040.12 | 102.71 |
| Proposed estimator $t_p$ | –5.90 | 1,986.22 | 100.00 |

The biases and MSEs of $t_p$ and the estimators in competition along with the percentage relative efficiency (PRE) of $t_p$ over competing estimators of $S_y^2$ using auxiliary information for the above population are presented in Table 3.

The MSE of various estimators and the PRE of $t_p$ over the competing estimators are shown in Figures 1 and 2, respectively.

**Figure 1** MSE of various estimators (see online version for colours)

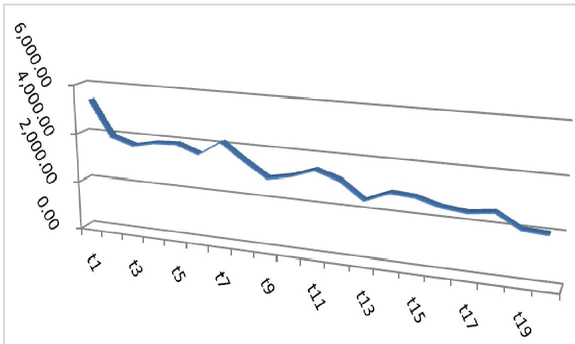
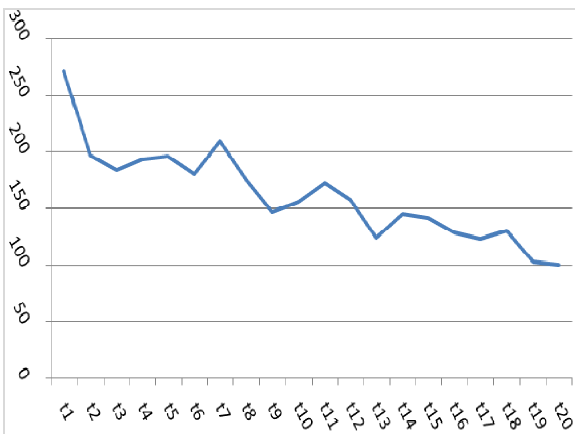



**Figure 2** PRE of proposed estimator over others (see online version for colours)




## 6 Simulation study

We have considered the parameter of the same natural population in the empirical analysis to produce a hypothetical population. The artificial population is simulated with its mean vector and variance-covariance matrix via bivariate normal distribution as:

$$\text{Means of } [Y, X] \text{ as } \mu = [51.8264, 11.2646]$$

Variances and covariance of $[Y, X]$ as $\sigma^2 = \begin{bmatrix} 336.9024 & 146.1035 \\ 146.1035 & 71.5090 \end{bmatrix}$ with correlation $\rho_{yx} = 0.9413$.

**Table 4** The PRE of proposed estimators over the estimators in competition

| *Estimator* | *PRE* |
|---|---|
| Sample variance $t_0$ | 270.94 |
| Isaki (1983) estimator $t_r$ | 198.02 |
| Upadhyaya and Singh (1999) estimator $t_1$ | 184.35 |
| Kadilar and Cingi (2006) estimator $t_2$ | 194.26 |
| Kadilar and Cingi (2006) estimator $t_3$ | 196.79 |
| Kadilar and Cingi (2006) estimator $t_4$ | 179.92 |
| Subramani and Kumarpandiyan (2012a) estimator $t_5$ | 210.03 |
| Subramani and Kumarpandiyan (2012b) estimator $t_6$ | 175.48 |
| Subramani and Kumarpandiyan (2012b) estimator $t_7$ | 147.88 |
| Subramani and Kumarpandiyan (2012b) estimator $t_8$ | 156.08 |
| Subramani and Kumarpandiyan (2012b) estimator $t_9$ | 173.04 |
| Subramani and Kumarpandiyan (2012b) estimator $t_{10}$ | 158.10 |
| Subramani and Kumarpandiyan (2013) estimator $t_{11}$ | 124.63 |
| Khan and Shabbir (2013) estimator $t_{12}$ | 145.18 |
| Maqbool and Javaid (2017) estimator $t_{13}$ | 142.06 |
| Khalil et al. (2018) estimator $t_{14}$ | 128.75 |
| Khalil et al. (2018) estimator $t_{15}$ | 123.86 |
| Khalil et al. (2018) estimator $t_{16}$ | 130.13 |
| Yadav et al. (2019) estimator $t_{17}$ | 103.12 |
| Proposed estimator $t_p$ | 100.00 |

The following steps have been used for the simulation:

a Bivariate normal distribution of $X$ and $Y$ of size $N = 5{,}000$ have been generated through these parameters using R Program.

b The parameters have been computed for this simulated population of size $N = 5{,}000$.

c A sample of size $n = 20$ has been selected from this simulated population.

d Sample statistics that is the sample mean, sample variance, and the values of the suggested and competing estimators $t_i$ of population variance are calculated for this sample.

e Steps c and d are repeated $m = 50{,}000$ times.

f The MSE of every estimator $t_i$ is calculated through the formula,

$$MSE(t_i) = \frac{1}{m}\sum_{j=1}^{m}\left(t_{ij} - \overline{T}_i\right)^2.$$

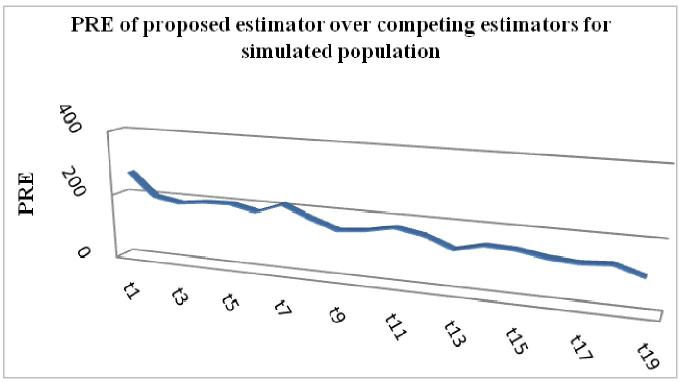
g The PRE of the suggested estimator $t_p$ over the competing estimators is given by,

$$PRE(t_p) = \frac{MSE(t_i)}{MSE(t_p)} \times 100, \quad i = 0, 1, \ldots, 17$$

The PRE of the proposed estimator $t_p$ over the competing estimators for the simulated population are presented in Table 4.

Figure 3 represents the PREs of the introduced estimator overt the competing estimators for the simulated population.

**Figure 3** PREs of $t_p$ over competing estimators (see online version for colours)




## 7 Results and discussion

From Table 3, it is established that the suggested estimator is having the smallest MSE among the competing estimators of the $S_y^2$ of the main variable under study. The MSEs of the competing estimators lie in the interval [2,040.12 5,393.89], while the MSE of the suggested estimator $t_p$ is 1,986.22. The PREs of the suggested estimators over the sample variance estimator is 271.57, whereas over the competing estimators, which make use of auxiliary information, ranges in interval [102.71 209.34]. Similar type of results are for the simulated population as well. Thus, we may observe that among the competing estimators that make use of along with the sample variance estimator, the estimator of Yadav et al. (2019) performs best, and the suggested estimator is better than Yadav et al. (2019) estimator of the population variance as it is having its MSE lesser than Yadav et al. (2019) for both the real and simulated populations.

## 8 Conclusions

In the study, we suggested an estimator for improved estimation of $S_y^2$ using Searls' technique for a simple random sampling without replacement (SRSWOR) scheme. We evaluated the expressions derived for the bias and MSE of $t_p$ to the first degree approximation. The theoretical, as well as empirical comparison of $t_p$, is carried out with the competing estimators of $S_y^2$. The computational algorithms have been developed for the biases and the MSEs for the suggested, as well as competing estimators in the R programming language. These R codes have been used for the natural population given in Yadav et al. (2019) to calculate the numerical values of biases and MSEs of the proposed and competing estimators. The results confirmed that the proposed estimator has the least MSE among the mentioned competing estimators of $S_y^2$. We expect that the use of proposed estimator ($t_p$) for enhance estimation of $S_y^2$ under a SRSWOR scheme will be beneficial in planning and making better business decisions. Thus, the purpose of searching for a more efficient estimator is accomplished by the study, and it may be applied in different areas of business decision making including life insurance, automobile, banking, marketing, and others.

## Acknowledgements

The authors are very much thankful to the editor and the anonymous referees for their valuable suggestions.

## References

Cohen, J.R. and Pant, L.W. (2018) *The Only Thing That Counts Is That Which Is Counted: A Discussion of Behavioral and Ethical Issues in Cost Accounting That Are Relevant for the OB Professor*, 18 September, DOI: 10.1.1.1026.5569&rep=rep1&type=pdf.

Conine, T.C. and McDonald, M. (2017) *The Application of Variance Analysis in FP&A Organizations: Survey Evidence and Recommendations for Enhancement*, 30 July, SSRN [online] hhttps://ssrn.com/abstract=3045928; http://dx.doi.org/10.2139/ssrn.3045928 (accessed 8 April 2021).

Grover, L.K. (2010) 'A correction note on improvement in variance estimation using auxiliary information', *Communications in Statistics Theory and Methods*, Vol. 39, No. 5, pp.753–764.

Gulzar, M.A., Abid, M., Nazir, H.Z., Zahid, F.M. and Riaz, M. (2020) 'On enhanced estimation of population variance using unconventional measures of an auxiliary variable', *Journal of Statistical Computation and Simulation*, Vol. 90, No. 12, pp.2180–2197.

Isaki, C.T. (1983) 'Variance estimation using auxiliary information', *Journal of American Statistical Association*, Vol. 78, No. 381, pp.117–123.

Jhajj, H.S., Sharma, M.K. and Grover, L.K. (2005) 'An efficient class of chain estimators of population variance under sub-sampling scheme', *Journal of Japan Statistical Society*, Vol. 35, No. 2, pp.273–286.

Kadilar, C. and Cingi, H. (2006) 'Improvement in variance estimation using auxiliary information', *Hacettepe Journal of mathematics and Statistics*, Vol. 35, No. 1, pp.111–115.

Kadilar, C. and Cingi, H. (2007) 'Improvement in variance estimation in simple random sampling', *Communications in Statistics – Theory and Methods*, Vol. 36, No. 11, pp.2075–2081.

Khalil, M., Ali, M., Shahzad, U., Hanif, M. and Jamal, N. (2018) 'Improved estimator of population variance using measure of dispersion of auxiliary variable', *Oriental Journal of Physical Sciences*, Vol. 3, No. 1, pp.33–39.

Khan, M. and Shabbir, J. (2013) 'A ratio type estimator for the estimation of population variance using quartiles of an auxiliary variable', *Journal of Statistics Applications & Probability*, Vol. 2, No. 3, pp.319–325.

Maqbool, S. and Javaid, S. (2017) 'Variance estimation using linear combination of tri-mean and quartile average', *American Journal of Biological and Environmental Statistics*, Vol. 3, No. 1, pp.5–9.

Murthy, M.N. (1967) *Sampling Theory and Methods, Calcutta Statistical Publishing Society*, Calcutta, India.

Naz, F., Nawaz, T., Pang, T. and Abid, M. (2020) 'Use of nonconventional dispersion measures to improve the efficiency of ratio-type estimators of variance in the presence of outliers', *Symmetry*, Vol. 12, No. 16, pp.1–26.

Oghenekevwe, E.H., Njideka, E.C. and Sylvia, O.C. (2020) 'Distribution effect on the efficiency of some classes of population variance estimators using information of an auxiliary variable under simple random sampling', *Science Journal of Applied Mathematics and Statistics*, Vol. 8, No. 1, pp.27–34.

Searls, D.T. (1964) 'The utilization of known coefficient of variation in the estimation procedure', *Journal of American Statistical Association*, Vol. 59, No. 308, pp.1125–1126.

Shabbir, J. and Gupta, S. (2007) 'On improvement in variance estimation using auxiliary information', *Communications in Statistics Theory and Methods*, Vol. 36, No. 12, pp.2177–2185.

Shahzad, U., Ahmad, I., Almanjahie, I.M., Noor, N.H.A. and Hanif, M. (2021) 'A new class of L-moments based calibration variance estimators', *Computers, Materials & Continua*, Vol. 66, No. 3, pp.3013–3028.

Singh, H.P. and Pal, S.K. (2016) 'An efficient class of estimators of finite population variance using quartiles', *Journal of Applied Statistics*, Vol. 43, No. 10, pp.1945–1958.

Singh, H.P. and Singh, R. (2001) 'Improved ratio-type estimator for variance using auxiliary information', *Journal of Indian Society of Agricultural Statistics*, Vol. 54, No. 3, pp.276–287.

Singh, H.P. and Singh, R. (2003) 'Estimation of variance through regression approach in two phase sampling', *Aligarh Journal of Statistics*, Vol. 23, pp.13–30.

Singh, H.P. and Solanki, R.S. (2012) 'A new procedure for variance estimation in simple random sampling using auxiliary information', *Statistical Papers*, DOI: 10.1007/s00362-012-0445-2.

Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2008) 'Almost unbiased ratio and product type estimator of finite population variance using the knowledge of kurtosis of an auxiliary variable in sample surveys', *Octogon Mathematical Journal*, Vol. 16, No. 1, pp.123–130.

Subramani, J. and Kumarapandiyan, G. (2012a) 'Variance estimation using median of the auxiliary variable', *International Journal of Probability and Statistics*, Vol. 1, No. 3, pp.36–40.

Subramani, J. and Kumarapandiyan, G. (2012b) 'Variance estimation using quartiles and their functions of an auxiliary variable', *International Journal of Statistics and Applications*, Vol. 2, No. 5, pp.67–72.

Subramani, J. and Kumarapandiyan, G. (2013) 'A new modified ratio estimator of population mean when median of the auxiliary variable is known', *Pakistan Journal of Statistics and Operation Research*, Vol. 9, No. 2, pp.137–145.

Teplicka, K. (2015) 'Utilization of variance analysis for managerial decision in the firm', *Comenius Management Review*, Vol. 9, No. 1, pp.5–11.

Upadhyaya, L.N. and Singh, H.P. (1999) 'Use of transformed auxiliary variable in estimating the finite population mean', *Biometrical Journal*, Vol. 41, No. 5, pp.627–636.

Yadav, S.K. and Kadilar, C. (2014) 'A two parameter variance estimator using auxiliary information', *Applied Mathematics and Computation*, Vol. 226, No. 1, pp.117–122.

Yadav, S.K., Misra, S., Mishra, S.S. and Khanal, S.P. (2017) 'Variance estimator using tri-mean and inter quartile range of auxiliary variable', *Nepalese Journal of Statistics*, Vol. 1, No. 1, pp.83–91.

Yadav, S.K., Sharma, D.K. and Mishra, S.S. (2019) 'Searching efficient estimator of population variance using tri-mean and third quartile of auxiliary variable', *Int. J. Business and Data Analytics*, Vol. 1, No. 1, pp.30–40.