Journal of King Saud University - Science 36 (2024) 103287

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Journal of King Saud University - Science

journal homepage: www.sciencedirect.com

Full Length Article

# The impact of transformations on the performance of variance estimators of finite population under adaptive cluster sampling with application to ecological data

Hameed Ali *, Sayed Muhammad Asim, Khazan Sher

*Department of Statistics, University of Peshawar, Pakistan*

ARTICLE INFO

*Keywords:*
Adaptive cluster sampling
Auxiliary information
Transformation
Dominance region
MSE
Simulation study

ABSTRACT

This paper aims to investigate the impact of transformed auxiliary variables on the performance of variance estimators of finite population under adaptive cluster sampling scheme. Further, the formulation of an efficient variance estimator of a finite population is also under consideration in this article. Specifically, we explore the gain in efficiency obtained through various transformations and define dominance space for each transformation. These dominance regions provide valuable insights into the circumstances under which one transformation prevails over another regarding precision and accuracy. The theoretical properties of the suggested estimators have been discussed along with the dominance region under each transformation. The bias and Mean Square Error (MSE) have been derived up to the first order of approximation. To evaluate and empirically validate our methodology, we conduct a numerical analysis using real-life ecological data of blue-winged teal. The finding reflects the superior performance of the suggested variance estimators over the competing estimators, thereby substantiating its importance in making informed decisions in real-world applications.

## 1. Introduction

Sampling plays a vital role in making informed decisions in real-life domains. Inferences about the statistical population or data are based on the information extracted from the sample. Therefore, a sample must be representative, mirroring every characteristic of the population of interest (Lohr, 2021). Consequently, special care must be taken in selecting a representative sample at the design and estimation stage. Adaptive cluster sampling (ACS) is of prime importance in the field of survey sampling, in situations when the variable of interest is rare, clumpy, and clustered with localized variability (Smith et al., 1995). Unlike traditional sampling methods like simple, systematic, and stratified random sampling, select units in the sample without observing it, resulting in high bias and mean square error. ACS allows the dynamic adjustment of sampling effort based on observed values to satisfy some pre-determined condition $C(y_i > 0)$, thereby enhancing the efficiency of data collection as well as parameter estimation in specific contexts. This paper investigates the domain of ACS, with a specific emphasis on the use of transformed auxiliary variables to formulate efficient variance and enhance efficiency Fig. 1.

In survey sampling, practitioners and researchers face the challenge of optimizing sampling efforts to gather meaningful data and estimate parameters precisely. The problem becomes more challenging in a situation when the population is rare and clustered where conventional sampling efforts like simple random sampling, systematic random sampling, etc. lose their effectiveness and result in high bias and low efficiency in estimating parameters (Thompson, 1990). Therefore, the use of conventional sampling strategies leads us to doubtful and misleading inferences. This inadequacy of the design and estimation problem of classical sampling methods demands the exploration of innovative methods at both the design and estimation stages. Such as ACS and the adequate use of auxiliary information in combination with the main study variable can cater to dynamic sampling requirements. It is revealed from the numerical analysis that the precision and efficacy of estimates of the variance of finite population under ACS can be enhanced remarkably.

The main objective of this study is to assess the impact of transformed auxiliary variables on the performance of variance estimators within the framework of ACS with implications for various persuasions, such as ecology, epidemiology, and geology, where ACS can offer enhanced insights into clustered or rare populations (Thompson, 1990). In this context, several sampling survey statisticians have done their

* Corresponding author.
*E-mail address:* hameedali@aup.edu.pk (H. Ali).

https://doi.org/10.1016/j.jksus.2024.103287
Received 30 November 2023; Received in revised form 4 June 2024; Accepted 5 June 2024
Available online 9 June 2024

remarkable contributions. (Diggle et al., 1976) works is regarded as a pioneered distance-based approach to assess spatial event randomness using adaptive cluster sampling. The work done by (Thompson, 1990) brings further innovation to sampling designs and unbiased estimators. In estimating parameters (Chao, 2004; Félix-Medina and Thompson, 2004) explored the importance of incorporating auxiliary variables in enhancing the efficiency of ratio estimators of population mean. The work done by (Chutiman et al., 2013),(Grover and Kaur, 2014), and later by (Yadav et al., 2016) encouraged the use of transformed auxiliary variables in the efficient formulation of estimators of parameters. A similar strategy of incorporating a transformed auxiliary variable with the study variable can also be seen in the work of (Gattone et al., 2016) for rare and clustered populations. (Noor-Ul-Amin et al., 2018) and (Yasmeen et al., 2018) suggested an effective variance estimator under adaptive cluster sampling (ACS) and Stratified adaptive cluster (SACS) sampling. Some recent work in the field of survey sampling on efficient formulation of variance under adaptive cluster sampling is due (Qureshi et al., 2020; Singh & Mishra, 2022; Yasmeen et al., 2022), (Ahmad et al., 2021), (Qureshi et al., 2020), (Singh and Mishra, 2022) with diverse applications specifically to ecological data and health data including COVID-19.

## 2. Methodology

Let us consider the population P of size N, where $P = (1, 2, \ldots, N)$. Let an initial sample of size n be drawn from the population using a Simple random sampling without replacement (SRSWOR) scheme such that $n < P$. Let $y_i$,$x_j$ be the unit observed in the initial sample of the main study variable and supplementary variable $\{x\}$. The supplementary variable $\{x\}$, where $x = \{x_1, x_2, \ldots x_N\}$, is supposed to be positively correlated with the study variable $\{y\}$, where $y = \{y_1, y_2, \ldots y_N\}$..

The selection of units in the primary sample and its neighboring components is based on some predefined condition $C(y_i > 0)$, according to ACS. If the unit selected by SRSWOR and observed satisfies the condition $C(y_i > 0)$ it is included in the sample. The additional sampling units vary adaptively selected in this way. A network of sampling units is therefore selected, consisting of all components that satisfy those conditions. The neighbouring components that fail to satisfy the condition $C(y_i > 0)$, is called the edge component. The network with its edge component is called a cluster, as a whole. The networks formed so, are non-overlapping and comprise the whole population.

Consider a network $\psi$ consisting of $m_k$components. Let $\psi_k$ be the $k_{th}$ network in the population contains component j. let us denote the average values of the elements of variables y and x by $w_{yj}$ and $w_{xj}$ respectively, as following

$$w_{yk} = \sum_{j\in\psi_k} \frac{y_j}{m_k} and w_{xk} = \sum_{j\in\psi_k} \frac{x_j}{m_k}. \quad (1)$$

The following terms and symbols will be used throughout this article while deriving Bias and MSE of the proposed estimators under ACS.

Suppose,

$$\left.\begin{aligned} & e_{0(w)} = \frac{s^2_{wy} - S^2_{wy}}{S^2_{wy}} \text{ and } e_{1(w)} = \frac{s^2_{wx} - S^2_{wx}}{S^2_{wx}} \\ & \text{such that } \mathrm{E}(e_{0(w)}) = \mathrm{E}(e_{1(w)}) = 0 \ , \ \mathrm{E}\left(e^2_{0(w)}\right) = \lambda(\beta_{2y} - 1) = V_{y(w)} \\ & \mathrm{E}\left(e^2_{1(w)}\right) = \lambda(\beta_{2x} - 1) = V_{x(w)} \ , \mathrm{E}\left(e_{0(w)} e_{1(w)}\right) = \lambda(\varphi_{22} - 1) = V_{yx(w)} \end{aligned}\right\} \quad (3)$$

| |
|---|
| $e_{0(w)} = \frac{s^2_{wy} - S^2_{wy}}{S^2_{wy}}, e_{1(w)} = \frac{s^2_{wx} - S^2_{wx}}{S^2_{wx}}$ error due to sampling of main study variable y and supplementary variable x respectively. |
| $\lambda = \frac{1}{n} - \frac{1}{P}$ is a finite population correction factor (fpc).<br>$\overline{y} = \frac{1}{n}\sum_{j=1}^{n} y_j$ and $\overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ are the sample mean of $y$ and $x$ respectively.<br>$\mu_{rq} = \frac{1}{S-1}\sum_{i=1}^{S} (w_{yj} - \overline{Y})^r (w_{xj} - \overline{X})^q$ is the second-order moments and $(r, q)$ is the non-negative integers.<br>$\beta_{2y} = \frac{\mu_{40}}{\mu^2_{20}}$ and $\beta_{2x} = \frac{\mu_{04}}{\mu^2_{02}}$ are the coefficients of kurtosis due to y and x respectively.<br>$\varphi_{22} = \frac{\mu_{22}}{\mu_{20}\mu_{02}}$ is the moment ratio?<br>$\overline{w}_y = \frac{1}{n}\sum_{j\varepsilon s_0} w_{yj}, \overline{w}_x = \frac{1}{n}\sum_{j\varepsilon s_0} w_{xj}$ The average of auxiliary variable x belonging to the sample $s_0$ where $s_0 \in S$ and $S$ is the collection of all samples.<br>$w_{y\kappa} = \frac{1}{m_\kappa}\sum_{j\in\psi_\kappa} y_j$, and $w_{x\kappa} = \frac{1}{m_\kappa}\sum_{j\in\psi_\kappa} x_j$ be the average values of the elements in the kth-network for variable $y$ and $x$, respectively.<br>$W_y = \sum_{j\varepsilon s_0} \frac{w_{yj}}{N}$ and $W_x = \sum_{j\varepsilon s_0} \frac{w_{xj}}{N}$ respectively.<br>$s^2_{wy} = \frac{1}{n-1}\sum_{j=1}^{n} (w_y - \overline{w_y})^2$ and $s^2_{wx} = \frac{1}{n-1}\sum_{j=1}^{n} (w_x - \overline{w_x})^2$ be the sample variances and $S^2_{wy} = \frac{1}{N-1}\sum_{j=1}^{N} (w_y - \overline{W}_y)^2$ and $S^2_{wx} = \frac{1}{N-1}\sum_{j=1}^{N} (w_x - \overline{W}_x)^2$ be the population variances of y and x respectively. |

Some existing estimators of variance of finite population under adaptive cluster sampling discussed in the literature are given as follows.

- The usual variance estimator of population variance is given by

$$t_0 = s^2_{y(w)} = \frac{1}{n-1} \sum_{j=1}^{n} \left(y_{j(w)} - \overline{y}\right)^2 \quad (1)$$

Which is an unbiased estimator with variance given by

$$var(t_0) = S^4_{y(w)} \lambda\left(\beta_{2y(w)} - 1\right) = S^4_{y(w)} V_{y(w)} \quad (2)$$

By letting $\lambda\left(\beta_{2y(w)} - 1\right) = V_{y(w)}$.

- (Isaki, 1983) suggested the ratio estimator of population variance in ACS design as follows

$$t_1 = s^2_{y(w)} \left(\frac{S^2_{x(w)}}{s^2_{x(w)}}\right) \quad (3)$$

With the following Bias and MSE

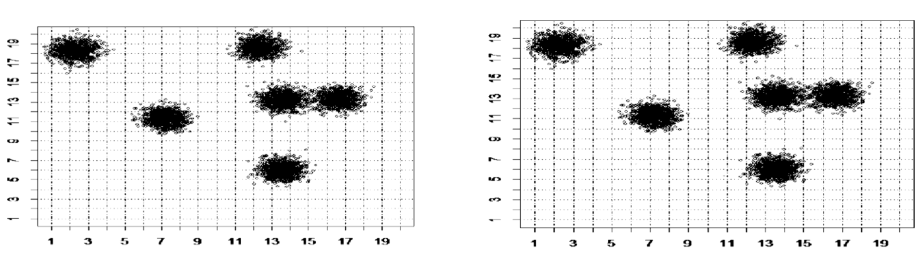


**Fig. 1.** Plot of survey variable (y) and auxiliary variable (x) in study region partitioned in 20*20 square cells generated by population-1.

$$Bias(t_1) = S^2_{y(w)} V_{y(w)} \left(1 - V_{yx(w)}\right), \tag{4}$$

And

$$MSE(t_1) = S^4_{y(w)} \left(V_{y(w)} + V_{x(w)} - 2V_{yx(w)}\right). \tag{5}$$

- (Yasmeen and Thompson, 2020) proposed the following class of estimators of finite population variance as following

$$t_{2,i} = s^2_{y(w)} \left( \frac{\alpha S_x^2 + \tau S_x^2}{\alpha s^2_{x(w)} + \tau S^2_{x(w)}} \right), \ i = 1, 2, 3, 4, 5. \tag{6}$$

Where are some suitable constants or some functions of auxiliary variables?

The Bias and MSE of $t_{2,i}$ is given by $Bias(t_{2,i}) \approx D + R^2 D V_{x(w)} - DRV_{yx(w)} + S_y^2 R^2 V_{x(w)} - S_y^2 R V_{yx(w)}$, (7)

$$MSE(t_{2,i}) \approx D^2 + \left(D^2 + S_y^2\right)^2 \left\{V_{y(w)} + R^2 V_{x(w)} - 2RV_{yx(w)}\right\}. \tag{8}$$

Where $R = \frac{\alpha}{\alpha+\tau}, D = \frac{S^2_{wy}}{S^2_{wx}} - S_y^2$ for different choices of $\alpha \& \tau$, $t_{2,i}$ takes the following special form listed in Table 1.

## 3. Proposed estimators

Motivated by (Isaki, 1983), the first estimators is proposed by taking the linear combination of usual ratio and exponential estimators in term of transformed auxiliary variable, and similarly in the second estimator is proposed by taking the linear combination of regression ratio and exponential form of transformed auxiliary variable with the main study variable as following

$$t_{P1,k} = \omega_{1k} \left( s^2_{y(w)} \frac{Z_{k(w)}}{z_{k(w)}} \right) + \omega_{2k} \left( s^2_{y(w)} \exp\left( \frac{Z_{k(w)} - z_{k(w)}}{Z_{k(w)} + z_{k(w)}} \right) \right), \tag{9}$$

$$t_{P2,k} = \omega_{3k} \left\{ s^2_{y(w)} + b\left(S^2_{x(w)} - s^2_{x(w)}\right) \right\} + \omega_{4k} \left( s^2_{y(w)} \frac{Z_{k(w)}}{z_{k(w)}} \right) + \omega_{5k} \left\{ s^2_{y(w)} \exp\left( \frac{Z_{k(w)} - z_{k(w)}}{Z_{k(w)} + z_{k(w)}} \right) \right\}, k = 1, 2, \dots, 7. \tag{10}$$

Taking motivation from (Ali et al., 2024; Cingi and Oncel Cekim, 2015; Gupta and Shabbir, 2008; Jhajj et al., 2006; Khan et al., 2015) the transformations, listed in Table 2, are suggested.

## 4. Asymptotic properties of the proposed estimators

The theoretical properties of the developed estimators are discussed along with the transformations given in Table 1, the properties of the error term will alter with each transformation and accordingly influence the sampling error as given in Table 3. Their corresponding superiority or dominance space bounds the validity of the transformation properties of the error due to sampling using the transformed auxiliary variable, we can now obtain the bias and mean square error (MSE) of $t_{P1,k}$ and $t_{P2,k}$ , k=1,2,..,7., Rewriting eq.(9) and eq. (10) in terms of the error due to sampling as following Table 4.

$$t_{P1,k} \cong S^2_{y(w)} \left(1 + e_{0(w)}\right) \left\{ \omega_{1k} \left(1 - g_k e_{1(w)} + g_k^2 e^2_{1(w)} + \cdots \right) + \omega_{2k} \left(1 - \frac{1}{2} g_k e_{1(w)} + \frac{3}{8} g_k^2 e^2_{1(w)} + \cdots \right) \right\} \tag{11}$$

And

$$t_{P2,k} \cong \omega_{3k} S^2_{y(w)} \left(1 + e_{0(w)} - \frac{V_{yx(w)}}{V_{x(w)}} e_{1(w)}\right) + \omega_{4k} S^2_{y(h)} \left( \left(1 + e_{0(w)}\right) \left(1 - g_k e_{1(w)} + g_k^2 e^2_{1(w)} + \cdots \right) \right) \tag{12}$$

Or

$$t_{P1,k} - S^2_{y(w)} \cong S^2_{y(w)} \left[ \begin{array}{l} (\omega_{1k} + \omega_{2k} - 1) + (\omega_{1k} + \omega_{2k}) e_{0(w)} - g_k \left(\omega_{1k} + \frac{\omega_{2k}}{2}\right) e_{1(w)} \\ + \left(\omega_{1k} + \frac{3\omega_{2k}}{8}\right) g_k^2 e^2_{1(w)} - g_k \left(\omega_{1k} + \frac{\omega_{2k}}{2}\right) e_{0(w)} e_{1(w)} \end{array} \right] \tag{13}$$

And

$$t_{P2,k} - S^2_{y(w)} \cong S^2_{y(w)} \left[ \begin{array}{l} (\omega_{3k} + \omega_{4k} + \omega_{5k} - 1) + (\omega_{3k} + \omega_{4k} + \omega_{5k}) \delta_{0(h)} - \left\{ \begin{array}{l} \omega_{3k} \frac{V_{22(h)}}{V_{04(h)}} \\ + \omega_{4k} + \frac{\omega_5}{2} \end{array} \right\} g_k e_{1(w)} \\ + \omega_{5k} S^2_{y(w)} \left( \left(1 + e_{0(w)}\right) \left(1 - \frac{1}{2} g_k e_{1(w)} + \frac{3}{8} g_k^2 e^2_{1(w)} + \cdots \right) \right) \end{array} \right] \tag{14}$$

Taking expectation of both sides of eq.(13) and eq.(14) and after simplification we get

$$Bias\left(t_{P1,k}\right) \cong S^2_{y(w)} \left[ (\omega_{1k} + \omega_{2k} - 1) + \left(\omega_{1k} + \frac{3\omega_{2k}}{8}\right) g_k^2 V_{x(w)} - g_k \left(\omega_{1k} + \frac{\omega_{2k}}{2}\right) V_{yx(w)} \right] \tag{15}$$

And

**Table 1**
some special cases of estimators for different transformations of auxiliary variables.

| S. No | Estimator $t_{2,i}$ | $R = \frac{\alpha}{(\alpha+\tau)}$ | Bias and MSE |
|---|---|---|---|
| 1 | $t_{2,1} = s_{wy}^2\left(\frac{S_x^2 + M_d S_x^2}{s_{wx}^2 + M_d S_{wx}^2}\right)$ | $R_1 = \frac{1}{(1+M_d)}$ | $Bias(t_{2,1}) \approx D + R_1^2 D V_{x(w)} - D R_1 V_{yx(w)} + S_y^2 R_1^2 V_{x(w)} - S_y^2 R_1^2 V_{yx(w)} MSE(t_{2,1}) \approx D^2 + \left(D^2 + S_y^2\right)^2 \{V_{y(w)} + R_1^2 V_{x(w)} - 2R_1 V_{yx(w)}\}$ |
| 2 | $t_{2,2} = s_{wy}^2\left(\frac{\rho S_x^2 + M_d S_x^2}{\rho s_{wx}^2 + M_d S_{wx}^2}\right)$ | $R_2 = \frac{\rho}{(\rho+M_d)}$ | $Bias(t_{2,2}) \approx D + R_2^2 D V_{x(w)} - D R_2 V_{yx(w)} + S_y^2 R_2^2 V_{x(w)} - S_y^2 R_2^2 V_{yx(w)} MSE(t_{2,2}) \approx D^2 + \left(D^2 + S_y^2\right)^2 \{V_{y(w)} + R_2^2 V_{x(w)} - 2R_2 V_{yx(w)}\}$ |
| 3 | $t_{2,3} = s_{wy}^2\left(\frac{C_x S_x^2 + M_d S_x^2}{C_x s_{wx}^2 + M_d S_{wx}^2}\right)$ | $R_3 = \frac{C_x}{(C_x+M_d)}$ | $Bias(t_{2,3}) \approx D + R_3^2 D V_{x(w)} - D R_3 V_{yx(w)} + S_y^2 R_3^2 V_{x(w)} - S_y^2 R_3^2 V_{yx(w)} MSE(t_{2,3}) \approx D^2 + \left(D^2 + S_y^2\right)^2 \{V_{y(w)} + R_3^2 V_{x(w)} - 2R_3 V_{yx(w)}\}$ |
| 4 | $t_{2,4} = s_{wy}^2\left(\frac{\beta_1 S_x^2 + M_d S_x^2}{\beta_1 s_{wx}^2 + M_d S_{wx}^2}\right)$ | $R_4 = \frac{\beta_1}{(\beta_1+M_d)}$ | $Bias(t_{2,4}) \approx D + R_4^2 D V_{x(w)} - D R_4 V_{yx(w)} + S_y^2 R_4^2 V_{x(w)} - S_y^2 R_4^2 V_{yx(w)} MSE(t_{2,4}) \approx D^2 + \left(D^2 + S_y^2\right)^2 \{V_{y(w)} + R_4^2 V_{x(w)} - 2R_4 V_{yx(w)}\}$ |
| 5 | $t_{2,5} = s_{wy}^2\left(\frac{\beta_2 S_x^2 + M_d S_x^2}{\beta_2 s_{wx}^2 + M_d S_{wx}^2}\right)$ | $R_5 = \frac{\beta_2}{(\beta_2+M_d)}$ | $Bias(t_{2,5}) \approx D + R_5^2 D V_{x(w)} - D R_5 V_{yx(w)} + S_y^2 R_5^2 V_{x(w)} - S_y^2 R_5^2 V_{yx(w)} MSE(t_{2,5}) \approx D^2 + \left(D^2 + S_y^2\right)^2 \{V_{y(w)} + R_5^2 V_{x(w)} - 2R_5 V_{yx(w)}\}$ |

**Table 2**
Transformed auxiliary variables and their impact on the error due to sampling and the dominance space.

| Transformed Auxiliary Variable | Error term | Transformer/normalizers | Properties of Error term | Dominance region |
|---|---|---|---|---|
| $z_{1(w)} = s_{x(w)}^2 + \alpha_1\left(S_{x(w)}^2 - s_{x(w)}^2\right)$<br>$Z_{1(w)} = S_{x(w)}^2$ | $e_{11(w)} = g_1 e_{1(w)}$ | $g_1 = 1 - \alpha_1$ | $E(e_{11(w)}) = 0$ and<br>$E(e_{11(w)}^2) = {g_1}^2 V_{x(w)} = V_{x(w),1}$<br>$E(e_{0(w)}e_{11(w)}) = g_1 V_{yx(w)} = V_{yx(w),1}$ | $0 < \alpha_1 < 1$ |
| $z_{2(w)} = \alpha_2 s_{x(w)}^2 + (1-\alpha_2)(S_{x(w)}^2 - s_{x(w)}^2)$<br>$Z_{2(w)} = \alpha_2 S_{x(w)}^2$ | $e_{12}(w) = g_2 e_{1(w)}$ | $g_2 = 2 - \frac{1}{\alpha_2}$ | $E(e_{12(w)}) = 0$ and<br>$E(e_{12(w)}^2) = {g_2}^2 V_{x(w)} = V_{x(w),2}$<br>$E(e_{0(w)}e_{12(w)}) = g_2 V_{yx(w)} = V_{yx(w),2}$ | $0.5 < \alpha_2 < \infty$ |
| $z_{3(w)} = s_{x(w)}^2 + S_{x(w)}^2(\alpha_3 - 1)$<br>$Z_{3(w)} = \alpha_3 S_{x(w)}^2$ | $e_{13(w)} = g_3 e_{1(w)}$ | $g_3 = \frac{1}{\alpha_3}$ | $E(e_{13(w)}) = 0 and$<br>$E\left(e_{13(w)}^2\right) = {g_3}^2 V_{x(w)} = V_{x(w),3}$<br>$E(e_{0(w)}e_{13(w)}) = g_3 V_{yx(w)} = V_{yx(w),3}$ | $0 < \alpha_3 < 1$ |
| $z_{4(w)} = \alpha_4 s_{x(w)}^2 + \beta_1(S_{x(w)}^2 - s_{x(w)}^2)$<br>$Z_{4(w)} = \alpha_4 S_{x(w)}^2$ | $e_{14(w)} = g_4 e_{1(w)}$ | $g_4 = 1 - \frac{\beta_1}{\alpha_4}$ | $E(e_{14(w)}) = 0 and$<br>$E\left(e_{14(w)}^2\right) = {g_4}^2 V_{x(w)} = V_{x(w),4}$<br>$E(e_{0(w)}e_{14(w)}) = g_4 V_{yx(w)} = V_{yx(w),4}$ | $\beta_1 < \alpha_4$ and both$\beta_1, \alpha_4 > 0$ |
| $z_{5(w)} = \alpha_5 s_{x(w)}^2 + \beta_2$<br>$Z_{5(w)} = \alpha_5 S_{x(w)}^2 + \beta_2$ | $e_{15(w)} = g_5 e_{1(w)}$ | $g_5 = \frac{\alpha_5 S_{x(w)}^2}{\alpha_5 S_{x(w)}^2 + \beta_2}$ | $E(e_{15(w)}) = 0 and$<br>$E\left(e_{15(w)}^2\right) = {g_5}^2 V_{x(w)} = V_{x(w),5}$<br>$E(e_{0(w)}e_{15(w)}) = g_5 V_{yx(w)} = V_{yx(w),5}$ | $\alpha_5, \beta_2 > 0$ |
| $z_{6(w)} = \alpha_6 s_{x(w)}^2 - \beta_3$<br>$Z_{6(w)} = \alpha_6 S_{x(w)}^2 - \beta_3$ | $e_{16(w)} = g_6 e_{1(w)}$ | $g_6 = \frac{\alpha_6 S_{x(w)}^2}{\alpha_6 S_{x(w)}^2 - \beta_3}$ | $E(e_{16(w)}) = 0$ and<br>$E(e_{16(w)}^2) = {g_6}^2 V_{x(w)} = V_{x(w),6}$<br>$E(e_{0(w)}e_{16(w)}) = g_6 V_{yx(w)} = V_{yx(w),6}$ | $\alpha_6 S_{x(w)}^2 - \beta_3 > 0$ |
| $z_{7(w)} = \alpha_7 s_{x(w)}^2 + (\alpha_7 + \beta_4)S_{x(w)}^2$<br>$Z_{7(w)} = (2\alpha_7 + \beta_4)S_{x(w)}^2$ | $e_{17(w)} = g_7 e_{1(w)}$ | $g_7 = \frac{\alpha_7}{2\alpha_7 + \beta_4}$ | $E(e_{17(w)}) = 0 and$<br>$E\left(e_{17(w)}^2\right) = {g_7}^2 E\left(\varepsilon_{1(w)}^2\right)$<br>$E\left(e_{17(w)}^2\right) = {g_7}^2 V_{x(w)} = V_{x(w),7}$<br>$E(e_{0(w)}e_{17(w)}) = g_7 V_{yx(w)} = V_{yx(w),7}$ | $\alpha_7, \beta_4 > 0$ |

$$Bias(t_{P2,k}) \cong S_{y(w)}^2\left[(\omega_{3k} + \omega_{4k} + \omega_{5k} - 1) + \left\{\omega_{4k} + \frac{3}{8}\omega_{5k}\right\} g_k^2 V_{x(w)} - \left(\omega_{4w} + \frac{\omega_{5w}}{2}\right) g_k V_{yx(w)}\right] \quad (16)$$

Squaring both sides of eq. (13) and eq.(14) and applying expectation, to obtain the MSE of $t_{P1,k}$ and $t_{P2,k}$,k=1,2,…,7. as following

$$MSE(t_{P1,k}) \cong S_y^4\left[A_{1k}\omega_{1k}^2 + A_{2k}\omega_{2k}^2 + A_{3k}\omega_{1k} + A_{4k}\omega_{2k} + A_{5k}\omega_{1k}\omega_{2k} + 1\right] \quad (17)$$

Where

$$A_{1k} = \left(3V_{x(w),k}^2 + V_{y(w)}^2 - 4V_{yx(w),k}\right) + 1, A_{2k} = \left(V_{x(w),k}^2 + V_{y(w)}^2 - 2V_{yx(w),k}\right) + 1$$

$$A_{3k} = 2\left(V_{yx(w),k} - V_{x(w),k}^2\right) - 2, A_{4k} = \left(V_{yx(w),k} - \frac{3}{4}V_{x(w),k}^2\right) - 2$$

$$A_{5k} = \left(\frac{15}{4}V_{x(w),k}^2 + 2V_{y(w)}^2 - 6V_{yx(w),k}\right) + 2$$

To find the optimum value of $\omega_1, \omega_2, \omega_3, \omega_4 \, and \, \omega_5$, we use calculus rule

$$MSE(t_{P2,k}) \cong S_{y(w)}^4\left(\begin{array}{c}(\omega_3 + \omega_4 + \omega_5 - 1)^2 + (\omega_3 + \omega_4 + \omega_5)^2 V_{y(w)} + \{\left(\omega_{3k}\frac{V_{yx(w)}}{V_{x(w)}} + \omega_{4k} + \frac{\omega_{5k}}{2}\right)^2 + \\ 2(\omega_{3k} + \omega_{4k} + \omega_{5k} - 1)\left(\omega_{4k} + \frac{3}{8}\omega_{5k}\right)\} V_{x(w),k} - 2\{(\omega_3 + \omega_4 + \omega_5 - 1)\left(\omega_4 + \frac{1}{2}\omega_5\right) \\ + (\omega_3 + \omega_4 + \omega_5)\left(\omega_3\frac{V_{yx(w)}}{V_{x(w)}} + \omega_4 + \frac{\omega_5}{2}\right)\} V_{yx(w),k}\end{array}\right) \quad (18)$$

of differentiating the squared loss functions (MSEs) and equating to zero to find the minimum value of MSEs function w.r.t $\omega_{1k}$, $\omega_{2k}$, $\omega_{3k}$, $\omega_{4k}$ *and* $\omega_{5k}$ . This gives

$$\omega_{1(opt)} = -\frac{2A_{2k}A_{3k} - A_{4k}A_{5k}}{4A_{1k}A_{2k} - A_{5k}^2},\ \omega_{2(opt)} = -\frac{2A_{1k}A_{4k} - A_{3k}A_{5k}}{4A_{1k}A_{2k} - A_{5k}^2}$$

And

$$\omega_{5(opt)} = \frac{-8\left\{\begin{array}{l} 4V_{x(w),k}^4 V_{yx(w),k} - 3V_{x(w),k}^4 V_{y(w)} - 17V_{x(w),k}^3 V_{yx(w),k}^2 - 12V_{x(w),k}^3 V_{y(w)} V_{yx(w),k} - \\ 30V_{x(w),k}^2 V_{yx(w),k}^3 + 5V_{x(w),k} V_{yx(w),k}^4 - 4V_{yx(w),k}^5 - 4V_{x(w),k}^3 V_{yx(w),k} - 8V_{x(w),k}^2 V_{yx(w),k}^2 \end{array}\right\}}{V_{x(w),k}\left(\begin{array}{l} 25V_{x(w),k}^4 - 112V_{x(w),k}^3 V_{yx(w),k} - 16V_{x(w),k}^3 V_{y(w)} + 240V_{x(w),k}^2 V_{yx(w),k}^2 \\ -192V_{x(w),k} V_{yx(w),k}^3 + 80V_{yx(w),k}^4 - 16V_{x(w),k}^3 \end{array}\right)}.$$

Substituting the optimum value of $\omega_1$ *and* $\omega_2$ $\omega_3$, $\omega_4$ *and* $\omega_5$ in eq.(17) and eq.(18), we get

$$MSE(t_{P1,k})_{min} \cong \lambda S_{y(w)}^4 \left(1 - \left(\frac{A_{2k}A_{3k}^2 + A_{1k}A_{4k}^2 - A_{3k}A_{4k}A_{5k}}{4A_{1k}A_{2k} - A_{5k}^2}\right)\right), \tag{19}$$

$$\begin{aligned} MSE(t_{P2,k})_{\min} \cong\ & S_{y(w)}^4 \Big[25V_{x(w),k}^5 V_{y(w)} - V_{x(w),k}^4 \Big(41V_{yx(w),k}^2 + \\ & 136V_{y(w)} V_{yx(w),k} + 16V_{y(w)}\Big) + V_{x(w),k}^3 \Big( \\ & 184V_{yx(w),k}^3 + 192V_{y(w)} V_{yx(w),k}^2 + 32V_{yx(w)}^2\Big) \\ & -V_{x(w),k}^2 \Big(153V_{yx(w),k}^4 - 64V_{yx(w),k}^3\Big) - V_{yx(w),k}^4 \\ & \{V_{x(w),k}(216V_{yx(w),k} - 80)\} - 64V_{yx(w),k}^2\Big] / \\ & [25V_{x(w),k}^5 - V_{x(w),k}^4\{16 + (112V_{yx(w),k} + \\ & 16V_{x(w),k}^4)\} + 240V_{x(w),k}^3 V_{yx(w),k}^2 - \\ & 192V_{x(w),k}^2 V_{yx(w),k}^3 + 80V_{x(w),k} V_{yx(w),k}^4]. \end{aligned} \tag{20}$$

This complete the final expression of minimmum MSEs of the proposed estimators for k=1,2,…,7. Howevere, as for practice it is observed that the MSEs can further be reduced if proper choice of auxiliary variable's parameter or constants are use in the transformation within the dominance region.

## 5. Theoretical comparisons

The theoretical comparison of the first and second proposed class of estimators given by eq.(9) to eq.(10) for k=1,2,…,6. against the competing estimators given by eq.(2), eq.(5) and eq.(8) and some special cases of eq.(8) for i=1,2,…,5., discussed in the literature under adaptive cluster sampling is given as following:

- The proposed estimator given by eq.(9) and eq.(10) well outperform the usual classical estimator $t_0$ given by eq.(2) in ACS, if

$$MSE(t_{P1,k}) \leqslant Var(t_0) \Rightarrow \frac{Var(t_0)}{MSE(t_{P1,k})} > 1, k = 1, 2, .., 7.$$

and

$$MSE(t_{P2,k}) \leqslant Var(t_0) \Rightarrow \frac{Var(t_0)}{MSE(t_{P2,k})} > 1, k = 1, 2, .., 7.$$

Or $\frac{Var(t_0)}{MSE(t_{P1,k})} \times 100 > 100 \Rightarrow PRE(t_{P1,k}, t_0) > 100.$

and

$$\frac{Var(t_0)}{MSE(t_{P1,k})} \times 100 > 100 \Rightarrow PRE(t_{P1,k}, t_0) > 100$$

- The proposed estimator given by eq.(9) and eq.(10) will outperform the ratio type estimator given by eq.(5) if

$$MSE(t_{P1,k}) \leqslant MSE(t_1) \Rightarrow \frac{MSE(t_1)}{MSE(t_{P1,k})} > 1, k = 1, 2, .., 7.$$

And

$$MSE(t_{P2,k}) \leqslant MSE(t_1) \Rightarrow \frac{MSE(t_1)}{MSE(t_{P2,k})} > 1, k = 1, 2, .., 7.$$

Or

$$\frac{MSE(t_1)}{MSE(t_{P1,k})} \times 100 > 100 \Rightarrow PRE(t_{P1,k}, t_1) > 100$$

And

$$\frac{MSE(t_1)}{MSE(t_{P2,k})} \times 100 > 100 \Rightarrow PRE(t_{P2,k}, t_1) > 100$$

- The proposed estimator will outperform the ratio type transformed class of estimator given by (8) and with special cases given in Table1 if

$$\frac{MSE(t_{2,m})}{MSE(t_{P1,k})} \times 100 > 100 \Rightarrow PRE(t_{P1,k}, t_{2,m}) > 100$$

$MSE(t_{P1}) \leqslant MSE(t_{2,m}) \Rightarrow \frac{MSE(t_{2,m})}{MSE(t_{P1,k})} > 1$, m=1, 2,..,5 and k=1,2,…,7.

The above conditions hold true for all types of data when there is a positive correlation between the main survey variable and auxiliary variable.

**Table 3**
Blue Winged Teal Data (Smith et al., 1995).

| 0 | 0 | 3 | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 24 | 14 | 0 | 0 | 10 | 103 | 0 |
| 0 | 0 | 0 | 0 | 2 | 3 | 2 | 0 | 13,639 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 37 | 14 | 122 |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 177 |

**Table 4**
Simulated y Values (Smith et al., 1995).

| 0 | 0 | 11 | 17 | 0 | 0 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 95 | 51 | 0 | 0 | 39 | 422 | 0 |
| 0 | 0 | 0 | 0 | 9 | 12 | 7 | 0 | 54,483 | 4 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 53 | 499 |
| 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 734 |

**Table 5**
Relative Efficiencies of the Proposed Estimators and Competing Estimators against the usual Variance under Simulated Model given by (21) using the first Population.

| Estimators | Relative efficiency Sample Size | | | |
|---|---|---|---|---|
| | 7 | 20 | 34 | 48 |
| $t_1$ | 2502.7 | 16063.8 | 61005.73 | 87095.37 |
| $t_{2,1}$ | 2663.8 | 25592.8 | 462054.1 | 607055.3 |
| $t_{2,3}$ | 2726.1 | 29603.3 | 409460.5 | 615805.4 |
| $t_{2,5}$ | 2715.3 | 24423.4 | 484324.2 | 629328.4 |
| $(t_{P1})_{\alpha_1=\rho_{yx(w)}}$ | 5426.7 | 37095.1 | 505865.4 | 682067.2 |
| $(t_{P1})_{\alpha_1=0.5}$ | 6020.2 | 37536.2 | 554446.3 | 683554.0 |
| $(t_{P1})_{\alpha_2=\rho_{yx(w)}}$ | 6065.0 | 37478.0 | 538798.2 | 683193.4 |
| $(t_{P1})_{\alpha_4=S^2_{x(w)},\beta_1=C^2_{x(w)}}$ | 6020.2 | 37536.2 | 554446.3 | 683554.01 |
| $(t_{P1})_{\alpha_4=N,\beta_1=n}$ | 6091.2 | 38273.11 | 509,529 | 700388.23 |
| $(t_{P1})_{\alpha_4=1/2,\beta_1=1}$ | 6141.42 | 38653.20 | 519458.05 | 682332.57 |
| $(t_{P1})_{\alpha_5=S^2_{x(w)},\beta_2=C^2_{x(w)}}$ | 6230.18 | 38707.73 | 511665.73 | 682800.41 |
| $(t_{P1})_{\alpha_5=\rho_{yx(w)},\beta_2=C^2_{x(w)}}$ | 6145.83 | 37209.67 | 513223.19 | 693910.56 |
| $(t_{P1})_{\alpha_7=S^2_{x(w)},\beta_4=C^2_{x(w)}}$ | 6151.97 | 37347.45 | 516632.00 | 708435.74 |
| $(t_{P1})_{\alpha_7=N,\beta_4=n}$ | 6065.51 | 38715.91 | 504457.21 | 697522.02 |
| $(t_{P1})_{\alpha_7=V_{x(w)},\beta_4=N}$ | 6044.42 | 37703.24 | 508780.34 | 685366.44 |
| $(t_{P2})_{\alpha_3=\rho_{yx(w)}}$ | 6091.22 | 37140.56 | 518742.73 | 706059.25 |
| $(t_{P2})_{\alpha_3=1}$ | 6250.19 | 38230.83 | 513023.41 | 702638.03 |
| $(t_{P2})_{\alpha_3=2/3}$ | 6067.62 | 38319.19 | 506546.24 | 685560.91 |
| $(t_{P2})_{\alpha_6=V_{x(w)},\beta_3=N}$ | 6065.08 | 37478.02 | 538798.01 | 683193.47 |
| $(t_{P2})_{\alpha_6=N,\beta_3=C^2_{x(w)}}$ | 7055.31 | 51024.07 | 601145.31 | 791147.51 |
| $(t_{P2})_{\alpha_6=1,\beta_3=1/2}$ | 7513.26 | 50963.81 | 602356.39 | 792064.30 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=1/2}$ | 7325.14 | 51167.29 | 602063.71 | 791072.11 |

**Table 6**
Relative Efficiencies of the Proposed Estimators and Competing Estimators against the usual variance under simulated model given by (21) using 2nd population.

| Estimators | Relative efficiency Sample size | | | |
|---|---|---|---|---|
| | 4 | 12 | 18 | 20 |
| $t_1$ | 45.0193 | 191.241 | 376.1015 | 423.7462 |
| $t_{2,1}$ | 49.5371 | 364.964 | 2894.187 | 5221.121 |
| $t_{2,3}$ | 54.6728 | 372.547 | 4010.763 | 3060.547 |
| $t_{2,5}$ | 52.7281 | 414.849 | 2261.723 | 3771.930 |
| $(t_{P1})_{\alpha_1=\rho_{yx(w)}}$ | 94.152 | 440.951 | 4058.425 | 5513.719 |
| $(t_{P1})_{\alpha_1=0.5}$ | 96.1619 | 445.719 | 4544.176 | 5520.819 |
| $(t_{P1})_{\alpha_2=\rho_{yx(w)}}$ | 98.5221 | 444.41 | 4387.849 | 5575.152 |
| $(t_{P1})_{\alpha_4=S^2_{x(w)},\beta_1=C^2_{x(w)}}$ | 99.2121 | 441.835 | 4282.176 | 5441.459 |
| $(t_{P1})_{\alpha_4=N,\beta_1=n}$ | 96.1619 | 455.700 | 4417.211 | 5511.004 |
| $(t_{P1})_{\alpha_4=1/2,\beta_1=1}$ | 99.8179 | 451.740 | 4514.267 | 5571.877 |
| $(t_{P1})_{\alpha_5=S^2_{x(w)},\beta_2=C^2_{x(w)}}$ | 96.124 | 443.591 | 4351.560 | 5591.416 |
| $(t_{P1})_{\alpha_5=\rho_{yx(w)},\beta_2=C^2_{x(w)}}$ | 98.3215 | 455.970 | 4543.618 | 5404.716 |
| $(t_{P1})_{\alpha_7=S^2_{x(w)},\beta_4=C^2_{x(w)}}$ | 98.3001 | 445.145 | 4516.673 | 5609.886 |
| $(t_{P1})_{\alpha_7=N,\beta_4=n}$ | 94.6021 | 450.581 | 4498.267 | 5590.5601 |
| $(t_{P1})_{\alpha_7=V_{x(w)},\beta_4=N}$ | 96.1619 | 449.883 | 4456.618 | 5518.7841 |
| $(t_{P2})_{\alpha_3=\rho_{yx(w)}}$ | 92.8013 | 454.910 | 4501.7814 | 5611.1708 |
| $(t_{P2})_{\alpha_3=1}$ | 89.1525 | 455.100 | 41201.568 | 5589.1355 |
| $(t_{P2})_{\alpha_3=2/3}$ | 96.2445 | 456.733 | 4414.3856 | 5567.7814 |
| $(t_{P2})_{\alpha_6=V_{x(w)},\beta_3=N}$ | 88.5128 | 484.407 | 4271.1943 | 5651.4589 |
| $(t_{P2})_{\alpha_6=N,\beta_3=C^2_{x(w)}}$ | 101.100 | 510.189 | 5135.9102 | 6610.7183 |
| $(t_{P2})_{\alpha_6=1,\beta_3=1/2}$ | 101.168 | 499.154 | 5210.6193 | 6680.8925 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=1/2}$ | 100.937 | 491.692 | 5219.7183 | 6639.7435 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=3/4}$ | 99.6571 | 501.315 | 5339.6391 | 6715.8492 |
| $(t_{P2})_{\alpha_6=N,\beta_3=\rho_{yx(w)}}$ | 98.4534 | 511.201 | 5115.1482 | 6698.4189 |
| $(t_{P2})_{\alpha_6=1,\beta_3=\rho_{yx(w)}}$ | 101.155 | 509.553 | 5209.4519 | 6701.1473 |
| $(t_{P2})_{\alpha_6=N,\beta_3=S^2_{x(w)}}$ | 101.765 | 493.981 | 5203.5167 | 6751.754 |
| $(t_{P2})_{\alpha_6=S^2_{x(w)},\beta_3=C^2_{x(w)}}$ | 99.0346 | 501.191 | 5318.8152 | 6705.6103 |

**Table 7**
Relative Efficiencies of the Proposed Estimators and Competing Estimators against the usual Variance under the Simulated Model given by (22) using the first Population.

| Estimators | Relative efficiency Sample size | | | | |
|---|---|---|---|---|---|
| | 4 | 8 | 12 | 18 | 20 |
| $t_1$ | 4.04E-06 | 3.07E-04 | 8.95E-05 | 2.99E-04 | 0.011 |
| $t_{2,1}$ | 3.58 | 0.01269 | 0.631 | 0.284 | 0.032 |
| $t_{2,2}$ | 3.68 | 0.01292 | 0.635 | 0.277 | 0.080 |
| $t_{2,3}$ | 3.581 | 0.01297 | 0.621 | 0.259 | 0.137 |
| $t_{2,4}$ | 3.567 | 0.01259 | 0.630 | 0.261 | 0.076 |
| $t_{2,5}$ | 3.577 | 0.01274 | 0.621 | 0.261 | 0.077 |
| $(t_{P1})_{\alpha_2=\rho_{yx(w)}}$ | 11.041 | 2.035 | 0.944 | 0.786 | 0.1939 |
| $(t_{P1})_{\alpha_4=S^2_{x(w)},\beta_1=C^2_{x(w)}}$ | 11.129 | 1.964 | 1.077 | 0.818 | 0.2244 |
| $(t_{P1})_{\alpha_4=N,\beta_1=n}$ | 11.247 | 1.942 | 1.179 | 0.761 | 0.1378 |
| $(t_{P1})_{\alpha_4=1/2,\beta_1=1}$ | 10.645 | 1.904 | 1.005 | 0.837 | 0.1143 |
| $(t_{P1})_{\alpha_5=S^2_{x(w)},\beta_2=C^2_{x(w)}}$ | 11.037 | 2.086 | 1.094 | 0.788 | 0.0703 |
| $(t_{P1})_{\alpha_5=\rho_{yx(w)},\beta_2=C^2_{x(w)}}$ | 11.093 | 1.964 | 1.856 | 0.788 | 0.0801 |
| $(t_{P1})_{\alpha_7=S^2_{x(w)},\beta_4=C^2_{x(w)}}$ | 10.847 | 2.045 | 1.071 | 0.734 | 0.082 |
| $(t_{P1})_{\alpha_7=N,\beta_4=n}$ | 10.132 | 1.905 | 1.106 | 0.816 | 0.1308 |
| $(t_{P1})_{\alpha_7=V_{x(w)},\beta_4=N}$ | 10.939 | 2.053 | 0.929 | 0.781 | 0.1045 |
| $(t_{P2})_{\alpha_3=\rho_{yx(w)}}$ | 10.269 | 2.094 | 1.092 | 0.838 | 0.2006 |
| $(t_{P2})_{\alpha_3=1}$ | 10.845 | 1.911 | 0.924 | 0.730 | 0.0865 |
| $(t_{P2})_{\alpha_3=2/3}$ | 11.133 | 1.904 | 1.123 | 0.713 | 0.0838 |
| $(t_{P2})_{\alpha_6=V_{x(w)},\beta_3=N}$ | 10.116 | 2.015 | 1.016 | 0.836 | 0.1253 |
| $(t_{P2})_{\alpha_6=N,\beta_3=C^2_{x(w)}}$ | 10.893 | 1.973 | 1.162 | 0.750 | 0.1765 |
| $(t_{P2})_{\alpha_6=1,\beta_3=1/2}$ | 11.319 | 2.013 | 0.911 | 0.704 | 0.1907 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=1/2}$ | 11.149 | 2.046 | 0.950 | 0.855 | 0.2139 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=3/4}$ | 10.209 | 1.996 | 1.075 | 0.786 | 0.2658 |
| $(t_{P2})_{\alpha_6=N,\beta_3=\rho_{yx(w)}}$ | 10.749 | 1.959 | 0.932 | 0.825 | 0.2642 |
| $(t_{P2})_{\alpha_6=1,\beta_3=\rho_{yx(w)}}$ | 10.564 | 1.902 | 1.149 | 0.763 | 0.2216 |
| $(t_{P2})_{\alpha_6=N,\beta_3=S^2_{x(w)}}$ | 11.073 | 2.077 | 0.970 | 0.877 | 0.2193 |
| $(t_{P2})_{\alpha_6=S^2_{x(w)},\beta_3=C^2_{x(w)}}$ | 10.603 | 1.929 | 1.061 | 0.857 | 0.1386 |
| $(t_{P2})_{\alpha_6=S^2_{x(w)},\beta_3=\rho_{yx(w)}}$ | 11.142 | 2.087 | 1.179 | 0.767 | 0.1642 |

## 6. Numerical analysis

The performance of the proposed estimator against competing estimators was demonstrated in a simulation study under the ACS design. Two populations were used: a Poisson cluster (Diggle et al., 1976) pages 55–57. Second population is taken from (Smith et al., 1995) in which 5000 $km^2$ of area distributed among $50 \times 100$ quadrants in central Florida. The data of blue-winged teal was used as an auxiliary variable to compare the efficiency of the estimators and the estimator suggested by (Isaki, 1983) in estimating variance under adaptive cluster sampling without replacement sampling. Denoting the j-th variate of interest $y$ and auxiliary variate $w_x$ by $y_j$ and $w_{xj}$. (Dryver & Chao, 2007).

The following two models generated the survey variable, given by

$$y_j = 4x_j + \varepsilon_j,\ \ \varepsilon_j \sim N(0, x_j) \tag{21}$$

$$y_j = 4w_{xj} + \varepsilon_j,\ \ \varepsilon_j \sim N(0, w_{xj}) \tag{22}$$

The two models given by eq.(21) and eq.(22) suggest a strong correlation of the survey variable with a subsidiary variable at both, the unit level and network level respectively. The comparison is made with the (Isaki, 1983) estimator of variance in adaptive sampling design. For neighboring units to be included if $\left[y; y_j > 0\right]$

$$\text{Relative Efficiency} \ = \frac{var(t_0)}{MSE(t_*)} \times 100 \tag{23}$$

**Table 8**
Relative Efficiencies of the Proposed Estimators and competing estimators against the usual Variance under the Simulated Model given by (22) using 2nd Population.

| Estimators | Relative efficiency | | | | |
|---|---|---|---|---|---|
| | Sample size | | | | |
| | 4 | 8 | 12 | 18 | 20 |
| $t_1$ | 1.04E-12 | 4.01E-11 | 1.95E-11 | 2.99E-11 | 2.11E-10 |
| $t_{2,1}$ | 3.071 | 1.319 | 0.7201 | 0.419 | 0.32 |
| $t_{2,2}$ | 3.801 | 1.288 | 0.7395 | 0.387 | 0.32 |
| $t_{2,3}$ | 3.846 | 1.290 | 0.7173 | 0.388 | 0.33 |
| $t_{2,4}$ | 3.782 | 1.337 | 0.7325 | 0.388 | 0.32 |
| $t_{2,5}$ | 3.715 | 1.301 | 0.7391 | 0.379 | 0.3 |
| $(t_{P1})_{\alpha_2=\rho_{yx(w)}}$ | 10.97 | 9.716 | 6.074 | 2.091 | 0.926 |
| $(t_{P1})_{\alpha_4=S^2_{x(w)},\beta_1=C^2_{x(w)}}$ | 10.63 | 8.239 | 6.172 | 1.272 | 0.922 |
| $(t_{P1})_{\alpha_4=N,\beta_1=n}$ | 10.29 | 8.164 | 5.977 | 1.501 | 0.928 |
| $(t_{P1})_{\alpha_4=1/2,\beta_1=1}$ | 10.35 | 9.244 | 4.721 | 1.259 | 0.937 |
| $(t_{P1})_{\alpha_5=S^2_{x(w)},\beta_2=C^2_{x(w)}}$ | 9.871 | 8.658 | 4.386 | 1.669 | 0.819 |
| $(t_{P1})_{\alpha_5=\rho_{yx(w)},\beta_2=C^2_{x(w)}}$ | 10.78 | 9.625 | 5.271 | 1.681 | 0.734 |
| $(t_{P1})_{\alpha_7=S^2_{x(w)},\beta_4=C^2_{x(w)}}$ | 10.89 | 8.691 | 4.808 | 1.473 | 0.716 |
| $(t_{P1})_{\alpha_7=N,\beta_4=n}$ | 10.48 | 9.463 | 6.077 | 1.412 | 0.827 |
| $(t_{P1})_{\alpha_7=V_{x(w)},\beta_4=N}$ | 10.48 | 9.104 | 5.803 | 1.369 | 0.906 |
| $(t_{P2})_{\alpha_3=\rho_{yx(w)}}$ | 12.61 | 10.43 | 7.914 | 2.764 | 1.035 |
| $(t_{P2})_{\alpha_3=1}$ | 12.55 | 9.941 | 7.524 | 2.618 | 1.023 |
| $(t_{P2})_{\alpha_3=2/3}$ | 11.96 | 10.87 | 6.049 | 2.491 | 1.340 |
| $(t_{P2})_{\alpha_6=V_{x(w)},\beta_3=N}$ | 11.99 | 10.86 | 6.568 | 2.128 | 0.907 |
| $(t_{P2})_{\alpha_6=N,\beta_3=C^2_{x(w)}}$ | 12.24 | 9.783 | 6.662 | 2.918 | 1.036 |
| $(t_{P2})_{\alpha_6=1,\beta_3=1/2}$ | 12.86 | 9.425 | 6.467 | 2.077 | 1.031 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=1/2}$ | 10.06 | 9.127 | 6.217 | 2.219 | 1.021 |
| $(t_{P2})_{\alpha_6=2/3,\beta_3=3/4}$ | 10.66 | 8.434 | 6.921 | 2.163 | 1.038 |
| $(t_{P2})_{\alpha_6=N,\beta_3=\rho_{yx(w)}}$ | 13.22 | 9.221 | 6.277 | 2.183 | 1.022 |
| $(t_{P2})_{\alpha_6=1,\beta_3=\rho_{yx(w)}}$ | 12.38 | 10.13 | 6.914 | 2.141 | 1.024 |
| $(t_{P2})_{\alpha_6=N,\beta_3=S^2_{x(w)}}$ | 9.843 | 9.731 | 5.801 | 3.023 | 1.016 |
| $(t_{P2})_{\alpha_6=S^2_{x(w)},\beta_3=C^2_{x(w)}}$ | 11.75 | 10.39 | 5.139 | 3.027 | 0.832 |
| $(t_{P2})_{\alpha_6=S^2_{x(w)},\beta_3=\rho_{yx(w)}}$ | 12.16 | 10.53 | 6.001 | 3.108 | 0.737 |

Where $t_* = t_{P1}, t_{P2}, t_1, t_{2,j},\ j = 1, 2, ..., 5.$ denote the proposed class of estimators and competing estimators of variance in adaptive cluster sampling in the formula for Percent Relative efficiency (PRE) given by eq.(28).

The following steps are used in R-Language to perform simulation:

Step 1: Generate response variable y using model (21) and (22) with supplementary variable x and $W_x$ from given populations.

Step 2: Consider initial sample sizes $n = 7;\ 20;\ 34$ *and* $48$ for 100,000 repetitions to calculate the variance estimator in adaptive cluster sampling.

Step 3: Calculate 100,000 values of $t_{P1i}, t_{P2i}, t_1, t_{2,j},\ i = 1, 2, ..., 7.\ j = 1, 2, ..., 5.$ using equations (1) to (10) for different choices of $\alpha_k, \beta_j, k = 1, 2, .., 7.$ *and* $j = 1, 2, 3, 4.$.

Step 4: Compute Mean Squared Error (MSE) for both conventional and proposed estimators for each sample.

Step 5: Calculate Percent Relative Efficiency (PRE using values from steps 3 and 4 and report in Table 5-8.

## 7. Results and discussion

Adaptive Cluster Sampling (ACS) is a complex sampling technique used in statistical estimation, particularly when the characteristic of interest is rare and clustered. However, the accuracy of estimation remains a major concern. The suggested estimators consistently outperform competing estimators of finite population variance under ACS. These estimators incorporate transformed auxiliary variables, reducing mean squared error and bias. Comparative analysis reveals that (Isaki, 1983) variance estimator performs poorly compared to competing estimators. The suggested class of estimators increases efficiency with sample size, outperforming inferior estimators. Zero values in the sample and a high correlation between the survey and auxiliary variables do not significantly affect the target function estimation.

The expected sample size is calculated using a formula that sums all quadrant inclusion probabilities is given by:

$$E(\nu) = \sum_{i=1}^{N} \pi_i.$$

Interestingly, the final sample size usually grows with the size of the primary sample and is usually greater than the former.

Two proposed classes of variance estimators have been developed, incorporating auxiliary variables and known population parameters. These estimators outperform the (Isaki, 1983) estimator when dealing with moderate sample sizes and using only the primary sample. The proposed estimators are flexible and can be adapted to other sampling scenarios, such as simple random sampling, stratified random sampling, and non-response sampling. These estimators represent a promising advancement in statistical estimation, offering better results for rare and patchy populations in practical scenarios. The suggested estimators are quite flexible can be seamlessly adapted into the estimation of other parameters such as mean, median, coefficient of variation etc. thereby making a significant contribution in parameter estimation using transformed auxiliary variable.

Disclosure of any funding to the study

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure instructions

During the preparation of this work the author(s) used AI in order to remove grammatical mistakes. After using this tool/service, the author (s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

### CRediT authorship contribution statement

**Hameed Ali:** Writing – original draft, Conceptualization. **Sayed Muhammad Asim:** Writing – review & editing, Supervision, Resources, Project administration. **Khazan Sher:** Methodology, Investigation, Formal analysis, Data curation.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### References

Ahmad, S., Arslan, M., Khan, A., Shabbir, J., 2021. A generalized exponential-type estimator for population mean using auxiliary attributes. PLOS ONE 16, e0246947.
Ali, H., Asim, S.M., Ijaz, M., Zaman, T., Iftikhar, S., 2024. Improvement in variance estimation using transformed auxiliary variable under simple random sampling. Sci. Rep. 14, 8117. https://doi.org/10.1038/s41598-024-58841-x.
Chao, C.-T., 2004. Ratio Estimation On Adaptive Cluster Sampling. 中國統計學報 42, 307–327. https://doi.org/10.29973/JCSA.200409.0006.
Chutiman, N., Chiangpradit, M., Suraphee, S., 2013. A New Estimator Using Auxiliary Information in Stratified Adaptive Cluster Sampling. Open J. Stat. 03, 278–282. https://doi.org/10.4236/ojs.2013.34032.
Cingi, H., Oncel Cekim, H., 2015. Some estimator types for population mean using linear transformation with the help of the minimum and maximum values of the auxiliary variable. Hacet. J. Math. Stat. 46, 1. https://doi.org/10.15672/HJMS.201510114186.
Diggle, P.J., Besag, J., Gleaves, J.T., 1976. Statistical Analysis of Spatial Point Patterns by Means of Distance Methods. Biometrics 32, 659–667. https://doi.org/10.2307/2529754.
Félix-Medina, M.H., Thompson, S.K., 2004. Adaptive Cluster Double Sampling. Biometrika 91, 877–891.
Gattone, S.A., Mohamed, E., Dryver, A.L., Münnich, R.T., 2016. Adaptive cluster sampling for negatively correlated data. Environmetrics 27, E103–E113. https://doi.org/10.1002/env.2381.

Grover, L.K., Kaur, P., 2014. A Generalized Class of Ratio Type Exponential Estimators of Population Mean Under Linear Transformation of Auxiliary Variable. Commun. Stat. - Simul. Comput. 43, 1552–1574. https://doi.org/10.1080/03610918.2012.736579.

Gupta, S., Shabbir, J., 2008. On improvement in estimating the population mean in simple random sampling. J. Appl. Stat. 35, 559–566. https://doi.org/10.1080/02664760701835839.

Isaki, C.T., 1983. Variance Estimation Using Auxiliary Information. J. Am. Stat. Assoc. 78, 117–123. https://doi.org/10.1080/01621459.1983.10477939.

Jhajj, H.S., Sharma, M.K., Grover, L.K., 2006. Dual of Ratio Estimators of Finite Population Mean Obtained on Using Linear Transformation to Auxiliary Variable. J. Jpn. Stat. Soc. 36, 107–119. https://doi.org/10.14490/jjss.36.107.

Khan, S., Ali, H., Manzoor, S., 2015. A class of transformed efficient ratio estimators of finite population mean.

Lohr, S.L., 2021. In: Sampling: Design and Analysis, 3rd ed. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9780429298899.

Noor-Ul-Amin, M., Yasmeen, U., Hanif, M., 2018. Generalized variance estimators in adaptive cluster sampling using single auxiliary variable. J. Stat. Manag. Syst. 21, 401–415. https://doi.org/10.1080/09720510.2017.1413045.

Qureshi, M.N., Kadilar, C., Hanif, M., 2020. Estimation of rare and clustered population mean using stratified adaptive cluster sampling. Environ. Ecol. Stat. 27, 151–170. https://doi.org/10.1007/s10651-019-00438-z.

Singh, R., Mishra, R., 2022. Transformed ratio type estimators under Adaptive Cluster Sampling: An application to COVID-19. J. Stat. Appl. Probab. Lett. 9, 63–70. https://doi.org/10.18576/jsapl/090201.

Smith, D.R., Conroy, M.J., Brakhage, D.H., 1995. Efficiency of Adaptive Cluster Sampling for Estimating Density of Wintering Waterfowl. Biometrics 51, 777–788. https://doi.org/10.2307/2532964.

Thompson, S.K., 1990. Adaptive Cluster Sampling. J. Am. Stat. Assoc. 85, 1050–1059. https://doi.org/10.1080/01621459.1990.10474975.

Yadav, S.K., Misra, S., Mishra, S.S., Chutiman, N., 2016. Improved Ratio Estimators of Population Mean In Adaptive Cluster Sampling.

Yasmeen, U., Noor-ul-Amin, M., Hanif, M., 2018. Exponential Estimators of Finite Population Variance Using Transformed Auxiliary Variables. Proc. Natl. Acad. Sci. India Sect. Phys. Sci. 89 https://doi.org/10.1007/s40010-017-0410-5.

Yasmeen, U., Noor-ul-Amin, M., Hanif, M., 2022. Variance estimation in stratified adaptive cluster sampling. Stat. Transit. New Ser. 23, 173–184. https://doi.org/10.2478/stattrans-2022-0010.

Yasmeen, U., Thompson, M., 2020. Variance estimation in adaptive cluster sampling. Commun. Stat. - Theory Methods 49, 2485–2497. https://doi.org/10.1080/03610926.2019.1576890.