A Class of Ratio-Type Estimator Using Two Auxiliary Variables For Estimating The Population Mean With Some Known Population Parameters

Toluwalase Janet Akingbade
Department of Statistics
University of Nsukka, Nigeria
akingbadejane@gmail.com

Fabian C. Okafor
Department of Statistics
University of Nigeria, Nsukka, Nigeria
fabian.okafor@gmail.com

Abstract

In this paper, we have suggested a class of ratio type estimators with a linear combination using two auxiliary variables with some known population mean of the study variable. The bias and the mean square error of the proposed estimators are derived. We identified sub-members of the class of ratio type estimators. The theoretical condition for which the the proposed the proposed estimators perform better than the sample mean, Olkin (1958) multivariate ratio, classical linear regression estimator, Singh(1965), Mohanty (1967) and Swain (2012) are derived. From the analysis, it is observed that the proposed estimators perform better than the sample mean and other existing ratio type estimators considered in this study.

Keywords: Bias; Two auxiliary variables; Mean square error; Ratio-type estimator

1. Introduction

It is well known that efficient use of auxiliary variable improves the performance of ratio estimators. Therefore, with the aim of improving on the classical ratio estimator by (Cochran 1977, p151), many authors have suggested some ratio type estimators with known population parameters using single auxiliary variable. For example Tripathi and Khare (1994), Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999), Singh and Tailor (2004), Kadilar and Cingi (2004,2006), Singh and Kakran (1993),Singh and Tailor (2003), Yan and Tian (2010), Subramani and Kumarapandiyan (2012a,2012b, 2012c, 2013), Jelani et al (2013), Raja et al (2017) and Abid et al (2016) have proposed some ratio type estimators based on single auxiliary variable and they showed that their estimators were more efficient than the classical ratio estimator in some cases.

Sometimes, we may have prior information on two or more auxiliary variables that can be used to improve the estimation of the population mean of the study variable. Olkin (1958) was the first author to propose ratio estimator with multi-auxiliary variables to improve estimation. Authors like Abu-Dayyeh et al (2003), Kadilar and Cingi (2005), Lu and Yan (2014) have discussed extensively on ratio estimation. Singh (1965, 1967), Perri (2007), Swain (2012) and Mohanty (1967) also suggested some ratio cum-product type estimators and regression cum-ratio estimator for estimating the population mean using two auxiliary variables. In this study, we shall propose a class of ratio estimator with availability of information on two auxiliary variables with known population parameters on one of the auxiliary variables.

Consider a finite population $S = (s_1, s_2, \ldots s_N)$ of size N. Let Y be the study variable, X and Z be the auxiliary variables, respectively, taking values (y_i, x_i, z_i) on the i[th] unit $S_i(i = 1,2, \ldots, N)$ of the population. Let $(\bar{Y}, \bar{X}, \bar{Z})$ be the population means of (y, x, z), respectively. It is assumed that the population means $\bar{X}$ and $\bar{Z}$of the auxiliary variables are known. For estimating the population mean $\bar{Y}$ of the study variable y, a simple random sample of size *n* is selected without replacement from the population *S*

The usual sample mean is

$$t_0 = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{1}$$

and the variance of the unbiased estimator of t_0 is given as

$$V(t_0) = f_1 \bar{Y}^2 C_y^2 \tag{2}$$

Olkin (1958) proposed the traditional multivariate ratio estimator using two auxiliary variables X and Z, as follows

$$t_{MR} = k_1 \bar{y}\frac{\bar{X}}{\bar{x}} + k_2 \bar{y}\frac{\bar{Z}}{\bar{z}} \tag{3}$$

where k_1 and k_2 are the weights that satisfy the condition $k_1 + k_2 = 1$

The MSE of t_{MR} is given by

$$\text{MSE}(t_{MR}) = f_1 \bar{Y}^2 \left[C_y^2 + k_1^2 C_x^2 + k_2^2 C_z^2 - 2k_1 \rho_{xy} C_y C_x - 2k_2 \rho_{yz} C_y C_z + k_1 k_2 \rho_{xz} C_x C_z\right] \tag{4}$$

The optimum values of k_1 and k_2 are given by

$$k_1^* = \frac{C_z^2 + \rho_{yx} C_y C_x - \rho_{yz} C_y C_z - \rho_{xz} C_x C_z}{C_x^2 + C_z^2 - 2\rho_{xz} C_x C_z}, k_2^* = 1 - k_1^*$$

$$\text{MSE}(t_{MR})_{opt} = f_1 \bar{Y}^2 \left[C_y^2 + k_1^{*2} C_x^2 + k_2^{*2} C_z^2 - 2k_1^* \rho_{xy} C_y C_x - 2k_2^* \rho_{yz} C_y C_z + k_1^* k_2^* \rho_{xz} C_x C_z\right] \tag{5}$$

where $\rho_{xy} = \frac{S_{xy}}{S_x S_y}$, $\rho_{xz} = \frac{S_{xz}}{S_x S_z}$, $\rho_{xy} = \frac{S_{zy}}{S_z S_y}$ are the population correlation coefficients between their respective subscripts. $C_y^2 = \frac{S_y^2}{\bar{Y}^2}$, $C_x^2 = \frac{S_x^2}{\bar{X}^2}$,

$S_y^2 = (N-1)^{-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ and $S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$,

$S_{xy} = (N-1)^{-1}\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})$, $f_1 = \frac{1-f}{n}$, $f = \frac{n}{N}$ (Sampling fraction)

Kadilar and Cingi (2005) presented the classical regression estimator using two auxiliary variables

$$t_{2lr} = \bar{y} + b_{xy}(\bar{X} - \bar{x}) + b_{zy}(\bar{Z} - \bar{z}) \tag{6}$$

The MSE of t_{2lr} is given by

$$\mathrm{MSE}(t_{2lr}) = f_1\bar{Y}^2C_y^2(1 - \rho_{xy}^2 - \rho_{zy}^2 + 2\rho_{xy}\rho_{zy}\rho_{xz}) \tag{7}$$

where $b_{xy} = \frac{s_{xy}}{s_x^2}$ and $b_{zy} = \frac{s_{zy}}{s_z^2}$ are the sample regression coefficients of their respective subscripts.

Singh (1965) suggested a chain ratio estimator given by $t_S = \frac{\bar{y}}{\bar{x}}\bar{X}\frac{\bar{y}}{\bar{z}}\bar{Z}$ (8)

The MSE of t_S is given by

$$\mathrm{MSE}(t_S) = f_1\bar{Y}^2\left(C_y^2 + C_x^2 + C_z^2 - 2\rho_{xy}C_yC_x - 2\rho_{yz}C_yC_z + 2\rho_{xz}C_xC_z\right) \tag{9}$$

Mohanty (1967) proposed ratio estimator using linear combination of the study variable and one of the auxiliary variables for estimating the population mean of the study variable.

$$t_M = \frac{\left(\bar{y} + b_{xy}(\bar{X} - \bar{x})\right)\bar{Z}}{\bar{z}} \tag{10}$$

$$\mathrm{MSE}(t_M) = f_1\bar{Y}^2\left(C_y^2\left(1 - \rho_{xy}^2\right) + C_z^2 - 2\rho_{yz}C_yC_z + 2\rho_{xz}\rho_{xy}C_zC_y\right) \tag{11}$$

Swain (2012) further proposed a generalized form of estimator proposed by Mohanty (1967), given by

$$t_{SW} = \frac{\left(\bar{y} + k_3(\bar{X} - \bar{x})\right)\bar{Z}}{\bar{z}} \tag{12}$$

The optimum MSE of t_{SW} at $k_3$optimum is given by

$$\mathrm{MSE}(t_{SW}) = f_1\bar{Y}^2\left((C_y^2 - 2C_{yz} + C_z^2) - (\rho_{xy}C_y - \rho_{xz}C_z)^2\right) \tag{13}$$

2. The proposed Estimator

Motivated by Kadilar and Cingi (2004, 2006) and Subramani and Kumarapandiyan (2013), we propose a class of ratio estimator for estimating the population mean in simple random sampling using linear combination of two auxiliary variables in single phase sampling with assumed known population means of the auxiliary variables as follows:

$$t_p = \frac{\left(\bar{y} - t_1(\bar{x}^\gamma - \bar{X}^\gamma)\right)}{\left(A\bar{x} + G - t_2(\bar{z}^\gamma - \bar{Z}^\gamma)\right)^\alpha}\left(A\bar{X} + G\right)^\alpha \tag{14}$$

Let A and G be either real values or assumed known population parameters of the auxiliary variable X such as skewness $(\beta_{1(x)})$, kurtosis $(\beta_{2(x)})$, coefficient of variation(C_x), first deciles$(D_{1(x),})$, second decile$(D_{2(x)})$, …, tenth decile $(D_{10(x),)}$, ρ_{xy}

$0<\gamma\leq 1$ and $t_1$and t_2 are the known constants. The scalar α takes values -1, (for product-type estimator) and + 1 (for ratio-type estimator).

To obtain the Bias and MSE of t_p, up to the first order of approximation, let us define

$\bar{y}=\bar{Y}(1+\Delta_y)$, $\bar{x}=\bar{X}(1+\Delta_x)$, $\bar{z}=\bar{Z}(1+\Delta_z)$ $\bar{x}^\gamma=\bar{X}^\gamma(1+\Delta_x)^\gamma$, $\bar{z}^\gamma=\bar{Z}^\gamma(1+\Delta_z)^\gamma$

Such that $E(\Delta_x)=E(\Delta_y)=E(\Delta_z)=0, E(\Delta_x^2)=f_1C_x^2, E(\Delta_y^2)=f_1C_y^2, E(\Delta_z^2)=f_1C_z^2$

$E(\Delta_x\Delta y)=f_1\rho_{xy}C_xC_y$, $E(\Delta_{zy})=f_1\rho_{zy}C_zC_y$, $E(\Delta_{xz})=f_1\rho_{xz}C_xC_z$

$$t_p=\frac{\left(\bar{Y}+\bar{Y}\Delta_y-t_1\left(\bar{X}^\gamma(1+\Delta_x)^\gamma-\bar{X}^\gamma\right)\right)}{\left[A\bar{X}+A\bar{X}\Delta_x+G-t_2\left(\bar{Z}^\gamma(1+\Delta_z)^\gamma-\bar{Z}^\gamma\right)\right]^\alpha}\left(A\bar{X}+G\right)^\alpha \tag{15}$$

We assume $|\Delta_x|<1$, $|\Delta_z|<1$, the expressions $(1+\Delta_z)^\gamma$ and $(1+\Delta_x)^\gamma$ are expandable to a convergent infinite series using binomial expansion. Retaining the terms $\Delta's$ up to power 2, we have

$$t_p=[\bar{Y}+\bar{Y}\Delta_y-t_1\bar{X}^\gamma\gamma\Delta_x-t_1\gamma(\gamma-1)\bar{X}^\gamma\Delta_x^2/2][1+(\lambda\Delta_x-\frac{t_2\bar{Z}^\gamma\gamma\Delta_z}{A\bar{X}+G}-\frac{t_2\gamma(\gamma-1)\bar{Z}^\gamma\Delta_z^2}{2(A\bar{X}+G)})]^{-\alpha}$$

$$t_p=[\bar{Y}+\bar{Y}\Delta_y-t_1\bar{X}^\gamma\gamma\Delta_x-\frac{t_1\gamma(\gamma-1)\bar{X}^\gamma\Delta_x^2}{2}-\alpha\lambda\bar{Y}\Delta_x-\alpha\lambda\bar{Y}\Delta_x\Delta_y+t_1\alpha\lambda\gamma\bar{X}^\gamma\Delta_x^2+$$
$$\frac{t_2\alpha\gamma\bar{Z}^\gamma\bar{Y}\Delta_z}{A\bar{X}+G}+\frac{t_2\alpha\gamma\bar{Z}^\gamma\bar{Y}\Delta_y\Delta_z}{A\bar{X}+G}-\frac{t_1t_2\alpha\gamma^2\bar{Z}^\gamma\bar{X}^\gamma\Delta_x\Delta_z}{A\bar{X}+G}+\frac{t_2\alpha\gamma(\gamma-1)\bar{Z}^\gamma\bar{Y}\Delta_z^2}{2(A\bar{X}+G)}+$$
$$\frac{\alpha(\alpha+1)\lambda^2\bar{Y}\Delta_x^2}{2}+\frac{t_2^2\alpha(\alpha+1)\gamma^2\bar{Z}^{2\gamma}\bar{Y}\Delta_z^2}{2(A\bar{X}+G)^2}-\frac{t_2\alpha(\alpha+1)\gamma\lambda\bar{Z}^\gamma\bar{Y}\Delta_x\Delta_z}{A\bar{X}+G}]$$

$$t_p-\bar{Y}=[\bar{Y}\Delta_y-t_1\bar{X}^\gamma\gamma\Delta_x-\frac{t_1\gamma(\gamma-1)\bar{X}^\gamma\Delta_x^2}{2}-\alpha\lambda\bar{Y}\Delta_x-\alpha\lambda\bar{Y}\Delta_x\Delta_y+t_1\alpha\lambda\gamma\bar{X}^\gamma\Delta_x^2+$$
$$\frac{t_2\alpha\gamma\bar{Z}^\gamma\bar{Y}\Delta_z}{A\bar{X}+G}+\frac{t_2\alpha\gamma\bar{Z}^\gamma\bar{Y}\Delta_y\Delta_z}{A\bar{X}+G}-\frac{t_1t_2\alpha\gamma^2\bar{Z}^\gamma\bar{X}^\gamma\Delta_x\Delta_z}{A\bar{X}+G}+\frac{t_2\alpha\gamma(\gamma-1)\bar{Z}^\gamma\bar{Y}\Delta_z^2}{2(A\bar{X}+G)}+$$
$$\frac{\alpha(\alpha+1)\lambda^2\bar{Y}\Delta_x^2}{2}+\frac{t_2^2\alpha(\alpha+1)\gamma^2\bar{Z}^{2\gamma}\bar{Y}\Delta_z^2}{2(A\bar{X}+G)^2}-\frac{t_2\alpha(\alpha+1)\gamma\lambda\bar{Z}^\gamma\bar{Y}\Delta_x\Delta_z}{A\bar{X}+G}] \tag{16}$$

The bias is obtained by taking the expectation in (16)

$$B(t_p)=f_1(-\frac{t_1\gamma(\gamma-1)\bar{X}^\gamma C_x^2}{2}-\alpha\lambda\bar{Y}\rho_{xy}C_xC_y+t_1\alpha\lambda\gamma\bar{X}^\gamma C_x^2+\frac{t_2\alpha\gamma\bar{Z}^\gamma\bar{Y}\rho_{zy}C_yC_z}{A\bar{X}+G}-$$
$$\frac{t_1t_2\alpha\gamma^2\bar{Z}^\gamma\bar{X}^\gamma\rho_{xz}C_xC_z}{A\bar{X}+G}+\frac{t_2\alpha\gamma(\gamma-1)\bar{Z}^\gamma C_z^2}{2(A\bar{X}+G)}+\frac{\alpha(\alpha+1)\lambda^2\bar{Y}C_x^2}{2}+$$
$$\frac{t_2^2\alpha(\alpha+1)\gamma^2\bar{Z}^{2\gamma}\bar{Y}C_z^2}{2(A\bar{X}+G)^2}-\frac{t_2\alpha(\alpha+1)\gamma\lambda\bar{Z}^\gamma\bar{Y}\rho_{xz}C_xC_z}{A\bar{X}+G})$$

Squaring both side of (16) and taking the expectation to the first degree of approximation. We get the mean square error of (t_p)

$$MSE(t_p) = E[t_p - \bar{Y}]^2 = f_1[\bar{Y}^2 C_Y^2 + t_1^2\gamma^2 \bar{X}^{2\gamma} C_x^2 + \alpha^2\lambda^2\bar{Y}^2 C_x^2 + \frac{t_2^2\alpha^2\gamma^2\bar{Z}^{2\gamma}\bar{Y}^2 C_z^2}{K^2} - 2t_1\gamma\bar{X}^{\gamma}\bar{Y}\rho_{xy}C_xC_y - 2\alpha\lambda\bar{Y}^2\rho_{xy}C_xC_y + \frac{2t_2\alpha\gamma\bar{Z}^{\gamma}\bar{Y}^2\rho_{zy}C_zC_y}{K} + 2t_1\alpha\lambda\gamma\bar{X}^{\gamma}\bar{Y}C_x^2 - \frac{2t_1t_2\alpha\gamma^2\bar{X}^{\gamma}\bar{Z}^{\gamma}\bar{Y}\rho_{xz}C_xC_z}{K} - \frac{2t_2\alpha^2\gamma\lambda\bar{Z}^{\gamma}\bar{Y}^2\rho_{xz}C_xC_z}{K}] \tag{17}$$

$$MSE(t_p) = f_1(\bar{Y}^2C_y^2 + t_1^2q_1^2C_x^2 + \alpha^2\lambda^2\bar{Y}^2C_x^2 + t_2^2\alpha^2q_3^2\bar{Y}^2C_z^2 - 2t_1q_1\bar{Y}\rho_{xy}C_xC_y - 2\alpha\lambda\bar{Y}^2\rho_{xy}C_xC_y + 2t_2\alpha q_3\bar{Y}^2\rho_{zy}C_zC_y + 2t_1\alpha\lambda q_1\bar{Y}C_x^2 - 2t_1t_2\alpha q_4\bar{Y}\rho_{xz}C_xC_z - 2t_2\alpha^2\lambda q_3\bar{Y}^2\rho_{xz}C_xC_z) \tag{18}$$

where $K = A\bar{X} + G, q_1 = \gamma\bar{X}^{\gamma}, q_2 = \gamma\bar{Z}^{\gamma}, q_3 = \frac{q_2}{K}$ and $q_4 = \frac{\gamma^2\bar{X}^{\gamma}\bar{Z}^{\gamma}}{K}$

To obtain t_1 and t_2 optimum, we differentiate $MSE(t_p)$ with respect to t_1 and t_2

Differentiating equation (18) w.r.t. t_1 and t_2 we get the optimum values of t_1 and t_2 respectively as

$$t_{1_0} = \frac{\bar{Y}C_y(\rho_{xy} - \rho_{xz}\rho_{zy}) - \alpha\lambda\bar{Y}C_x(1-\rho_{xz}^2)}{\gamma\bar{X}^{\gamma}C_x(1-\rho_{xz}^2)} \tag{19}$$

$$t_{2_0} = -\frac{C_y(\rho_{zy} - \rho_{xy}\rho_{xz})(A\bar{X}+G)}{\alpha\gamma\bar{Z}^{\gamma}C_z(1-\rho_{xz}^2)} \tag{20}$$

Let $l_1 = \frac{\bar{Y}C_y(\rho_{xy} - \rho_{xz}\rho_{zy})}{C_x(1-\rho_{xz}^2)}$ and $l_2 = \frac{C_y(\rho_{zy} - \rho_{xy}\rho_{xz})}{C_z(1-\rho_{xz}^2)}$ then $t_{1_0} = \bar{Y}\left(\frac{l_1 - \alpha\lambda}{q_1}\right)$

$$t_{2_0} = -\frac{l_2K}{\alpha q_2}$$

Substituting t_1 and t_2 optimum in (18) we have the optimum $MSE(t_p)$ as

$$MSE(t_p)_{opt} = f_1\bar{Y}^2\begin{bmatrix} C_y^2 + l_1^2C_x^2 + l_2^2C_z^2 - 2l_1\rho_{xy}C_xC_y - \\ 2l_2\rho_{zy}C_zC_y + 2l_1l_2\rho_{xz}C_xC_z \end{bmatrix} \tag{21}$$

2.2 Sub members of the Class

Estimators in this class are obtained by setting $t_1 = 0, t_2 = 0$, $\alpha = 0$ in (14) and (18), we have the usual sample per mean estimator as in (1) and the variance as in (2) respectively Also setting $t_1 = b_{xy} = \frac{s_{xy}}{s_x^2}$ (sample regression coefficient of y on x), $t_2 = b_{xz} = \frac{s_{xz}}{s_z^2}$, (regression coefficient of x on z), where s_x^2 and s_z^2 are the sample variances of x and z respectively, s_{xy} and s_{xz} are the sample covariances between x and y and between x and z, respectively and $\alpha = 1$(ratio estimator), A=A_j , G=G_j in (14) then we have jth members of the class of ratio estimator for estimating the population mean using two auxiliary variables where j=1,..,20 and are given as follows:

$$t_{pj} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x}A_j + G_j - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}\left(\bar{X}A_j + G_j\right),\ 0 < \gamma \le 1\ ,\ j=1,2,\dots,20 \qquad (22)$$

and the jth estimator t_p can be seen in Table 1. Using large sample approximation, b_{xy} tends to β_{xy} (population regression coefficient of y on x) and b_{xz} tends to β_{xz} (population regression coefficient of x on z), the mean square error of t_{pj} from (18) is as follows:

$$\mathrm{MSE}(t_{pj}) = f_1\begin{bmatrix} \bar{Y}^2C_y^2 + \beta_{xy}^2 q_1^2 C_x^2 + \lambda_j^2\bar{Y}^2C_x^2 + \beta_{xz}^2 q_3^2\bar{Y}^2C_z^2 - 2\beta_{xy}q_1\bar{Y}\rho_{xy}C_xC_y \\ -2\lambda_j\bar{Y}^2\rho_{xy}C_xC_y + 2\beta_{xz}q_3\bar{Y}^2\rho_{zy}C_zC_y + 2\beta_{xy}\lambda_j q_1\bar{Y}C_x^2 - \\ 2b_{xy}\beta_{xz}q_4\bar{Y}\rho_{xz}C_xC_z - 2\beta_{xz}\lambda_j q_3\bar{Y}^2\rho_{xz}C_xC_z \end{bmatrix} \qquad (23)$$

further simplification of (23), we have

$$\mathrm{MSE}(t_{pj}) = f_1\left(\bar{Y}^2C_y^2 + B_1^2C_x^2 + B_2^2C_z^2 - 2B_3\rho_{xy}C_xC_y - 2B_4\rho_{xz}C_xC_z + 2B_5\rho_{zy}C_zC_y\right)$$

where $\mathrm{B}_1 = \lambda_j + q_1\bar{Y}b_{xy}$, $\mathrm{B}_2 = q_3\bar{Y}b_{xz}$, $\mathrm{B}_3 = \lambda_j\bar{Y}^2 + q_1\bar{Y}b_{xy}$,

$\mathrm{B}_4 = \left(q_4\bar{Y}b_{xy} + \lambda_j q_3\bar{Y}^2\right)b_{xz}$, $\mathrm{B}_5 = q_3\bar{Y}^2b_{xz}$

where $\lambda_j = \dfrac{A_j\bar{X}}{A_j\bar{X} + G_j}$, A_j and G_j are the chosen constant or parameter of the auxiliary variable X for jth estimator.

Table 1: Members of the proposed class of ratio estimators using two auxiliary variables

j	New Estimators $0 < \gamma \le 1$	A_j	G_j
1	$t_{p1} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x} - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}\bar{X}$	1	0
2	$t_{p2} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x} + C_x - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}(\bar{X} + C_x)$	1	C_x
3	$t_{p3} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x} + \beta_{2(x)} - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}(\bar{X} + \beta_{2(x)})$	1	$\beta_{2(x)}$
4	$t_{p4} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x}\beta_{2(x)} + C_x - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}(\beta_{2(x)}\bar{X} + C_x)$	$\beta_{2(x)}$	C_x
5	$t_{p5} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x}C_x + \beta_{2(x)} - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}(C_x\bar{X} + \beta_{2(x)})$	C_x	$\beta_{2(x)}$
6	$t_{p6} = \frac{\bar{y} - b_{xy}(\bar{x}^{\gamma} - \bar{X}^{\gamma})}{\bar{x} + \rho_{xy} - b_{xz}(\bar{z}^{\gamma} - \bar{Z}^{\gamma})}(\bar{X} + \rho_{xy})$	1	ρ_{xy}

7	$t_{p7}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}C_x+\rho_{xy}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}C_X+\rho_{xy})$	C_x	ρ_{xy}
8	$t_{p8}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}\rho_{xy}+C_X-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}\rho_{xy}+C_x)$	ρ_{xy}	C_x
9	$t_{p9}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}\beta_{2(x)}+\rho_{xy}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}\beta_{2(x)}+\rho_{xy})$	$\beta_{2(x)}$	ρ_{xy}
10	$t_{p10}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}\rho_{xy}+\beta_{2(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}\rho_{xy}+\beta_{2(x)})$	ρ_{xy}	$\beta_{2(x)}$
11	$t_{p11}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{1(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{1(x)})$	1	$D_{1(x)}$
12	$t_{p12}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{2(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{2(x)})$	1	$D_{2(x)}$
13	$t_{p13}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{3(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{3(x)})$	1	$D_{3(x)}$
14	$t_{p14}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{4(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{4(x)})$	1	$D_{4(x)}$
15	$t_{p15}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{5(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{5(x)})$	1	$D_{5(x)}$
16	$t_{p16}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{6(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{6(x)})$	1	$D_{6(x)}$
17	$t_{p17}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{7(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{7(x)})$	1	$D_{7(x)}$
18	$t_{p18}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{8(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{8(x)})$	1	$D_{8(x)}$
19	$t_{p19}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{9(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{9(x)})$	1	$D_{9(x)}$
20	$t_{p20}=\dfrac{\bar{y}-b_{xy}(\bar{x}^{\gamma}-\bar{X}^{\gamma})}{\bar{x}+D_{10(x)}-b_{xz}(\bar{z}^{\gamma}-\bar{Z}^{\gamma})}(\bar{X}+D_{10(x)})$	1	$D_{10(x)}$

3. Efficiency Comparison

In this section we have derived the theoretical conditions for which the sub member of the proposed class of ratio estimators t_p in (20) perform better than the existing estimators $t_0, t_{MR},\ t_S,\ t_M,\ t_{2lr}$ and the condition which the generalized class $t_{p(opt)}$ performed better than all the estimators considered in this study

(i) The proposed estimators t_{pj} , j=1,...,20 will be better than the sample mean t_o

iff $MSE(t_{pj}) < MSE(t_0)$

$$\left[B_1^2C_x^2+B_2^2C_z^2-2B_3\rho_{xy}C_xC_y-2B_4\rho_{xz}C_xC_z+2B_5\rho_{zy}C_zC_y\right]\le 0$$

(ii) The proposed estimators t_{pj} , j=1,...,20 will be better than t_{MR}

iff $MSE(t_{pj}) < MSE(t_{MR})$

$$[\bar{Y}^2C_y^2+(B_1^2-\bar{Y}^2k_1^{*2})C_x^2+(B_2^2-\bar{Y}^2k_2^{*2})C_z^2-2(B_3-\bar{Y}^2k_1^*)\rho_{xy}C_xC_y$$
$$-2(B_4+\bar{Y}^2k_1^*k_2^*)\rho_{xz}C_xC_z+2(B_5+\bar{Y}^2k_2^*)\rho_{zy}C_zC_y]\le 0$$

(iii)The proposed estimators t_{pj} , j=1,...,20 will be better than t_s

iff $MSE(t_{pj}) < MSE(t_S)$

$$\begin{bmatrix} (B_1^2 - \bar{Y}^2)C_x^2 + (B_2^2 - \bar{Y}^2)C_z^2 - 2(B_3 - \bar{Y}^2)\rho_{xy}C_xC_y \\ -2(B_4 + \bar{Y}^2)\rho_{xz}C_xC_z + 2B_5\rho_{zy}C_zC_y \end{bmatrix} \le 0$$

(iv) The proposed estimators t_{pj} , j=1,...,20 will be better than t_M

iff $MSE(t_{pj}) < MSE(t_M)$

$$\begin{bmatrix} \bar{Y}^2\rho_{xy}^2C_y^2 + B_1^2C_x^2 + (B_2^2 - \bar{Y}^2)C_z^2 - 2B_3\rho_{xy}C_xC_y - 2B_4\rho_{xz}C_xC_z \\ +2(B_5 + \bar{Y}^2)\rho_{zy}C_zC_y - 2\bar{Y}^2\rho_{xz}\rho_{xy}C_zC_y \end{bmatrix} \le 0$$

(v) The proposed estimators t_{pj} , j=1,...,20 will be better than t_{SW}

iff $MSE(t_{pj}) < MSE(t_{Sw})$

$$\begin{bmatrix} B_1^2C_x^2 + (B_2^2 - \bar{Y}^2)C_z^2 + 2\bar{Y}^2C_{zy} - 2B_3\rho_{xy}C_xC_y - 2B_4\rho_{xz}C_xC_z \\ +2B_5\rho_{zy}C_zC_y + \bar{Y}^2(\rho_{xy}C_y - \rho_{xz}C_z)^2 \end{bmatrix} \le 0$$

(vi) The proposed estimators t_{pj} , j=1,...,20 will be better than t_{2lr}

iff $MSE(t_{pj}) < MSE(t_{2lr})$

$$\begin{bmatrix} B_1^2C_x^2 + B_2^2C_z^2 - 2B_3\rho_{xy}C_xC_y - 2B_4\rho_{xz}C_xC_z - 2B_5\rho_{zy}C_zC_y \\ -\bar{Y}^2C_y^2(1 - \rho_{xy}^2 - \rho_{zy}^2 + 2\rho_{xy}\rho_{zy}\rho_{xz}) \end{bmatrix} \le 0$$

(vii)The optimum MSE $(t_p)_{\text{opt}}$ will be better than the sub members t_{pj} , j=1,...,20

iff $MSE(t_p)_{\text{opt}} < \text{MSE}(t_{pj})$

$$\begin{pmatrix} (\bar{Y}^2l_1^2 - B_1^2)C_x^2 + (\bar{Y}^2l_2^2 - B_2^2)C_z^2 - 2(\bar{Y}^2l_1 - B_3)\rho_{xy}C_xC_y - \\ 2(\bar{Y}^2l_2 + B_5)\rho_{zy}C_zC_y + 2(\bar{Y}^2l_1l_2 + B_4)\rho_{xz}C_xC_z \end{pmatrix} \le 0$$

The Percent Relative Efficiencies (PREs) of t_{MR}, t_S, t_M, t_{2lr} and the proposed t_{pj},$t_{p(opt)}$ with respect to the usual sample mean t_0, are compared using the formula

PRE = $\frac{Var(t_0)}{MSE(.)} * 100$ where (.) =t_{MR}or t_S or $t_M$or t_{2lr} or t_{SW} or t_{pj} or $t_{p(opt)}$

The higher the percent relative efficiency, the more efficient is the estimator.

4. Empirical Study

In this section, we computed mean square errors and percent relative efficiencies of the existing and new estimators and the results are given in Tables 2 and 3. We considered the real life data sets by Chattefuee and Hadi (2006, p.187) and details of the data is as shown below

Y- Per capita expenditure on education in 1975

X- Per capita income in 1973

$Z -$ Number of residents per thousand living in urban areas in 1970

$N = 50$ $n = 15$ $\bar{Y} = 284.0612$ $\bar{X} = 4675.12$ $\bar{Z} = 657.8$

$\rho_{xy} = 0.60679$ $\rho_{zy} = 0.31675$ $\rho_{xz} = 0.61937$ $S_y = 733.1407$ $\beta_{xy} = 0.058228$

$\beta_{xz} = 2.764116 \quad \beta_{zy} = 0.13756 \quad C_y = 0.21776 \quad C_x = 0.13786 \quad C_z = 0.2204$
$D_{1(x)} = 3817 \quad D_{2(x)} = 3967 \quad D_{3(x)} = 4243 \quad D_{4(x)} = 4504 \quad D_{5(x)} = 4697$
$D_{6(x)} = 4827 \quad D_{7(x)} = 4989 \quad D_{8(x)} = 5309 \quad D_{9(x)} = 5560 \quad D_{10} = 5889$
$$\beta_{2(x)} = -0.94843 \quad \beta_{1(x)} = 0.05675$$

The results of the computation are presented below

Table 2: The MSE and PRE with respect to t_0 of the existing estimators

Estimators	MSE	PRE
t_0 (Sample mean)	178.573	100
t_{2lr}	137.427	129.94
t_{MR}	357.103	50.00
t_S	323.077	54.59
t_M	317.130	56.30
t_{SW}	246.914	72.32

Table 3: The MSE and PRE with respect to t_0 of the proposed estimators

j	MSE and Bias of the proposed estimators t_{pj} for j=1,…,20 at $\alpha = 1$									
	$\gamma = 0.1$		$\gamma = 0.3$		$\gamma = 0.5$		$\gamma = 0.8$		$\gamma = 1$	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	112.93	158.1	112.90	158.16	112.77	158.34	112.503	158.73	148.5	120.25
2	112.93	158.1	112.90	158.16	112.77	158.34	112.504	158.73	148.5	120.25
3	112.93	158.1	112.90	158.16	112.77	158.34	112.5	158.73	148.48	120.27
4	112.93	158.1	112.90	158.16	112.77	158.34	112.503	158.73	148.5	120.25
5	112.92	158.1	112.89	158.17	112.77	158.35	112.483	158.76	148.37	120.36
6	112.93	158.1	112.90	158.16	112.77	158.34	112.505	158.72	148.51	120.24
7	112.94	158.1	112.91	158.15	112.78	158.33	112.517	158.71	148.58	120.18
8	112.93	158.1	112.90	158.16	112.77	158.34	112.504	158.73	148.51	120.25
9	112.93	158.1	112.90	158.16	112.77	158.34	112.501	158.73	148.49	120.26
10	112.93	158.1	112.90	158.16	112.77	158.34	112.499	158.73	148.47	120.27
11	165.39	107.9	165.31	108.02	165.18	108.11	171.652	104.03	267.86	66.667
12	169.39	105.4	169.32	105.46	169.18	105.55	175.923	101.51	274.5	65.055
13	177.15	100.8	177.08	100.84	176.94	100.92	184.167	96.963	287.1	62.2
14	184.95	96.55	184.87	96.592	184.73	96.663	192.421	92.804	299.47	59.631
15	191.00	93.49	190.92	93.531	190.78	93.598	198.811	89.821	308.9	57.809
16	195.22	91.47	195.13	91.512	194.99	91.577	203.252	87.858	315.39	56.619
17	200.62	89.00	200.54	89.046	200.40	89.107	208.942	85.466	323.64	55.177
18	211.80	84.30	211.71	84.34	211.58	84.4	220.686	80.918	340.43	52.455
19	221.04	80.78	220.95	80.819	220.81	80.869	230.366	77.517	354.07	50.435
20	233.78	76.38	233.68	76.416	233.54	76.461	243.679	73.282	372.57	47.93
$t_{p(opt)}$	MSE =111.8, PRE =159.73 t_{1opt} =0.00322, t_{2opt}=0.67471, $\lambda = 1, A = 1, G = 0$									

5. Results and Discussion

Table 2: Results

We observed that all the existing estimators t_S, t_M, t_{SW} have no significant improvement on the classical regression estimator (t_{2lr}) using two auxiliary variables, because they have higher MSE and lower percent relative efficiency but they are more efficient than the traditional multivariate ratio estimator t_{MR}.

Table 3: Results

All the proposed estimators t_{pj}, j=1,…,20 at γ =0.1 to 1 are more efficient than Olkin (1958) traditional multivariate ratio estimator (t_{MR}), Singh(1965) estimator t_S, Mohanty (1967) estimator t_M and Swain (2012) estimator t_{SW}, because they have smaller MSE and higher percent relative efficiency except for estimators t_{pj}, j=11,…,20 at $\gamma = 1$ and t_{pj}, j=20 at $\gamma = 0.8$. We observed that the MSE of the proposed estimator t_{pj}, j=1,…,10, reduces as γ increases from 0.1 to 0.8 and the estimators are more efficient than the classical regression estimator(t_{2lr}) because they have smaller MSE and higher PRE. The proposed estimators t_{pj}, j=1,…,10 at $\gamma = 0.1$ to 0.8 are the estimators that utilized known population parameters such as coefficient of variation, coefficient of kurtosis, coefficient of skewness and correlation coefficient and they are approximately equal to the MSE of t_{pj} at t_1 and t_2 optimum values. The MSE of estimators t_{pj}, j=11,…,20 also reduces as γ increases from 0.1 to 0.5 but they are not efficient as the optimum MSE of t_p.

In general, the proposed class of estimator t_p at $\mathrm{t}_1$and t_2 optimum values, is the most efficient estimator because it has the least MSE and highest PRE. Alternatively, a good guess of γ for the proposed estimator t_{pj}, j=1,…,10 at γ between 0.1 to 0.8 inclusive is as efficient as estimator t_p at $\mathrm{t}_1$and t_2 optimum values and they all perform better than the classical regression estimator with two auxiliary variables.

6. Conclusion

In this work, we have proposed a class of ratio type estimator with a linear combination using two auxiliary variables with known population parameters of the auxiliary variable(X). The conditions under which the proposed estimators perform theoretically better than other existing estimators given in this work are mentioned in section 3.0. Here, we conclude that the proposed class of ratio type estimator t_p at $\mathrm{t}_1$and t_2 optimum values and some sub-members of the class of estimators t_{pj}, for stable transformation(γ) perform better than all the existing estimators considered in this study. Therefore, we strongly recommend the proposed estimators over the existing ratio type estimators for practical applications.

References

1. Abid, M., Abbas, N., Sherwani, R. A. K., & Nazir, H. Z. (2016). Improved ratio estimators for the population mean using non-conventional measures of dispersion. *Pakistan of Journal of Statistics and Operation Research*, *12*(2), 353-367
2. Abu-Dayyeh, W. A., Ahmed, M. S., Ahmed, R. A., & Muttlak, H. A. (2003). Some estimators of finite population mean using auxiliary information. *Applied Mathematics and Computation*, *139*, 287-298.
3. Cochran, W.G. (1977). *Sampling Techniques* (3rd ed.). Wiley Eastern Limited, New York, NY

4. Chattefuee, S., & Hadi, A. S. (2006). *Regression analysis by example* (4th ed.). John Wiley & Sons, New York, NY.
5. Jelani, M. I., Maqbool, S. And Mir, S. A. (2013). Modified ratio estimators of population mean using linear combination of coefficient of skewness and quartile deviation. *International Journal of Modern Mathematical Sciences 6*(3), 174-183.
6. Kadilar, C. & Cingi, H. (2004), Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, 151(3), 893–902.
7. Kadilar, C. & Cingi, H(2006). An improvement in estimating the population mean by using the correlation coefficient. *Hacettepe Journal of Mathematics and Statistics 35*(1), 103–109.
8. Lu, J. & Yan, Z. (2014). A class of ratio estimators of a finite population mean using two auxiliary variables. *Plos one, 9*(2), e89538.
9. Kadilar, C., & Cingi, H. (2005). A new estimator using two auxiliary variables. *Applied Mathematics and Computation, 162*, 901-908.
10. Mohanty, S. (1967). Combination of regression and ratio estimate. *Journal of Indian Statistical Association, I5*, 16-19.
11. Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika*, *45*, 154-165.
12. Perri, P. F. (2007). Improved ratio-cum-product type estimators. *Statistics in Transition: New Series*, 8(1): 51-69.
13. Raja, T. A., Subair, M., Maqbool, S., & Hakak, A. (2017). Enhancing the mean ratio estimator for estimating population mean using conventional parameters. *International Journal of Mathematics and Statistics Invention*, *5*(1), 58-61.
14. Singh, M. P. (1965). On the estimation of ratio and product of population parameters. *Sankhya*, *C*(27), 321-328.
15. Singh, M. P. (1967). Multivariate product method of estimation for finite population. *Journal of the Indian Society of Agriculture Statistics, 19,* 1-10.
16. Singh, H. P. & Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population means. *Statistics in Transition*, *6*(4), 555–560.
17. Singh, H. P. & Kakran, M. S. (1993). A modified ratio estimator using known coefficient of kurtosis of an auxiliary character. Revised version submitted to *Journal of Indian Society of Agricultural Statistics*, New Delhi, India.
18. Sisodia, B. V. S. & Dwivedi, V. K. (1981) 'A modified ratio estimator using coefficient of variation of auxiliary variable. *Journal of the Indian Society of Agricultural Statistics*, *33*(1), 13–18.
19. Subramani, J. & Kumarapandiyan, G. (2012a). Estimation of population mean using co-efficient of variation and median of an auxiliary variable. *International Journal of Probability and Statistics, 1*(4), 111–118.
20. Subramani, J. & Kumarapandiyan, G. (2012b). Estimation of population mean using known median and co-efficient of skewness. *American Journal of Mathematics and Statistics, 2*(5), 101–107.
21. Subramani, J. & Kumarapandiyan, G. (2012c). Modified ratio estimators using known median and co-efficient of kurtosis. *American Journal of Mathematics and Statistics, 2*(4), 95–100.
22. Subramani, J. & Kumarapandiyan, G. (2013). Estimation of finite population mean using deciles of an auxiliary variable. *Statistics In Transition:New Series, 14*(1), 75–88.

23. Swain A.K.P.C. (2012). On classes of modified ratio type and regression-cum-ratio type estimators in sample surveys using two auxiliary variables. *Statistics In Transition-New series*, *13*(3), 473—494.
24. Tripathi, T. P., Das, A. K. & Khare, B. B. (1994). Use of auxiliary information in sample surveys: A review. *Aligarh Journal of Statistics,* 14, 79-134.
25. Yan, Z. & Tian, B., (2010). Ratio method to the mean estimation using coefficient of skewness of auxiliary variable, *ICICA 2010, Part II, CCIS* 106, 103-110.