

Original Research Article

MODIFIED CLASSES OF REGRESSION-TYPE ESTIMATORS OF POPULATION MEAN IN THE PRESENCES OF AUXILIARY ATTRIBUTE

ABSTARCT

The use of relevant information from auxiliary variable at the estimation stage and design stage to obtain reliable and efficient estimate is a common practice in a sample survey. But situations arise when the available auxiliary information are attribute in nature. Some existing estimators based on auxiliary attribute include; Naik and Gupta (1996), Singh et al. (2007), Singh et al. (2013), Zaman and Kadilar (2019), Zaman (2020). However, the proposed estimators by Zaman and Kadilar (2019) are less efficient when the bi-serial correlation between the study variable and auxiliary attribute is negative. While Zaman (2020) depends on an unknown parameter of the study variable (C_y), hence the applicability of the estimator in real life situation is not possible unless if the value is estimated using a large sample which requires additional resources. In this work, the concept of regression base estimator was used to modified the estimator of Zaman and Kadilar (2019) to obtain estimators that is applicable for both negative and positive correlation. Also, the estimator of Zaman (2020) was modified to obtain estimator that is independent of unknown population parameter. The properties (Biases and MSEs) of the modified estimators were derived up to the first order of approximation using Taylor series approach. The efficiency conditions of the proposed estimation over the existing estimator considered in the study were established. The empirical studies were conducted using both existing population parameters and stimulation to investigate the efficiency of the proposed estimators over the efficiency of the existing estimators. The results revealed that the proposed estimators have minimum MSEs and higher PREs among all the competing estimators. These imply that the proposed estimators are more efficient and can produce better estimate of the population mean compared to other existing estimators considered in the study.

Keywords: Auxiliary Attribute, Bias, Mean Square Error (MSE), Population Mean

1. Introduction

It is well known that when auxiliary information is used appropriately in probability sampling, the variances of the estimators of population parameters such as population mean, median, variance, regression coefficient, and population correlation coefficient are significantly reduced. Auxiliary information recorded from the population elements can be successfully used to design an efficient sampling design and after sample selection, to further improve the efficiency of the estimators. If auxiliary variable is available and strongly correlated with the study variable, it is possible to improve the efficiency by using ratio, product, dual-to-ratio or regression methods of estimation. The use of these techniques can considerably improve the accuracy of estimates, i.e. produce estimates that are close to the corresponding population values. There are several

authors who have suggested estimators using some known population parameters of auxiliary variable(s). The first attempt was made by Cochran (1940) to investigate the problem of estimation of the population mean when auxiliary variables are present and he proposed the usual ratio estimator of population mean. Other authors like, Hartley-Ross (1954), Quenouille's (1956), Singh (1965, 1967), Abu-Dayeh et al. (2003), Kadilar and Cingi (2005), Khoshnevisan et al. (2007), Perri (2007), Singh et al. (2007a), Singh et al. (2020), Singh and Kumar (2011), Singh et al. (2015), Tailor et al. (2012), Lu (2013), Sharma and Singh (2014), Lu and Yan (2014), Verma et al. (2015), Audu and Singh (2015), Audu et al. (2020a,c), Audu et al. (2021a,b,c,e,f,g), Audu and Adewara (2017), Audu et al. (2016a,b,c), Yunusa et al (2021), Zaman et al (2021), have also considered the problem of estimating the mean of a survey variable when auxiliary variables are made available. There are many practical situations when auxiliary information is qualitative in nature, i.e, auxiliary information is available in the form of an attribute, such as the height of a person may depend on the fact that whether the person is male or female, the efficiency of a dog may depend on the particular breed of that dog, or the yield of wheat crop produced may depend on a particular variety of wheat, etc. Many authors have proposed estimators based on auxiliary attribute. Naik and Gupta (1996), Singh et al., (2007b), Singh et al., (2013), Zaman and Kadilar (2019), Zaman (2020) Audu et al. (2020b), Audu et al. (2021d) suggested estimators by using auxiliary attributes. All the above-mentioned authors assumed known population attribute (P). However, the estimators proposed by Zaman and Kadilar (2019) is less efficient when the bi-serial correlation between the study variable and the auxiliary attribute is negative and the unknown α in the Zaman (2020) estimator depends on an unknown parameter of the study variable (C_y) which makes it impracticable in real life situations unless if estimated using a large sample which requires additional resources. To address the above weaknesses, new classes of modified estimators were proposed using regression function to transform the aforementioned estimators.

1.1 Notations

Y: Study Variable

\bar{Y} : Population mean of variable.

\bar{y} : Sample mean

N: Population Size

n: Sample size

n' : Sample Size at first phase

$A = \sum_{i=1}^N \phi_i$: Total number of unit in the population possessing attribute ϕ

$a = \sum_{i=1}^n \phi_i$: Total number of unit in the sample possessing attribute ϕ

$P = \frac{A}{N}$: Population proportion attributes

$p = \frac{a}{n}$: Sample proportion of attributes

$p' = \frac{a}{n'}$: Sample proportion at first phase

$C_y = \frac{S_y}{\bar{Y}}$: Coefficient of population variation of the study variable

$C_\phi = \frac{S_\phi}{\bar{X}}$: Coefficient of variation of auxiliary attribute ϕ

$f = \frac{n}{N}$: Sampling fraction

$B_{2(\phi)} = \frac{\mu_4}{\sigma^4}$: Coefficient of population kurtosis of the form of the auxiliary attribute ϕ

$\rho_{y\phi} = \frac{S_{y\phi}}{S_y S_\phi}$: The population point of bi-serial correlation between the study variable y and auxiliary attribute ϕ

$\lambda = \left(\frac{1}{n} - \frac{1}{N} \right)$: Correction Factor

$p^* = \frac{NP - np}{N - n}$: Dual to ratio in single phase.

2. Literature Review

Let y_i be i -th characteristic of the population and ϕ_i is the case of possessing certain attributes. If i th unit has the desired characteristic, it takes the value 1; if not then the value 0. That is:

$$\phi_i = \begin{cases} 1, & \text{if } i\text{th unit of the population possesses attribute} \\ 0, & \text{otherwise} \end{cases} \quad (2.1)$$

Let $A = \sum_{i=1}^N \phi_i$ and $a = \sum_{i=1}^n \phi_i$ be the total count of the units that possess certain attribute in the population and the sample, respectively. $P = \frac{A}{N}$ and $p = \frac{a}{n}$ denote the proportion of unit in the population and sample possessing attribute ϕ respectively.

Naik and Gupta (1996) defined ratio estimator denoted by t_{NG} of population mean when the prior information of population proportion of units, possessing the same attribute is available which

modifies ratio estimator by Cochran (1940) as in (2.2). The Bias and MSE of t_{NG} are given as in (2.1) and (2.2)

$$t_{NG} = \frac{\bar{y}}{p} P,$$

$$Bias(t_{NG}) = \bar{Y} \lambda (C_\phi^2 - \rho_{y\phi} C_y C_\phi) \quad (2.3)$$

$$MSE(t_{NG}) = \lambda \bar{Y}^2 (C_y^2 - 2\rho_{y\phi} C_y C_\phi + C_\phi^2) \quad (2.4)$$

However, the estimator t_{NG} is less efficient when the point bi-serial correlation ($\rho_{y\phi}$) is negative and when the difference between sample and population proportion is large due to sample information.

Singh *et al.* (2007b) modified the estimator of Naik and Gupta (1996) using exponential transformation to reduce the effect of distance between sample and population proportions. The estimator denoted t_{STS} as well as its Bias and MSE as given in (2.5), (2.6) and (2.7) respectively.

$$t_{STS} = \bar{y} \exp\left(\frac{P-p}{P+p}\right) \quad (2.5)$$

The bias and MSE of estimator t_{STS} are respectively given by (2.6) and (2.7)

$$Bias(t_{STS}) = \lambda \bar{Y} \left(\frac{1}{4} C_p^2 - \frac{1}{2} \rho_{pb} C_y C_p \right) \quad (2.6)$$

$$MSE(t_{STS}) = \lambda \bar{Y}^2 \left(\frac{1}{4} C_p^2 - \rho_{pb} C_y C_p + C_y^2 \right) \quad (2.7)$$

However; t_{STS} is less efficient when the point bi-serial correlation ($\rho_{y\phi}$) is negative.

Singh *et al.*, (2013) proposed a ratio-product exponential type estimator which is applicable when bi-serial correlation is either negative or positive. The estimator denoted t_{ST} as well as its Bias and MSE as given in (2.8), (2.9) and (2.10) respectively.

$$t_{ST} = \bar{y} \left(\alpha \exp\left(\frac{P-p}{P+p}\right) + (1-\alpha) \exp\left(\frac{p-P}{p+p}\right) \right) \quad (2.8)$$

$$Bias(t_{ST}) = \lambda \bar{Y} \left(\frac{C_\phi^2}{8} + \rho_{y\phi} C_y C_\phi \left(\frac{1}{2} - \alpha \right) \right) \quad (2.9)$$

$$MSE(t_{ST}) = \lambda \bar{Y}^2 \left(C_y^2 + C_\phi^2 \left(\frac{1}{4} + \alpha^2 - \alpha \right) + 2\rho_{y\phi} C_y C_\phi \left(\frac{1}{2} - \alpha \right) \right) \quad (2.10)$$

$$K_p = \rho_{y\phi} \frac{C_y}{C_\phi}$$

$$\alpha = \frac{2K_p + 1}{2} = \alpha_0 \text{ (say)}$$

$$MSE(t_{ST})_{\min} = \lambda \bar{Y}^2 C_y^2 (1 - \rho_{y\phi}^2) \quad (2.11)$$

However, t_{ST} is a function of α which depends on unknown parameter of the study population C_y (coefficient of variation)

Zaman and Kadilar (2019) proposed family of ratio exponential estimators using known information of auxiliary attribute. The estimator denoted by t_{ZK} as well as its Bias and MSE are given in (2.12), (2.13) and (2.14) respectively

$$t_{ZK} = \bar{y} \exp\left(\frac{(kP+l)-(kp+l)}{(kP+l)+(kp+l)}\right) \quad (2.12)$$

$$Bias(t_{ZKi}) = \lambda \bar{Y} (\theta_i^2 C_\phi^2 - \theta_i \rho_{y\phi} C_y C_\phi) \quad i = 1, 2, \dots, 9 \quad (2.13)$$

$$MSE(t_{ZKi}) = \lambda \bar{Y}^2 (\theta_i^2 C_\phi^2 - 2\theta_i \rho_{y\phi} C_y C_\phi + C_y^2) \quad i = 1, 2, \dots, 9 \quad (2.14)$$

Where

$$\theta_1 = \frac{P}{2(P + \beta_2(\phi))}, \quad \theta_2 = \frac{P}{2(P + C_\phi)}, \quad \theta_3 = \frac{P}{2(P + \rho_{\phi y})}; \quad \theta_4 = \frac{\beta_2(\phi)P}{2(\beta_2(\phi)P + C_\phi)};$$

$$\theta_5 = \frac{C_\phi P}{2(C_\phi P + \beta_2(\phi))}; \quad \theta_6 = \frac{C_\phi P}{2(C_\phi P + \rho_{y\phi})}; \quad \theta_7 = \frac{\rho_{y\phi} P}{2(\rho_{y\phi} P + C_\phi)}; \quad \theta_8 = \frac{\beta_2(\phi)P}{2(\beta_2(\phi)P + \rho_{y\phi})}$$

$$; \theta_9 = \frac{\rho_{y\phi} P}{2(\rho_{y\phi} P + \beta_2(\phi))}$$

However, the estimator of t_{ZK} is less efficient when the point bi-serial correlation ($\rho_{y\phi}$) is negative.

Zaman (2020) suggested an improved class of estimator for the population mean by modifying Zaman and Kadilar (2019) as in (2.15). The Bias and MSE of t_{NG} are given as in (2.15) and (2.16).

$$t_{Zi} = \bar{y} \left(\frac{P}{P}\right)^\alpha \exp\left(\frac{(kP+l)-(kp+l)}{(kP+l)+(kp+l)}\right) \quad (2.15)$$

$$MSE(t_{Zi}) = \bar{Y}^2 (\alpha^2 C_\phi^2 + \theta_i^2 C_\phi^2 - 2\alpha\theta_i C_\phi^2 + C_y^2 + 2\alpha\rho_{y\phi} C_y C_\phi - 2\theta_i C_y C_\phi); \quad i = 0, 1, 2, \dots, 9 \quad (2.16)$$

$$MSE_{\min}(\bar{y}_{Zi}) \cong \frac{1-f}{n} \bar{Y}^2 C_y^2 (1 - \rho_{y\phi}^2) \quad (2.17)$$

$$\alpha_{opt} = \frac{\theta_i C_\phi - \rho_{y\phi} C_y}{C_\phi}; \quad i = 0, 2, \dots, 9$$

However, the estimator t_{Zi} is a function α which depends on unknown parameters $\rho_{y\phi}$ and C_y

3. Suggested estimators

Having studied the estimator of Zaman and Kadilar (2019) and Zaman (2020), the improved class of estimators denoted by t_{pi} and t_{qi} for the population mean are proposed as follows.

$$t_{pi} = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{(kP + l) - (kp + l)}{(kP + l) + (kp + l)}\right) \quad i = 1, 2, \dots, 9 \quad (3.1)$$

$$t_{qi} = \frac{(\bar{y} + b_\phi(P - p)(kP + l))}{kp^* + l} \exp\left(\frac{(kp^* + l) - (kP + l)}{(kp^* + l) + (kP + l)}\right) \quad i = 0, 1, 2, \dots, 9 \quad (3.2)$$

Where $p^* = \frac{NP - np}{N - n}$

Table 1: Members of the proposed estimator t_{pi}

Estimators	Values of k and l	
	k	l
$t_1 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{P - p}{P + p + 2\rho_{y\phi}}\right)$	1	$\beta_2(\phi)$
$t_2 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{P - p}{P + p + 2C_\phi}\right)$	1	C_ϕ
$t_3 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{P - p}{P + p + 2\rho_{y\phi}}\right)$	1	$\rho_{y\phi}$
$t_4 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{\beta_2(\phi)(P - p)}{\beta_2(\phi)(P + p) + 2C_\phi}\right)$	$\beta_2(\phi)$	C_ϕ
$t_5 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{C_\phi(P - p)}{C_\phi(P + p) + 2\beta_2(\phi)}\right)$	C_ϕ	$\beta_2(\phi)$
$t_6 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{C_\phi(P - p)}{C_\phi(P + p) + 2\rho_{y\phi}}\right)$	C_ϕ	$\rho_{y\phi}$
$t_7 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{\rho_{y\phi}(P - p)}{\rho_{y\phi}(P + p) + 2C_\phi}\right)$	$\rho_{y\phi}$	C_ϕ
$t_8 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{\beta_2(\phi)(P - p)}{\beta_2(\phi)(P + p) + 2\rho_{y\phi}}\right)$	$\beta_2(\phi)$	$\rho_{y\phi}$
$t_9 = (\bar{y} + b_\phi(P - p)) \exp\left(\frac{\rho_{y\phi}(P - p)}{\rho_{y\phi}(P + p) + 2\beta_2(\phi)}\right)$	$\rho_{y\phi}$	$\beta_2(\phi)$

Table 2: Members of the proposed estimator t_{qi}

Estimators	Values of k and l	
	k	l
$t_0 = \frac{\bar{y} + b_\phi (P - p)(kP)}{p^*} \exp\left(\frac{p^* - P}{P + p^*}\right)$	1	0
$t_1 = \frac{(\bar{y} + b_\phi (P - p)(P + \rho_{y\phi}))}{p^* + \rho_{y\phi}} \exp\left(\frac{p^* - P}{P + p^* + 2\rho_{y\phi}}\right)$	1	$\beta_2(\phi)$
$t_2 = \frac{(\bar{y} + b_\phi (P - p)(P + C_\phi))}{p^* + C_\phi} \exp\left(\frac{p^* - P}{P + p^* + 2C_\phi}\right)$	1	C_ϕ
$t_3 = \frac{(\bar{y} + b_\phi (P - p)(P + \rho_{y\phi}))}{p^* + \rho_{y\phi}} \exp\left(\frac{p^* - P}{P + p^* + 2\rho_{y\phi}}\right)$	1	$\rho_{y\phi}$
$t_4 = \frac{(\bar{y} + b_\phi (P - p)(\beta_2(\phi)P + C_\phi))}{\beta_2(\phi)p^* + C_\phi} \exp\left(\frac{\beta_2(\phi)(p^* - P)}{\beta_2(\phi)(P + p^*) + 2C_\phi}\right)$	$\beta_2(\phi)$	C_ϕ
$t_5 = \frac{(\bar{y} + b_\phi (P - p)(C_\phi P + \beta_2(\phi)))}{C_\phi p^* + \beta_2(\phi)} \exp\left(\frac{C_\phi(p^* - P)}{C_\phi(P + p^*) + 2\beta_2(\phi)}\right)$	C_ϕ	$\beta_2(\phi)$
$t_6 = \frac{(\bar{y} + b_\phi (P - p)(C_\phi P + \rho_{y\phi}))}{C_\phi p^* + \rho_{y\phi}} \exp\left(\frac{C_\phi(p^* - P)}{C_\phi(P + p^*) + 2\rho_{y\phi}}\right)$	C_ϕ	$\rho_{y\phi}$
$t_7 = \frac{(\bar{y} + b_\phi (P - p)(\rho_{y\phi}P + C_\phi))}{\rho_{y\phi}p^* + C_\phi} \exp\left(\frac{\rho_{y\phi}(p^* - P)}{\rho_{y\phi}(P + p^*) + 2C_\phi}\right)$	$\rho_{y\phi}$	C_ϕ
$t_8 = \frac{(\bar{y} + b_\phi (P - p)(\beta_2(\phi)P + \rho_{y\phi}))}{\beta_2(\phi)p^* + \rho_{y\phi}} \exp\left(\frac{\beta_2(\phi)(p^* - P)}{\beta_2(\phi)(P + p^*) + 2\rho_{y\phi}}\right)$	$\beta_2(\phi)$	$\rho_{y\phi}$
$t_9 = \frac{(\bar{y} + b_\phi (P - p)(\rho_{y\phi}P + \beta_2(\phi)))}{\rho_{y\phi}p^* + \beta_2(\phi)} \exp\left(\frac{\rho_{y\phi}(p^* - P)}{\rho_{y\phi}(P + p^*) + 2\beta_2(\phi)}\right)$	$\rho_{y\phi}$	$\beta_2(\phi)$

3.1 Properties (Bias and MSE) of t_{pi} and t_{qi}

Let S be the set of n -paired observations (y_i, ϕ_i) drawn with SRSWOR from a set of population Ω of size N . To obtain the properties of t_{pi} and t_{qi} the following error terms are defined.

$$e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}, \quad e_1 = \frac{p - P}{P} \text{ Such that, } \bar{y} = (1 + e_0)\bar{Y}, \quad p = (1 + e_1)P,$$

The expectation (E) of the error terms were obtained as in (3.3)

$$\left. \begin{aligned} E(e_0) = E(e_1) = 0, \quad E(e_0^2) = \lambda C_y^2, \\ E(e_1^2) = \lambda C_\phi^2, \quad E(e_0 e_1) = \lambda \rho_{y\phi} C_y C_\phi, \end{aligned} \right\} \quad (3.3)$$

Express t_1 and t_2 in term of error terms e_1 and e_2

$$t_1 = (\bar{Y} + \bar{Y}e_0 - b_\phi e_1 P) \exp\left(\frac{-kP}{2(kP+l)} e_1 \left(1 + \frac{kP}{2(kP+l)} e_1\right)^{-1}\right) \quad (3.4)$$

$$t_2 = (\bar{Y} + \bar{Y}e_0 - b_\phi e_1 P) \left(1 - \frac{kPJ}{(kP+l)} e_1\right)^{-1} \exp\left(\frac{-kPJ}{(kP+l)} e_1 \left(1 - \frac{kPJ}{2(kP+l)} e_1\right)^{-1}\right) \quad (3.5)$$

Simplify (3.4) and (3.5) up to the first order of approximation, to obtain

$$t_1 = (\bar{Y} + \bar{Y}e_0 - b_\phi e_1 P) \exp(-\varpi_1 e_1 + \varpi_1^2 e_1^2) \quad (3.6)$$

$$t_2 = (\bar{Y} + \bar{Y}e_0 - b_\phi e_1 P) (1 - J\varpi_2 e_1)^{-1} \exp(-J\varpi_2 e_1 - J^2 \varpi_1 \varpi_2 e_1^2) \quad (3.7)$$

$$\text{Where } \varpi_1 = \left(\frac{kP}{2(kP+l)}\right), \quad \varpi_2 = \left(\frac{kP}{kP+l}\right) \text{ and } J = \frac{n}{N-n}.$$

Simplify (3.6) and (3.7), to obtain (3.8) and (3.9)

$$t_1 = \bar{Y} + \bar{Y}e_0 - \bar{Y}\varpi_1 e_1 + \frac{3}{2}\varpi_1^2 \bar{Y}e_1^2 - \varpi_1 \bar{Y}e_0 e_1 - b_\phi e_1 P + b_\phi P e_1^2 \quad (3.8)$$

$$t_2 = \bar{Y} + \bar{Y}e_0 - \bar{Y} \left(J^2 \varpi_1 \varpi_2 - \frac{J^2 \varpi_2^2}{2} \right) e_1^2 - b_\phi P e_1 \quad (3.9)$$

Subtract \bar{Y} from both sides of equation (3.8) and (3.9) to obtain (3.10) and (3.11)

$$t_1 - \bar{Y} = \bar{Y}e_0 - (\bar{Y} + b_\phi P) e_1 + \left(\frac{3}{2}\varpi_1^2 \bar{Y} + b_\phi \varpi_2 P\right) e_1^2 - \bar{Y}\varpi_2 e_0 e_1 \quad (3.10)$$

$$t_2 - \bar{Y} = \bar{Y}e_0 - \bar{Y} \left(J^2 \varpi_1 \varpi_2 - \frac{J^2 \varpi_2^2}{2} \right) e_1^2 - b_\phi P e_1 \quad (3.11)$$

Take expectation of (3.10) and (3.11) and apply result of (3.3), the Bias of t_p and t_q are obtained as

$$Bias(t_{pi}) = \lambda \left(\left(\frac{3}{2} \varpi_2 \bar{Y} + b_\phi \varpi_2 P \right) C_\phi^2 - \bar{Y} \varpi_2 \rho_{y\phi} C_y C_\phi \right) \quad i = 1, 2, 3, \dots, 9 \quad (3.12)$$

$$Bias(t_{qi}) = -\bar{Y} \lambda \left(J^2 \varpi_1 \varpi_2 - \frac{J^2 \varpi_2^2}{2} \right) C_\phi^2 \quad i = 0, 1, 2, \dots, 9 \quad (3.13)$$

Square (3.12) and (3.13) up to the first order of approximation, take expectations and apply the result of (3.3), the MSE of t_1 and t_2 are obtained as

$$MSE(t_1) = \bar{Y}^2 \lambda \left(C_y^2 + \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} \right)^2 C_\phi^2 - 2 \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} \right) \rho_{y\phi} C_y C_\phi \right) \quad (3.14)$$

$$MSE(t_2) = \lambda \left(\bar{Y}^2 C_y^2 + b_\phi^2 P^2 C_\phi^2 - 2 \bar{Y} b_\phi P \rho_{y\phi} C_y C_\phi \right) \quad (3.15)$$

3.2 Efficiency comparison of t_{pi} and t_{qi}

The efficiency conditions of the proposed estimators t_{pi} and t_{qi} over some existing estimators of Naik and Gupta (1996), Singh et al. (2007), Singh et al. (2013), Zaman and Kadilar (2019) and Zaman (2020) were established in this section.

$$i. \quad MSE(t_{pi}) - MSE(t_{NG}) < 0 \Rightarrow |\rho_{y\phi}| > C_\phi \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} + 1 \right) / 2C_y \quad (3.16)$$

$$MSE(t_{qi}) - MSE(t_{NG}) < 0 \Rightarrow |\rho_{y\phi}| > C_\phi (b_\phi P + \bar{Y}) / 2\bar{Y}C_y \quad (3.17)$$

$$ii. \quad MSE(t_{pi}) - MSE(t_{STS}) < 0 \Rightarrow |\rho_{y\phi}| > C_\phi \left(2 \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} \right) + 1 \right) / 4C_y \quad (3.18)$$

$$MSE(t_{qi}) - MSE(t_{STS}) < 0 \Rightarrow |\rho_{y\phi}| > C_\phi (2b_\phi P + \bar{Y}) / 4\bar{Y}C_y \quad (3.19)$$

$$iii. \quad MSE(t_{pi}) - MSE(t_{ST})_{\min} < 0 \Rightarrow |\rho_{y\phi}| > C_\phi \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} \right) / C_y \quad (3.20)$$

$$MSE(t_{qi}) - MSE(t_{ST})_{\min} < 0 \Rightarrow |\rho_{y\phi}| > C_\phi b_\phi P \bar{Y} / C_y$$

(3.21)

$$iv. \quad MSE(t_{pi}) - MSE(t_{ZKi}) < 0 \Rightarrow |\rho_{y\phi}| > C_\phi \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} + \theta_i \right) / 2C_y \quad (2.22)$$

$$MSE(t_{qi}) - MSE(t_{ZKi}) < 0 \Rightarrow |\rho_{y\phi}| > C_\phi (b_\phi P + \bar{Y}\theta_i) / 2\bar{Y}C_y \quad (3.23)$$

$$v. \quad MSE(t_{pi}) - MSE(t_{Zi})_{\min} < 0 \Rightarrow |\rho_{y\phi}| > C_\phi \left(\varpi_1 + \frac{b_\phi P}{\bar{Y}} \right) / C_y \quad (3.24)$$

$$MSE(t_{qi}) - MSE(t_{Zi})_{\min} < 0 \Rightarrow |\rho_{y\phi}| > C_\phi b_\phi P / \bar{Y}C_y \quad (3.25)$$

If conditions (3.16), (3.18), (3.20), (3.22), (3.24) and (3.17), (3.19), (3.21), (3.23), (3.25) are satisfied respectively, then the proposed estimators t_{pi} and t_{qi} are more efficient than the of Naik and Gupta (1996), Singh et al. (2007), Singh et al. (2013), Zaman and Kadilar (2019) and Zaman (2020) respectively.

4. Empirical Study

In this section, empirical studies were conducted to examine the superiority of the proposed estimators over some estimators considered in the study.

4.1 Efficiency Comparison

Empirical Study using existing population parameters (Sukhatme and Sukhatme 1970).

Population 1: The data is defined as follows:

$y =$ the number of villages in the circles

$$\phi_i = \begin{cases} 1, & \text{if a circle consist of morethan five villages} \\ 0 & \text{otherwise} \end{cases}$$

$$N = 89, n = 20, \bar{Y} = 3.3596, P = 0.1236, b_\phi = 3.492, C_y = 0.6008, C_\phi = 2.6779, \rho_{y\phi} = 0.766$$

(Sukhatme and Sukhatme 1970).

Population 2: The data is defined as follows:

$y =$ the number of teachers

$$\phi_i = \begin{cases} 1, & \text{if the number of teachers is morethan 60} \\ 0 & \text{otherwise} \end{cases}$$

$$N = 111, n = 30, \bar{Y} = 29.279, P = 0.117, b_\phi = 3.898, C_y = 0.872, C_\phi = 2.758, \rho_{y\phi} = 0.797$$

Table 3 MSEs and PREs of Proposed and Existing Estimators Using Data from Pop. 1

Estimators		MSEs	PREs
Sample Mean	\bar{y}	0.1579298	100
Naik and Gupta (1996)	t_{NG}	2.217078	7.123331
Singh et al. (2007)	t_{STS}	0.4031122	39.17762
Singh et al. (2013)	t_{ST}	0.06526354	241.9878
Zaman and Kadilar (2019)	t_{ZKi}		

t_{ZK1}		0.1404135	112.4748
t_{ZK2}		0.1356671	116.4098
t_{ZK3}		0.09815459	160.899
t_{ZK4}		0.09819807	160.8278
t_{ZK5}		0.1171256	134.838
t_{ZK6}		0.06664636	236.9669
t_{ZK7}		0.1404316	112.4603
t_{ZK8}		0.06548182	241.1811
t_{ZK9}		0.1442428	109.4888
Zaman (2020)	t_Z	0.06526354	241.9878
Propose Estimators		t_{pi}	
t_{11}		0.0661802	238.636
t_{12}		0.06679036	236.456
t_{13}		0.08040544	196.4168
t_{14}		0.08037597	196.4888
t_{15}		0.07114321	221.9886
t_{16}		0.1366728	115.5532
t_{17}		0.0661782	238.6432
t_{18}		0.167143	94.48783
t_{19}		0.06581011	239.978
Proposed Estimator	t_q	0.06526354	241.9878

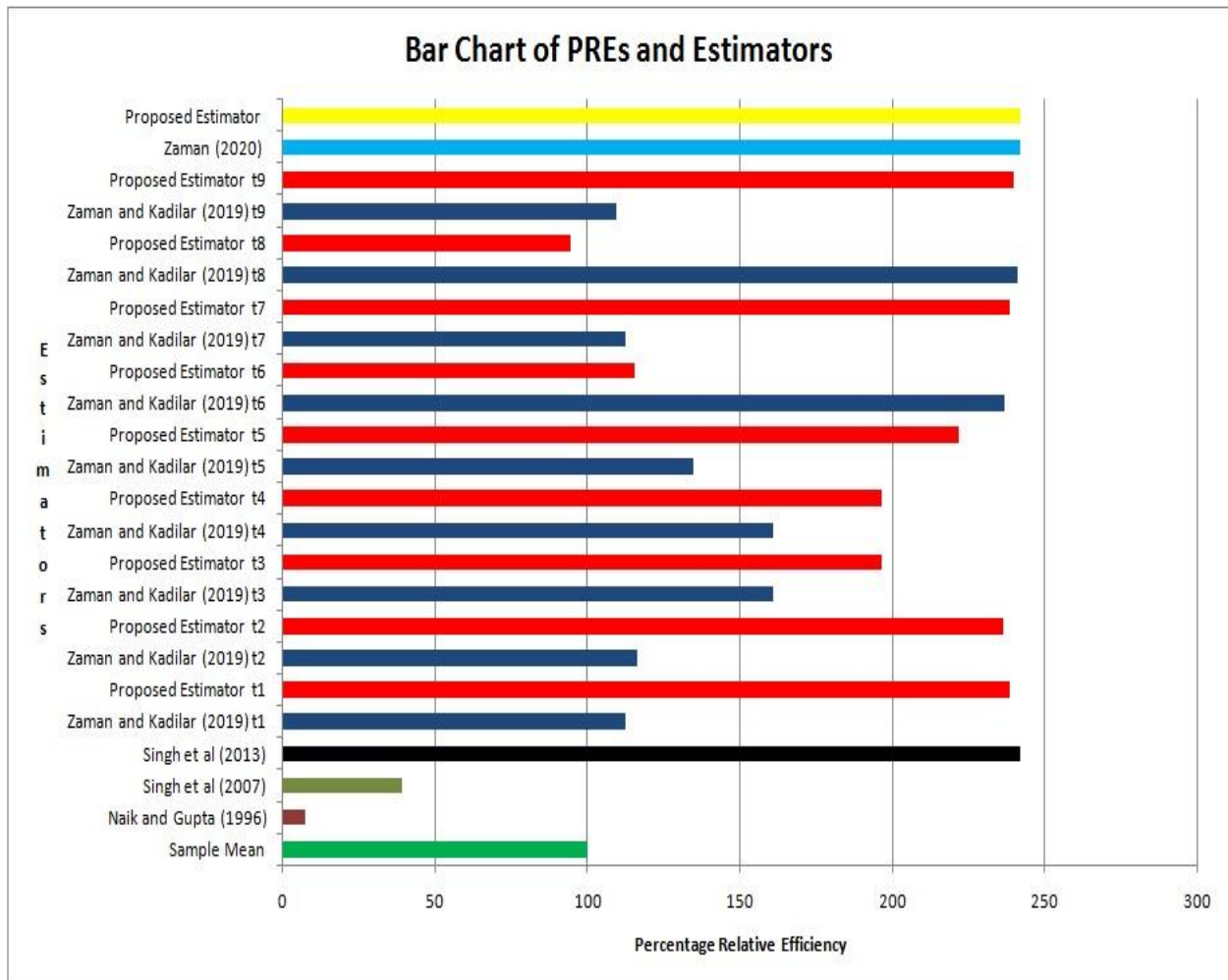


Figure 1: Represent the Bar Chart of PREs and Estimator using population 1

Table 4 MSEs and PREs of Proposed and Existing Estimators Using Data from Pop. 2

Estimators	MSEs	PREs
Sample Mean \bar{y}	15.85573	100
Naik and Gupta (1996) t_{NG}	94.532	16.77287
Singh et al. (2007) t_{STS}	15.54034	102.0295
Singh et al. (2013) t_{ST}	5.784028	274.1296
Zaman and Kadilar (2019) t_{ZKi}		
t_{ZK1}	14.72468	107.6813
t_{ZK2}	14.29484	110.9192
t_{ZK3}	11.38913	139.2181
t_{ZK4}	10.98268	144.3703
t_{ZK5}	13.03175	121.67
t_{ZK6}	7.630395	207.797
t_{ZK7}	14.59097	108.6681
t_{ZK8}	6.561422	241.6508
t_{ZK9}	14.94357	106.1041
Zaman (2020) t_Z	5.784028	274.1296
Propose Estimators t_{pi}		
t_{11}	5.817701	272.5429
t_{12}	5.849699	271.0521
t_{13}	6.4338	246.4443
t_{14}	6.582441	240.8792
t_{15}	6.015808	263.5678
t_{16}	9.077466	174.6713
t_{17}	5.826441	272.1341
t_{18}	11.03681	143.6623
t_{19}	5.805672	273.1076
Proposed Estimator t_q	5.784028	274.1296

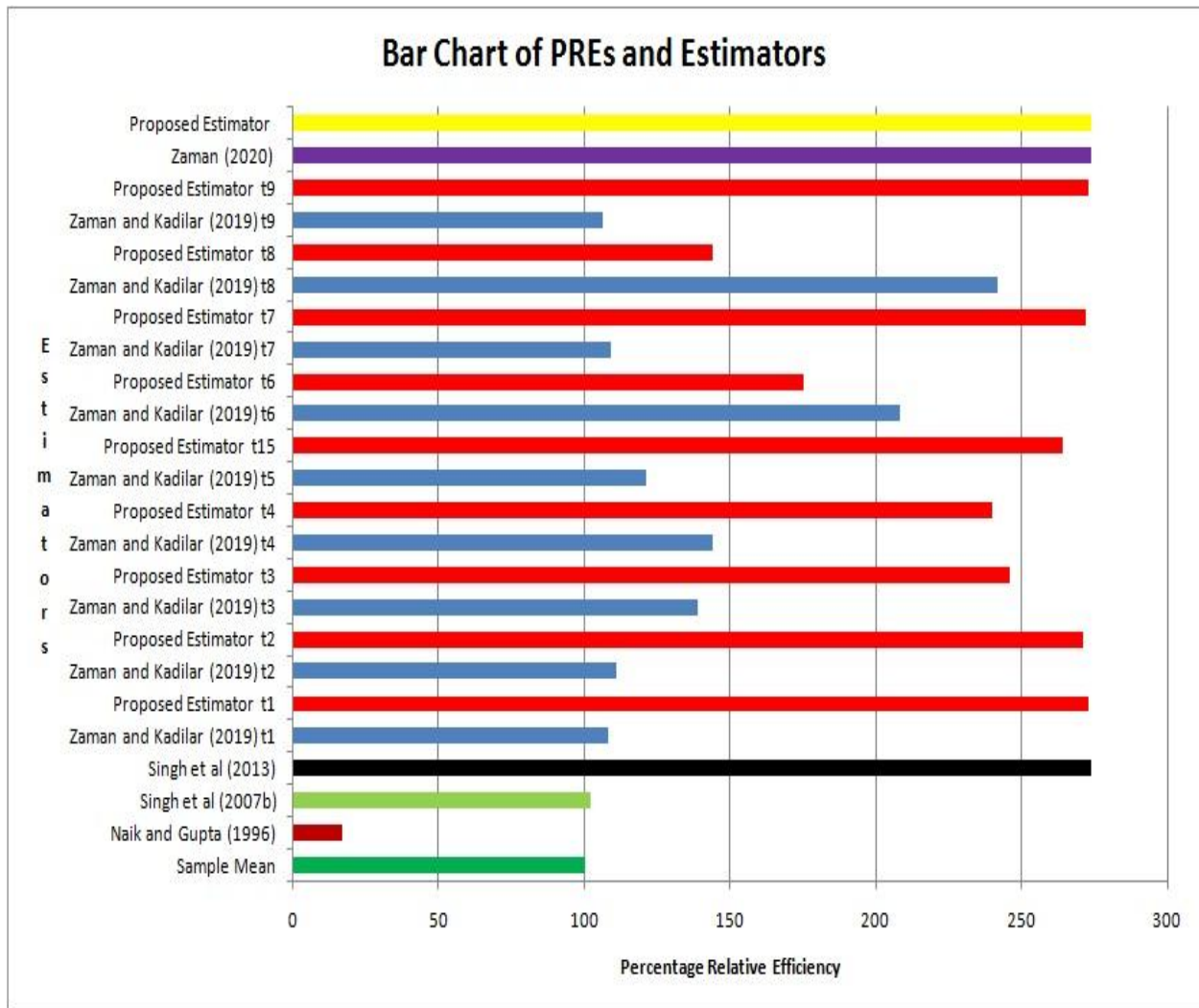


Figure 2: Represent the Bar Chart of PREs and Estimator using population 1

Tables 3 and 4 show the results of MSEs and PREs of the proposed estimators and that of some existing estimators using populations 1 and 2 respectively. The results revealed that the proposed estimators have minimum MSEs and higher PREs compared to conventional and other related estimators considered in this study. This implies that the proposed estimators are more efficient and can produce better estimate of the population parameters than the existing estimators. The results were also presented using Figures 1 and 2.

4.3 Empirical Study using Stimulated data

In this sub-section, stimulation studies were conducted to assess the performance of the proposed estimators over other estimators considered in the study. Data of size 500 units were

generated for study population using functions defined in Table 5 Sample of size 50 were selected by method of Simple Random Sampling without replacement (SRSWOR) 1000 times. The Biases, MSEs and PREs of the considered estimators were computed using (4.1), (4.2) and (4.3).

$$Bias(T) = \frac{1}{1000} \sum_{j=1}^{1000} (T - \bar{Y}) \quad (4.1)$$

$$MSE(T) = \frac{1}{1000} \sum_{j=1}^{1000} (T - \bar{Y})^2 \quad (4.2)$$

$$PREs(T) = \left(\frac{MSE(t_o)}{MSE(T)} \right) \times 100 \quad (4.3)$$

Where T are any of the exiting or proposed estimators.

Table 5: Population used for Simulation Study

Models	Auxiliary variable (x)	Study Variable (y)
1	$X \sim rben(500, 0.7)$	$Y = 10 + 40X + e,$
2		$Y = 10 - 40X + e,$ where, $e \sim (0, 4)$

Table 6: Biases, MSEs and PREs of the Proposed and some Existing Estimators using model 1 with the Correlation ($\rho_{y\phi} = 0.9775665$)

Estimators	Biases	MSEs	PREs	Estimators	Biases	MSEs	PREs
Sample Mean \bar{y}	0.05549	64.07467	100	t_{z8}	0.1808972	2.925655	2190.097
Naik and Gupta (1996) t_{NG}	1.005351	14.26301	449.2368	t_{z9}	0.2185038	2.93286	2184.716
Singh et al. (2007b) t_{STS}	0.110366	8.312005	770.869	Proposed Estimators t_{pi}			
Singh et al. (2013) t_{ST}	268.518	7221.504	0.8872759	t_{31}	0.384611	0.931629	2185.634
Zaman and Kadilar(2019) t_{zki}				t_{32}	0.6210778	3.015386	2124.925
t_{zK1}	-0.21617	41.50172	154.3904	t_{33}	0.5174373	2.973801	2154.639
t_{zK2}	-0.17660	28.8395	222.1768	t_{34}	0.7598552	3.083468	2078.006
t_{zK3}	-0.178769	34.07176	188.058	t_{35}	0.300506	2.911192	2200.977
t_{zK4}	-0.218887	22.60134	283.4994	t_{36}	0.4149634	2.940186	2179.272
t_{zK5}	-0.213759	46.63139	137.4067	t_{37}	0.6145036	3.012516	2126.948
t_{zK6}	-0.195723	39.73171	161.2683	t_{38}	0.6570131	3.031632	2113.538
t_{zK7}	-0.139944	29.15678	219.7591	t_{39}	0.3792218	2.930175	2186.718
t_{zK8}	-0.189597	27.14009	236.0886	Proposed Estimators t_{qi}			
t_{zK9}	-0.219134	41.82048	153.2136	t_{40}	0.04764066	2.877944	2226.404
Zaman (2020) t_{zi}				t_{41}	0.04764457	2.877942	2226.406
t_{z0}	0.3810892	2.970742	2156.857	t_{42}	0.04764409	2.877942	2226.406
t_{z1}	0.2168675	2.932537	2184.957	t_{43}	0.04764436	2.877942	2226.406
t_{z2}	0.180437	2.925648	2190.102	t_{44}	0.04764359	2.877943	2226.405
t_{z3}	0.1879284	2.927082	2189.029	t_{45}	0.04764465	2.877942	2226.406
t_{z4}	0.1908273	2.927097	2189.018	t_{46}	0.04764454	2.877942	2226.406
t_{z5}	0.2465373	2.938672	2180.396	t_{47}	0.04764411	2.877942	2226.406
t_{z6}	0.2083266	2.930878	2186.193	t_{48}	0.04764398	2.877943	2226.405
t_{z7}	0.1805225	2.925674	2190.083	t_{49}	0.04764458	2.877942	2226.406

Table 7: Biases, MSEs and PREs of the Proposed and some Existing Estimators using Model 2 with Correlation ($\rho_{y\phi} = -0.9791406$)

Estimators	Biases	MSEs	PREs	Estimators	Biases	MSEs	PREs
Sample Mean \bar{y}	-0.585156	68.99155	100	t_{Z8}	-683.4372	1024063	0.006737044
Naik and Gupta (199) t_{NG}	0.1426181	3.077269	2241.973	t_{Z9}	-0.6003564	2.982463	2313.241
Singh et al. (2007b) t_{STS}	0.06668405	21.67639	318.2797	Proposed Estimators t_{pi}			
Singh et al. (2013) t_{ST}	-672.4029	6016708	0.001146666	t_{31}	-0.3273392	2.809121	2455.984
Zaman and Kadilar(2019) t_{ZKi}				t_{32}	-0.6951248	2.81572	2391.174
t_{ZK1}	-0.2132775	51.79144	133.2104	t_{33}	2.895604	5.412566	1274.655
t_{ZK2}	-0.052035	42.01574	164.2041	t_{34}	-0.700835	2.946906	2341.152
t_{ZK3}	-8.672273	334.4563	20.62797	t_{35}	-0.241003	2.790299	2472.551
t_{ZK4}	0.01186464	36.86736	187.1345	t_{36}	1.132527	3.089483	2233.11
t_{ZK5}	-0.2885746	55.6611	123.9493	t_{37}	66.30169	4400.433	1.567835
t_{ZK6}	-2.669514	140.9721	48.93988	t_{38}	-252.9175	149222.7	0.04623397
t_{ZK7}	-1039.578	1150628	0.00599599	t_{39}	0.8997042	2.959664	2331.061
t_{ZK8}	-1395.647	4542102	0.001518934	Proposed Estimators t_{qi}			
t_{ZK9}	-2.11172	123.0889	56.05017	t_{40}	0.03228859	2.761002	2498.787
Zaman (2020) t_{Zi}				t_{41}	-0.1555318	2.67584	2498.791
t_{Z0}	0.1190468	2.844141	2425.743	t_{42}	0.03228801	2.760998	2498.79
t_{Z1}	0.2357007	2.851608	2419.391	t_{43}	0.0322362	2.760946	2498.837
t_{Z2}	0.2598529	2.855944	2415.718	t_{44}	0.03228811	2.760999	2498.79
t_{Z3}	-4.228685	8.862594	778.4578	t_{45}	0.03228787	2.760998	2498.791
t_{Z4}	0.2539563	2.855833	2415.812	t_{46}	0.03228559	2.760994	2498.794
t_{Z5}	0.2159155	2.848987	2421.617	t_{47}	-3.955063	939.421	7.34405
t_{Z6}	-0.890612	3.13349	2201.748	t_{48}	0.01424415	2.771623	2489.211
t_{Z7}	-631.3627	334987.7	0.02059525	t_{49}	0.03228681	2.760996	2498.792

Table 6 and 7 show the results of Biases, MSEs and PREs of the proposed estimators and that of the existing estimators using stimulated data. The result revealed that the proposed estimators has minimum MSEs and higher PREs compared to conventional and other related estimators considered in this study. This implies that the proposed estimators are more efficient and can produce better estimate of the population parameters than the existing estimators.

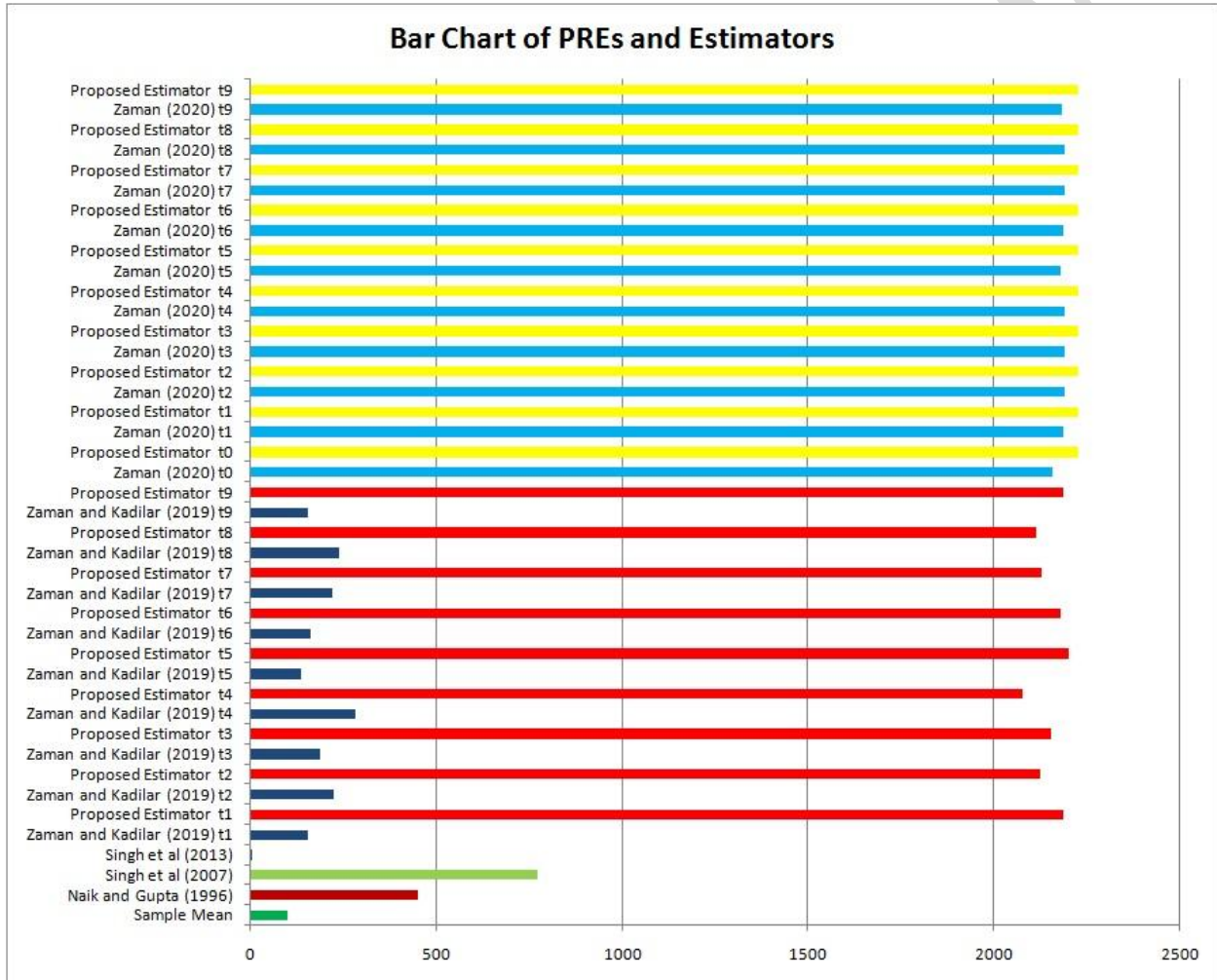


Figure 3: Represent the Bar Chart of PREs and Estimator with the correlation ($\rho_{y\phi} = 0.9775665$)

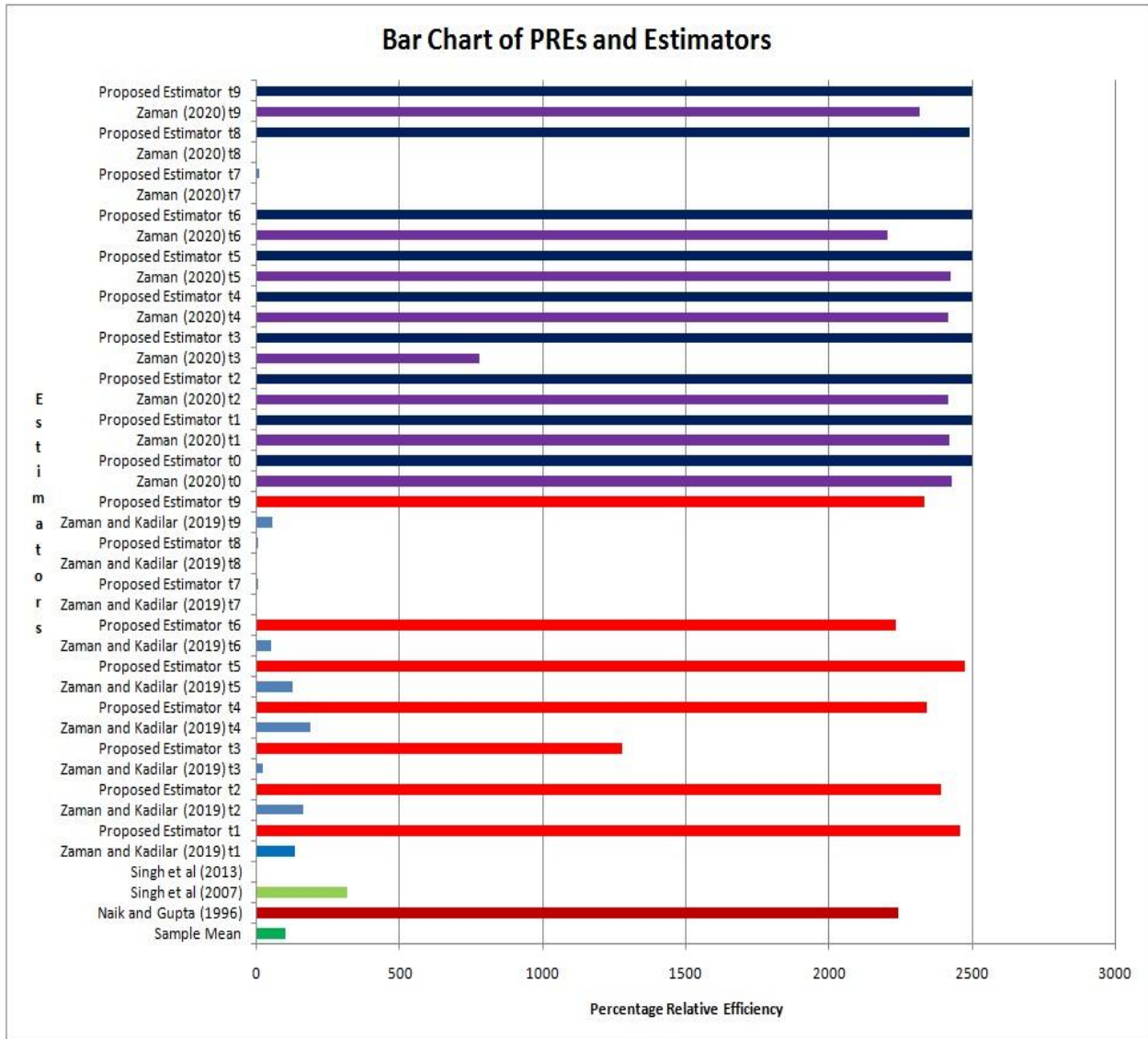


Figure 4: Represent the Bar Chart of PREs and Estimator with the correlation ($\rho_{y\phi} = -0.9791406$)

5. Conclusion

From the results of the empirical studies, it revealed that the suggested estimators have minimum MSEs compared to other competing estimators considered in literature. Hence, the suggested estimators demonstrated high level of efficiency over the other estimators considered in the study. In the other words, the suggested estimators have higher chance of producing estimate that is closer to the true value of the population mean than other estimators considered in the literature of this study and therefore recommended for use in practice.

REFERENCES

- [1] Abu-Dayeh WA, Ahmed MS, Ahmed RAm Muttlak, HA. (2003). Some Estimators of a Finite Population Mean Using Auxiliary Information. *Appl. Math Compute*, 13(9), 287–298.
- [2] Audu A., Singh R., and Khare S. (2021a). New Regression-Type Compromised Imputation Class of Estimators with known Parameters of Auxiliary Variable. *Communication in Statistics-Simulation and Computation*, 1-13. [DOI:10.1080/03610918.2021.1970182](https://doi.org/10.1080/03610918.2021.1970182)
- [3] Audu A. and Singh R. V. K. (2021b). Exponential-Type Regression Compromised Imputation Class of Estimators. *Journal of Statistics and Mgt Systems- Taylor and Francis*, 24(6), 1253-1266. [DOI: 10.1080/09720510.2020.1814501](https://doi.org/10.1080/09720510.2020.1814501)
- [4] Audu A., Singh R., and Khare S., Dauran N. S. (2021c). Almost unbiased estimators for population mean in the presence of non-response and measurement error. *Journal of Statistics and Management Systems*, 24:3, 573-589, [DOI: 10.1080/09720510.2020.1759209](https://doi.org/10.1080/09720510.2020.1759209)
- [5] Audu A., Ishaq O. O., Abubakar A., Akintola K. A., Isah U., Rashida A., and Muhammad S. (2021d). Regression-type Imputation Class of Estimators using Auxiliary Attribute. *Asian Research Journal of Mathematics*, 17(5): 1-13. [DOI: 10.9734/ARJOM/2021/v17i530296](https://doi.org/10.9734/ARJOM/2021/v17i530296)
- [6] Audu A., Danbaba A, Ahmad S. K., Musa N, Shehu A, Ndatsu A. M. and Joseph A. O. (2021e). On the Efficient of Almost Unbiased Mean Imputation When the Population Mean of Auxiliary Variable is Unknown. *Asian Journal of Probability and Statistics*, 15(4): 235-250. [DOI:10.9734/AJPAS/2021/v15i430377](https://doi.org/10.9734/AJPAS/2021/v15i430377)
- [7] Audu A., Yunusa M. A., Ishaq O. O., Lawal M. K., Rashida A., Muhammad A. H., Bello A. B., Hairullahi M. U. and Muili J. O. (2021f). Difference-Cum-Ratio Estimators for Estimating Finite Population Coefficient of Variation in Simple Random Sampling. *Asian Journal of Probability and Statistics*, 13(3): 13-29. [DOI: 10.9734/AJPAS/2021/v13i330308](https://doi.org/10.9734/AJPAS/2021/v13i330308)
- [8] Audu A., Yunusa M. A., Zoramawa A.B., Buda S., and Ran Singh R. V. K. (2021g). Exponential-Ratio-Type Imputation Class of Estimators using Nonconventional Robust

Measures of Dispersions. *Asian Journal of Probability and Statistics*, 15(2): 59-74. DOI: [10.9734/AJPAS/2021/v15i230351](https://doi.org/10.9734/AJPAS/2021/v15i230351)

- [9] Audu, A., Muili, J. O., Agwamba, E. N., Erinola, Y. A., Yunusa, M. A., and Hamzat, M. A. (2020a). Modified Ratio-Cum-Product Estimators of Population Mean Using Two Auxiliary Variables. *Asian Journal of Research in Computer Science*, 6(1), 55-65.
- [10] Audu, A., Ishaq O. O., Muili, J., Zakari Y., Ndatsu A. M., and Muhammed S. (2020b): On the Efficiency of Imputation Estimators using Auxiliary Attribute. *Continental J. Applied Sciences*, 15 (1), 1-13. DOI: [10.5281/zenodo.3721046](https://doi.org/10.5281/zenodo.3721046)
- [11] Audu A., Olawoyin I. O., Zakari Y., Wisdom D. D., Muili J., and Ndatsu A. M. (2020c): Regression-cum-exponential ratio imputation class of estimators of population mean in the presence of non-response. *Science Forum Journal of Pure and Applied Science*, 20, 58-63. DOI: <http://dx.doi.org/10.5455/sf.71109>
- [12] Audu A., and Adewara A. A. (2017): A Factor-Type Estimator with Known Functions of Auxiliary variable under Single-Phase Sampling. *ABACUS, Mathematics Science Series*, 44(1), 171-177. ISSN: 0001-3099.
- [13] Audu A., Adewara A. A. and Singh R. V. K. (2016a): Class of Ratio Estimators with known Functions of Auxiliary Variable for Estimating Finite Population Variance. *Asian Journal of Mathematics and Computer research*, 12(1), Pp. 63-70.
- [14] Audu A., Adewara A. A., and Singh R. V. K. (2016b). Class of Ratio Estimators with known Functions of Auxiliary Variable for Estimating Finite Population Variance. *Asian Journal of Mathematics and Computer research*, 12(1), Pp. 63-70.
- [15] Audu A., Adewara A. A., and Singh R. V. K. (2016c). Ratio and Product type Exponential Estimators of Population Variance under transformed Sample Information of Study and Supplementary Variables. *Asian Journal of Mathematics and Computer research*, 11(3), Pp. 175-183.
- [16] Audu A., and Singh, R. V. K. (2015): Improved Exponential Type Estimators for Estimating Population Variance Survey Sampling. *International Journal of Advance Research*, 3(4), Pp. 1-16.
- [17] Cochran, W. G. (1977). *Sampling techniques*. New York, NY: John Wiley and Sons.
- [18] Hartley, H. O., and Ross, A. (1954). Unbiased ratio estimators. *Nature*, 17(4), 270- 271.
- [19] Kadilar, C., and Cingi, H. (2005). A new estimator using two auxiliary variables. *Applied Mathematics and Computation*, 16(2), 901–908.
- [20] Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N., and Smarandache, F. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s). *Far East Journal of Theoretical Statistics*, 22(2), 181–191.

- [21] Lu, J. (2013). The Chain Ratio Estimator and Regression Estimator with linear Combination of Two Auxiliary Variables. PLoS ONE, 8(11), 890-899.
- [22] Lu, J., and Yan, Z. (2014). A Class of Ratio Estimators of a Finite Population Mean Using Two Auxiliary Variables. PLoS ONE, 9(2), 890-899.
- [23] Naik, V. D., and P. C. Gupta. (1996). A note on estimation of mean with known population proportion of an auxiliary character. Journal of the Indian Society of Agricultural Statistics, 48(2), 151-8.
- [24] Neyman, J. (1938). Contribution to the theory of sampling human population. Journal of the American Statistical Association, 33(201), 101-16. doi:10.2307/22791s17.
- [25] Perri, P. F. (2007). Improved ratio-cum-product type estimators. Statist Trans, 8(2), 51-69.
- [26] Quenouille, M. H. (1956). Notes on bias in estimation. Biometrika, 4(3), 353-360.
- [27] Singh R. V. K., Saleh B. B., and Audu A. (2021). Modified Classes of Regression Type Exponential Estimators of Population Mean. Journal of Scientific Research, 65(5), 244-251. DOI: 10.37398/JSR.2021.650530
- [28] Sharma, P., and Singh, R. (2014). Improved Ratio Type Estimators Using Two Auxiliary Variables under Second Order Approximation. Mathematical Journal of Interdisciplinary Sciences, 2(2), 193-204.
- [29] Singh, M. P. (1965). On the estimation of ratio and product of the population parameters. Sankhya B, 2(7), 231-328.
- [30] Singh, M. P. (1967). Ratio cum product method of estimation. Metrika 1(2), 34-42.
- [31] Singh R, Mishra P, Audu A., and Khare S. (2020): Exponential Type Estimator for Estimating Finite Population Mean. Int. J. Comp. Theo. Stat., 7(1), 37-41
- [32] Singh, R., and Kumar, M. (2011). A note on transformations on auxiliary variable in survey sampling. MASA, 6(1), 17-19.
- [33] Singh, R., Chauhan, P., and Sawan, N. (2013). Ratio-Product type exponential for estimating finite population mean using information on auxiliary attribute. Journal of Statistics and Management Systems, 81(7), 679-1067.
- [34] Singh, R., Cauhan, P., Sawan, N., & Smarandache, F. (2007). Auxiliary information and a priori values in construction of improved estimators. Renaissance High Press.
- [35] Sukhatme, P.V. and Sukhatme, B.V. (1970) Sampling Theory of Surveys with Applications, Iowa State University Press, Ames, USA, 1970.
- [36] Tailor, R., Chouhan, S., and Garg, N. (2012). A ratio-cum-product estimator of population

mean in stratified random sampling using two auxiliary variables. *STATISTICA*, anno LXXII, n. 3.

- [37] Verma, H. K., Sharma, P. and Singh, R. (2015). Some Families of Estimators Using Two auxiliary variables in stratified random sampling. *Revista Investigacion Operacional*, 36(2), 140-150.
- [38] Yunusa M. A., Audu A., Musa N., Beki D. O., Rashida A., Bello A. B., and Hairullahi M. U. (2021). Logarithmic Ratio-Type Estimator of Population Coefficient of Variation. *Asian Journal of Probability and Statistics*, 14(2): 13-22. [DOI: 10.9734/AJPAS/2021/v14i230323](https://doi.org/10.9734/AJPAS/2021/v14i230323)
- [39] Zaman, T., and Kadilar, C., (2019). Novel family of exponential estimators using information of auxiliary attribute, *Journal of Statistics and Management Systems*, 34(7), 978-1078.
- [40] Zaman, T. (2020). Generalized exponential estimators for the finite population mean. *Statistics in transition*, 21(1), 159-168.
- [41] Zaman T., Dunder E., Audu A., Alilah D. A., Shahzad U. and Hanif M. (2021). Robust Regression-Ratio-Type Estimators of the Mean Utilizing Two Auxiliary Variables: A Simulation Study. *Hindawi Mathematical Problems in Engineering*, 1-9. <https://doi.org/10.1155/2021/6383927>