J. Japan Statist. Soc.
Vol. 36 No. 1 2006 107–119

DUAL OF RATIO ESTIMATORS OF FINITE POPULATION MEAN OBTAINED ON USING LINEAR TRANSFORMATION TO AUXILIARY VARIABLE

H. S. Jhajj*, M. K. Sharma* and Lovleen Kumar Grover**

The efficiencies of the ratio- type estimators have been increased by using linear transformation on auxiliary variable in the literature. But such type of estimators requires the additional knowledge of unknown population parameters, which restricts their applicability. Keeping in view such restrictions, we have proposed two unbiased estimators of population mean of study variable on applying linear transformation to auxiliary variable by using its extreme values in the population that are generally available in practice. The comparison of the proposed estimators with the existing ones have been done with respect to their variances. It has also been shown that the proposed estimators have greater applicability and are more efficient than the mean per unit estimator even when the existing estimators are less efficient. We have also shown that under some known conditions the choice of most efficient estimators among the considered ones can be made for a given population. The theoretical results obtained are shown diagrammatically and have been verified numerically by taking some empirical populations.

Key words and phrases: Bias, efficiency, most efficient estimator, preference region, simple random sampling without replacement, unbiased estimator, variance.

1. Introduction

For estimating the population mean $\bar{Y}$ of the variable under study y, the ratio estimator $\bar{y}_R$ has been widely used when there is a high positive correlation between study variable y and auxiliary variable x. In literature, it has been shown by various authors viz Mohanty and Das (1971), Reddy (1974), Reddy and Rao (1977), Srivenkataramana (1978), Chaudhuri and Adhikari (1979) that the bias and the mean square error of the ratio estimator $\bar{y}_R$ can be reduced with the application of transformation on the auxiliary variable x. By using such transformation on auxiliary variable, the construction of the estimator of population mean $\bar{Y}$ requires the knowledge of unknown parameters, which restrict the applicability of these estimators. To overcome such type of restrictions, Mohanty and Sahoo (1995) have defined two ratio estimators by making the transformation on auxiliary variable x, using its minimum value X_m and the maximum value X_M in the population, when the values X_m and X_M in addition to its population mean $\bar{X}$ are available in advance. The information of extreme values X_m and X_M is generally available in practice, otherwise it can be obtained approximately from

Received September 16, 2004. Revised November 8, 2004. Accepted April 8, 2005.
*Department of Statistics, Punjabi University, Patiala 147 002, India.
**Department of Mathematics, G.N.D. University, Amritsar 143 005, India. Email: lovleen_2@yahoo.co.in

either the past experience or pilot sample survey, inexpensively. On assuming auxiliary variable x as positive variable, Mohanty and Sahoo (1995) made the following transformations on x as

$$\left.\begin{aligned} v &= \frac{x + X_m}{X_M + X_m} \\ \omega &= \frac{x + X_M}{X_M + X_m} \end{aligned}\right\}. \tag{1.1}$$

From (1.1), we see that $v \in (0, 1]$ and $\omega \in [1, 2)$. It should be noted that the correlation coefficients for the bivariate (y, v) and (y, ω) remain same as that of (y, x).

Assuming that the variables y and x are positively correlated and prior information on population mean $\bar{X}$ of the auxiliary variable x is available then ratio estimator $\bar{y}_R$ is defined as

$$\bar{y}_R = \frac{\bar{y}}{\bar{x}}\bar{X} \tag{1.2}$$

where $\bar{y}$ and $\bar{x}$ are sample means of variables y and x respectively.

Under the transformations as given in (1.1), Mohanty and Sahoo (1995) defined the following ratio estimators of $\bar{Y}$ by using the known values of population means $\bar{V}$ and $\bar{\Omega}$ of variables v and ω respectively,

$$t_{MS1} = \frac{\bar{y}}{\bar{v}}\bar{V} \tag{1.3}$$

and

$$t_{MS2} = \frac{\bar{y}}{\bar{\omega}}\bar{\Omega} \tag{1.4}$$

where $\bar{v}$ and $\bar{\omega}$ denote the sample means of variables v and ω respectively.

Srivenkataramana (1980) defined a dual to the conventional ratio estimator $\bar{y}_R$ as

$$\bar{y}_{RD} = \frac{\bar{y}(N\bar{X} - n\bar{x})}{(N - n)\bar{X}} \tag{1.5}$$

where n and N denote the sample size and the population size respectively. He has shown that the exact expression for the bias of the estimator $\bar{y}_{RD}$ can be obtained even for a positive correlation between variables y and x whereas the exact expression for the bias of the estimator $\bar{y}_R$ is not available.

In the present paper, we have proposed two unbiased estimators of population mean $\bar{Y}$, which are respectively dual to the ratio estimators t_{MS1} and t_{MS2}. The expressions for their variances have been obtained. The comparisons of the proposed estimators with the existing estimators have been made with respect to their variances. The preference regions of various estimators have also been obtained. The results have also been illustrated diagrammatically as well as numerically.

2. Proposed estimators and their variances

Suppose a simple random sample of size n is drawn from a finite population of size N without replacement and observation on variable y and x are taken. When the values of X_m, X_M and $\bar{X}$ are known in advance then under the transformation $v = \frac{x+X_m}{X_M+X_m}$ and $\omega = \frac{x+X_M}{X_M+X_m}$ (same as given in (1.1)), the values of $\bar{V}$ and $\bar{\Omega}$ will also be known. Since the variables y and x are positively correlated therefore by using the known values of $\bar{V}$ and $\bar{\Omega}$, we consider the following estimators of $\bar{Y}$ which are dual to the ratio estimators t_{MS1} and t_{MS2} respectively

$$\hat{y}_{D1} = \frac{\bar{y}(N\bar{V} - n\bar{v})}{(N-n)\bar{V}} \tag{2.1}$$

and

$$\hat{y}_{D2} = \frac{\bar{y}(N\bar{\Omega} - n\bar{\omega})}{(N-n)\bar{\Omega}}. \tag{2.2}$$

The biases of estimator $\hat{y}_{D1}$ and $\hat{y}_{D2}$ are obtained as

$$B(\hat{y}_{D1}) = -\frac{1}{N}\bar{Y}\rho_{yx}C_y\frac{\bar{X}C_x}{X_m+\bar{X}} \tag{2.3}$$

$$B(\hat{y}_{D2}) = -\frac{1}{N}\bar{Y}\rho_{yx}C_y\frac{\bar{X}C_x}{X_M+\bar{X}} \tag{2.4}$$

where C_y and C_x are coefficients of variations of variables y and x respectively; ρ_{yx} is the correlation coefficient between y and x. The biases of estimators $\hat{y}_{D1}$ and $\hat{y}_{D2}$ given in (2.3) and (2.4) are constant so can be estimated on the basis of same sample. Hence the corresponding proposed unbiased estimators in place of $\hat{y}_{D1}$ and $\hat{y}_{D2}$ are

$$\hat{y}^*_{D1} = \frac{\bar{y}(N\bar{V} - n\bar{v})}{(N-n)\bar{V}} + \frac{s_{xy}}{N(X_m+\bar{X})} \tag{2.5}$$

$$\hat{y}^*_{D2} = \frac{\bar{y}(N\bar{\Omega} - n\bar{\omega})}{(N-n)\bar{\Omega}} + \frac{s_{xy}}{N(X_M+\bar{X})}. \tag{2.6}$$

Srivenkataramana (1980) has also defined the unbiased estimator in place of $\bar{y}_{RD}$ as

$$\bar{y}^*_{RD} = \frac{\bar{y}(N\bar{X} - n\bar{x})}{(N-n)\bar{X}} + \frac{s_{xy}}{N\bar{X}}. \tag{2.7}$$

The results obtained are given in the following theorem.

THEOREM 2.1. *For the simple random sampling without replacement (SRSWOR), the variances of the proposed estimators $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$, up to the terms of order n^{-1}, are*

$$V(\hat{y}^*_{D1}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \left(\frac{n}{N-n}\right)^2\frac{\bar{X}^2C_x^2}{(X_m+\bar{X})^2} - 2\left(\frac{n}{N-n}\right)\rho_{yx}C_y\frac{\bar{X}C_x}{X_m+\bar{X}}\right] \tag{2.8}$$

and

$$V(\hat{\bar{y}}^*_{D2}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \left(\frac{n}{N-n}\right)^2 \frac{\bar{X}^2C_x^2}{(X_M+\bar{X})^2} - 2\left(\frac{n}{N-n}\right)\rho_{yx}C_y\frac{\bar{X}C_x}{X_M+\bar{X}}\right]. \tag{2.9}$$

Remark. It is noted that under the given transformation i.e. x to v or x to ω, as defined in (1.1), there is no change in the form of usual linear regression estimator $\bar{y}_{lr} = \bar{y} + b_{yx}(\bar{X} - \bar{x})$.

3. Comparison of proposed estimators with existing ones

For comparing the proposed estimators with existing ones, we first write the expressions for the variances of the existing estimators, namely $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1} and t_{MS2} (under the same sampling scheme) up to the terms of order n^{-1} (the biases of these estimators if exist is of order n^{-1}, so their contributions to the mean square errors will be of order n^{-2}) as

$$V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2C_y^2 \tag{3.1}$$

$$V(\bar{y}_R) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2(C_y^2 + C_x^2 - 2\rho_{yx}C_yC_x) \tag{3.2}$$

$$V(\bar{y}^*_{RD}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \left(\frac{n}{N-n}\right)^2 C_x^2 - 2\left(\frac{n}{N-n}\right)\rho_{yx}C_yC_x\right] \tag{3.3}$$

$$V(t_{MS1}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \frac{\bar{X}^2C_x^2}{(X_m+\bar{X})^2} - 2\rho_{yx}C_y\frac{\bar{X}C_x}{X_m+\bar{X}}\right] \tag{3.4}$$

$$V(t_{MS2}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + \frac{\bar{X}^2C_x^2}{(X_M+\bar{X})^2} - 2\rho_{yx}C_y\frac{\bar{X}C_x}{X_M+\bar{X}}\right]. \tag{3.5}$$

Reddy (1978) has shown that the values of parameter $K = \rho_{yx}\frac{C_y}{C_x}$ remain stable in any repetitive survey. So we find the conditions on the values of K under which the proposed estimators are superior to the existing ones. For the present situation, we note that $0 < K < \frac{C_y}{C_x}$. In the survey sampling situations, usually $\frac{n}{N-n} < 0.1$ but we assume that $\frac{n}{N-n} < 1$ which hold good in almost all the situations of survey sampling. For the sake of comparison in the compact form, we take

$$\left.\begin{aligned} m_1 &= 1 + \frac{X_m}{\bar{X}} \\ m_2 &= 1 + \frac{X_M}{\bar{X}} \end{aligned}\right\} \tag{3.6}$$

$$\ell = \frac{n}{N-n}. \tag{3.7}$$

Noting that $m_1 \in (1, 2]$ and m_2 is a finite number so that $m_2 \geq 2$ and by assumption we note that $0 < \ell < 1$.

Using the expressions (2.8), (2.9) and (3.1) to (3.5), the results obtained are given in the following theorems.

THEOREM 3.1. *Up to the terms of order n^{-1}, we have*

$$V(\hat{y}^*_{D1}) < V(\bar{y}) \quad \textit{for} \quad K > \frac{\ell}{2m_1}$$
$$V(t_{MS1}) < V(\bar{y}) \quad \textit{for} \quad K > \frac{1}{2m_1}$$
$$V(\hat{y}^*_{D2}) < V(\bar{y}) \quad \textit{for} \quad K > \frac{\ell}{2m_2}$$
$$V(t_{MS2}) < V(\bar{y}) \quad \textit{for} \quad K > \frac{1}{2m_2}.$$

Remark. From the results of above theorem, we see that the proposed estimators $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ which are dual to t_{MS1} and t_{MS2} respectively are superior to the mean per unit estimator $\bar{y}$ even for smaller values of K when t_{MS1} and t_{MS2} are inferior than $\bar{y}$.

THEOREM 3.2. *Up to the terms of order n^{-1}, the variance of the proposed estimator $\hat{y}^*_{D1}$ is less than the variances of all the other estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2} and $\hat{y}^*_{D2}$ for $K \in I_1$, where*

$$I_1 = \begin{cases} \left(\frac{\ell}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right), \frac{\ell}{2}\left(\frac{1}{m_1}+1\right)\right) & \textit{when} \quad 0 < \ell < \frac{1}{m_2} \\ \left(\frac{\ell}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right), \frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right)\right) & \textit{when} \quad \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \\ \left(\frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right), \frac{\ell}{2}\left(\frac{1}{m_1}+1\right)\right) & \\ \quad \textit{when} \quad \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \quad \textit{with} \quad m_1^2 < m_2 & \\ \left(\frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right), \frac{1}{2m_1}(\ell+1)\right) & \\ \quad \textit{when either} \quad \frac{1}{m_1} < \ell < 1 \quad \textit{with} \quad m_1^2 < m_2 & \\ \quad \textit{or} \quad \frac{m_1}{m_2} < \ell < 1 \quad \textit{with} \quad m_1^2 > m_2. & \end{cases} \tag{3.8}$$

THEOREM 3.3. *Up to the terms of order n^{-1}, the variance of the proposed estimator $\hat{y}^*_{D2}$ is less than the variances of all the other estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2} and $\hat{y}^*_{D1}$ for $K \in I_2$, where*

$$I_2 = \begin{cases} \left(\frac{\ell}{2m_2}, \frac{\ell}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\right) & \textit{when} \quad 0 < \ell < \frac{m_1}{m_2} \\ \left(\frac{\ell}{2m_2}, \frac{1}{2m_2}(\ell+1)\right) & \textit{when} \quad \frac{m_1}{m_2} < \ell < 1 \end{cases}. \tag{3.9}$$

THEOREM 3.4. *Up to the terms of order* n^{-1}, *the estimator* $\bar{y}$ *will become most efficient among the estimators* $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ *and* $\hat{y}^*_{D2}$ *for*

$$K \in \left(0, \frac{\ell}{2m_2}\right) = I_3. \tag{3.10}$$

THEOREM 3.5. *Up to the terms of order* n^{-1}, *the estimator* $\bar{y}_R$ *will become most efficient among the estimators* $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ *and* $\hat{y}^*_{D2}$ *for* $K \in I_4$, *where*

$$I_4 = \begin{cases} \left(\frac{1}{2}\left(1+\frac{1}{m_1}\right), \frac{C_y}{C_x}\right) & \text{when} \quad \ell < \frac{1}{m_1} \\ \left(\frac{1}{2}\left(1+\ell\right), \frac{C_y}{C_x}\right) & \text{when} \quad \ell > \frac{1}{m_1} \end{cases}. \tag{3.11}$$

THEOREM 3.6. *Up to the terms of order* n^{-1}, *the estimator* $\bar{y}^*_{RD}$ *will become most efficient among the estimators* $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ *and* $\hat{y}^*_{D2}$ *for* $K \in I_5$, *where*

$$I_5 = \begin{cases} \left(\frac{\ell}{2}\left(1+\frac{1}{m_1}\right), \frac{1}{2}\left(\ell+\frac{1}{m_2}\right)\right) & \text{when} \quad 0 < \ell < \frac{1}{m_2} \\ \left(\frac{1}{2}\left(\ell+\frac{1}{m_2}\right), \frac{1}{2}\left(\ell+\frac{1}{m_1}\right)\right) & \\ \qquad \text{when either} \quad \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \ \text{with} \ m_1^2 < m_2 & \\ \qquad \text{or} \quad \frac{1}{m_2} < \ell < \frac{1}{m_1} \ \text{with} \ m_1^2 > m_2 & \\ \left(\frac{1}{2}\left(\ell+\frac{1}{m_1}\right), \frac{1}{2}\left(\ell+1\right)\right) & \\ \qquad \text{when} \quad \frac{1}{m_1} < \ell < \frac{m_1}{m_2} \ \text{with} \ m_1^2 > m_2 & \\ \left(\frac{1}{2}\left(\ell+\frac{1}{m_1}\right), \frac{1}{2}\left(\ell+1\right)\right) & \\ \qquad \text{when} \quad \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \ \text{with} \ m_1^2 < m_2 & \\ \left(\frac{1}{2}\left(\ell+\frac{1}{m_1}\right), \frac{1}{2}\left(\ell+1\right)\right) & \\ \qquad \text{when either} \quad \frac{1}{m_1} < \ell < 1 \ \text{with} \ m_1^2 < m_2 & \\ \qquad \text{or} \quad \frac{m_1}{m_2} < \ell < 1 \ \text{with} \ m_1^2 > m_2 & \end{cases}. \tag{3.12}$$

THEOREM 3.7. *Up to the terms of order n^{-1}, the estimator t_{MS1} will become most efficient among the estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ for $K \in I_6$, where*

$$I_6 = \begin{cases} \left(\frac{1}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right), \frac{1}{2}\left(1+\frac{1}{m_1}\right)\right) \quad \textit{when} \quad 0 < \ell < \frac{1}{m_2} \\ \left(\frac{1}{2}\left(\ell+\frac{1}{m_1}\right), \frac{1}{2}\left(1+\frac{1}{m_1}\right)\right) \\ \quad \textit{when either} \quad \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \;\; \textit{with} \;\; m_1^2 < m_2 \\ \quad \textit{or} \quad \frac{1}{m_2} < \ell < \frac{1}{m_1} \;\; \textit{with} \;\; m_1^2 > m_2 \\ \left(\frac{1}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right), \frac{1}{2}\left(\ell+\frac{1}{m_1}\right)\right) \\ \quad \textit{when} \quad \frac{1}{m_1} < \ell < \frac{m_1}{m_2} \;\; \textit{with} \;\; m_1^2 > m_2 \\ \left(\frac{1}{2}\left(\ell+\frac{1}{m_1}\right), \frac{1}{2}\left(1+\frac{1}{m_1}\right)\right) \\ \quad \textit{when} \quad \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \;\; \textit{with} \;\; m_1^2 < m_2 \\ \left(\frac{1}{2m_1}(\ell+1), \frac{1}{2}\left(\frac{1}{m_1}+\ell\right)\right) \\ \quad \textit{when either} \quad \frac{1}{m_1} < \ell < 1 \;\; \textit{with} \;\; m_1^2 < m_2 \\ \quad \textit{or} \quad \frac{m_1}{m_2} < \ell < 1 \;\; \textit{with} \;\; m_1^2 > m_2 \end{cases} \tag{3.13}$$

THEOREM 3.8. *Up to the terms of order n^{-1}, the estimator t_{MS2} will become most efficient among the estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ for $K \in I_7$, where*

$$I_7 = \begin{cases} \left(\frac{1}{2}\left(\frac{1}{m_2}+\ell\right), \frac{1}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\right) \quad \textit{when} \quad 0 < \ell < \frac{1}{m_2} \\ \left(\frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right), \frac{1}{2}\left(\frac{1}{m_2}+\ell\right)\right) \\ \quad \textit{when either} \quad \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \;\; \textit{with} \;\; m_1^2 < m_2 \\ \quad \textit{or} \quad \frac{1}{m_2} < \ell < \frac{1}{m_1} \;\; \textit{with} \;\; m_1^2 > m_2 \\ \left(\frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right), \frac{1}{2}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\right) \\ \quad \textit{when} \quad \frac{1}{m_1} < \ell < \frac{m_1}{m_2} \;\; \textit{with} \;\; m_1^2 > m_2 \\ \left(\frac{1}{2m_2}(\ell+1), \frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right)\right) \\ \quad \textit{when} \quad \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \;\; \textit{with} \;\; m_1^2 < m_2 \\ \left(\frac{1}{2m_2}(\ell+1), \frac{1}{2}\left(\frac{\ell}{m_1}+\frac{1}{m_2}\right)\right) \\ \quad \textit{when either} \quad \frac{1}{m_1} < \ell < 1 \;\; \textit{with} \;\; m_1^2 < m_2 \\ \quad \textit{or} \quad \frac{m_1}{m_2} < \ell < 1 \;\; \textit{with} \;\; m_1^2 > m_2 \end{cases} \tag{3.14}$$

4. Diagrammatic representation of preference regions of various estimators

The preference region for the estimator is the interval of K on which the estimator is more efficient than the other estimators. To have a bird's eye view of the preference regions for the different estimators obtained earlier in the Section 3, we have made an effort by representing them diagrammatically. For the sake of convenience to mark the limits of the preference regions in the diagrams, we take

$$P_1 = \frac{l}{2m_2}, \quad P_2 = \frac{l}{2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right), \quad P_3 = \frac{l}{2}\left(1 + \frac{1}{m_1}\right),$$
$$P_4 = \frac{1}{2}\left(\frac{1}{m_2} + l\right), \quad P_5 = \frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_2}\right), \quad P_6 = \frac{1}{2}\left(\frac{1}{m_1} + 1\right),$$
$$P_7 = \frac{1}{2}\left[\frac{l}{m_1} + \frac{1}{m_2}\right], \quad P_8 = \frac{1}{2}\left(\frac{1}{m_1} + l\right), \quad P_9 = \frac{1}{2}(1 + l),$$
$$P_{10} = \frac{1}{2m_2}(1 + l), \quad P_{11} = \frac{1}{2m_1}(1 + l).$$

The order of the various preference regions is shown in the following diagrams by considering the whole range of K under various situations.

Diagram 4.1. When $0 < \ell < \frac{1}{m_2}$ then we have

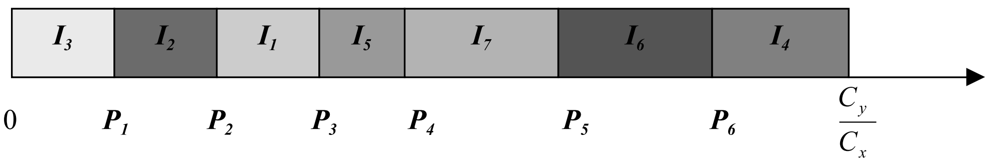


Diagram 4.2. When either $\frac{1}{m_2} < \ell < \frac{m_1}{m_2}$ with $m_1^2 < m_2$ or $\frac{1}{m_2} < \ell < \frac{1}{m_1}$ with $m_1^2 > m_2$ then we have

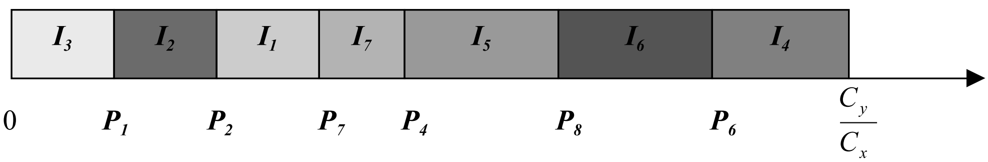


Diagram 4.3. When $\frac{1}{m_1} < \ell < \frac{m_1}{m_2}$ with $m_1^2 > m_2$ then we have

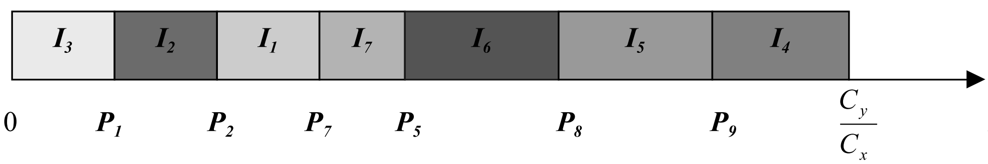


Diagram 4.4. When $\frac{m_1}{m_2} < \ell < \frac{1}{m_1}$ with $m_1^2 < m_2$ then we have

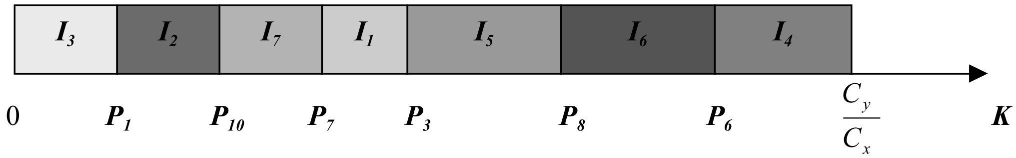
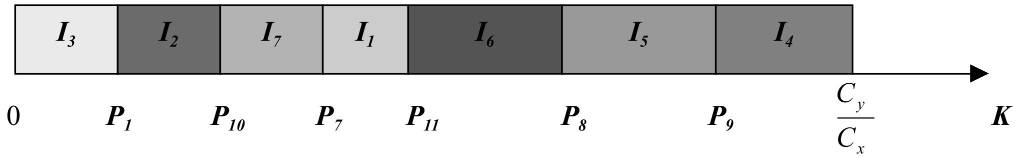


Diagram 4.5. When either $\frac{1}{m_1} < \ell < 1$ with $m_1^2 < m_2$ or $\frac{m_1}{m_2} < \ell < 1$ with $m_1^2 > m_2$ then we have



The above diagrams 4.1 to 4.5 clearly indicate that

(i) There is no advantage of using known auxiliary information for improving the estimator of $\bar{Y}$ when the value of K is very small i.e. close to zero.

(ii) The use of ordinary ratio estimator $\bar{y}_R$ is very beneficial for estimating $\bar{Y}$ when the value of K is very large.

(iii) The proposed estimators $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ which are dual to t_{MS1} and t_{MS2} respectively are superior to the mean per unit estimator $\bar{y}$ even for smaller values of K when t_{MS1} and t_{MS2} are inferior to $\bar{y}$.

(iv) The interval I_2 is always adjacent to I_3 in the diagrams which shows that the estimator $\hat{y}^*_{D2}$ is better than $\bar{y}$ even when $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2} and $\hat{y}^*_{D1}$ are not better than $\bar{y}$.

Hence we conclude that the use of the proposed estimators $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ for estimating $\bar{Y}$ will be more beneficial for moderate values of K. So the choice of sample size can be made accordingly for which the proposed estimators $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ will become most efficient with known values of K, m_1 and m_2. On the other hand, for a given sample size, the choice of most efficient estimator among the considered estimators namely $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ can also be made accordingly with known values of K, m_1 and m_2.

5. An empirical study

To get a rough idea about the efficiencies of the estimators, we have taken seven empirical populations from the literature. The description of the populations and the values of population constants are given in Tables 1 and 2 respectively. The expressions of biases of ratio estimator and ratio estimators defined by Mohanty and Sahoo (1995) are given in Appendix A, so the biases of these estimators are calculated which are given in Table 3. The preference

Table 1. Description of populations.

Population number	Source of population	N	Variable y	Variable x	n	ℓ
1	Singh and Chaudhary (1986) p-166	16	Area under wheat during 1979–80	Total cultivated area during 1978–79	5	0.4545
2	Cochran (1977) p-325	10	Number of persons in a block	Number of rooms in a block	4	0.6667
3	Singh and Chaudhary (1986) p-306	10	Number of inhabitants ('000) in 1981–82	Number of inhabitants ('000) in 1980–81	4	0.6667
4	Singh and Chaudhary (1986) p-155	17	Number of milch animals in survey (1977–78)	Number of milch animals in census (1976)	6	0.5454
5	Singh and Chaudhary (1986) p-177(1–15)	15	Area under wheat in 1973	Area under wheat in 1971	5	0.5000
6	Sampford (1962) p-61(1–9)	9	Acreage under oats in 1957	Acreage of crops and grass in 1947	3	0.5000
7	Panse and Sukhatme (1967) p-124 (1–10)	10	Parental plot mean (mm)	Parental plant value (mm)	4	0.6667

Table 2. Constants of the populations.

Population number	ρ_{yx}	C_x	C_y	m_1	m_2
1	0.9600	0.74	0.6900	1.21	4.06
2	0.6500	0.14	0.1500	1.82	2.22
3	0.8800	0.60	0.6400	1.54	3.43
4	0.4371	0.02	0.0165	1.98	2.04
5	0.1732	0.82	0.8903	1.08	4.59
6	0.0700	0.10	0.2900	1.86	2.12
7	0.0833	0.07	0.0647	1.92	2.13

Table 3. Biases of estimators.

Population number	$\frac{\lvert B(\cdot)\rvert}{(\frac{1}{n}-\frac{1}{N})\bar{Y}} \times 100$		
	Estimator		
	$\bar{y}_R$	t_{MS1}	t_{MS2}
1	5.742400	**3.108596**	8.751211
2	0.595000	**0.158284**	0.217170
3	**2.208000**	6.763232	6.791946
4	0.025576	0.002918	**0.002541**
5	54.5956	45.93969	**0.436784**
6	0.797000	0.179911	**0.126744**
7	0.452273	0.113272	**0.090291**

Table 4. Preference regions of superiority and percentage relative efficiencies with respect to $\bar{y}$ of various estimators.

Population number	K	Preference region and efficiency, **say E**, of estimators							Most efficient estimator
		$\bar{y}$	$\bar{y}_R$	$\bar{y}^*_{RD}$	t_{MS1}	t_{MS2}	$\hat{y}^*_{D1}$	$\hat{y}^*_{D2}$	
1	0.8951	$(0, 0.0560)$ E = 100	$(0.9132, 0.9324)$ E = 1098.3210	$(0.4151, 0.6405)$ E = 331.4861	$(0.6405, 0.9132)$ **E = 1192.9370**	$(0.1791, 0.3110)$ E = 177.7455	$(0.3110, 0.4151)$ E = 257.2123	$(0.0560, 0.1791)$ E = 127.5701	t_{MS1}
2	0.6964	$(0, 0.1501)$ E = 100	$(0.8333, 1.0714)$ E = 152.0270	$(0.6080, 0.8333)$ **E = 172.9291**	$(0.5000, 0.6080)$ E = 167.6957	$(0.4084, 0.5000)$ E = 158.6781	$(0.3333, 0.4084)$ E = 148.7127	$(0.1501, 0.3333)$ E = 140.0182	$\bar{y}^*_{RD}$
3	0.9387	$(0, 0.0972)$ E = 100	$(0.8333, 1.0667)$ **E = 436.8601**	$(0.6580, 0.8333)$ E = 344.0860	$(0.5411, 0.6580)$ E = 334.2606	$(0.2429, 0.3622)$ E = 168.4477	$(0.3622, 0.5411)$ E = 222.0132	$(0.0972, 0.2429)$ E = 140.3503	$\bar{y}_R$
4	0.3606	$(0, 0.1337)$ E = 100	$(0.7727, 0.8250)$ E = 70.9420	$(0.5252, 0.7727)$ E = 116.3948	$(0.4976, 0.5252)$ E = 119.1047	$(0.3828, 0.4976)$ E = 119.9592	$(0.2714, 0.3828)$ **E = 122.0122**	$(0.1337, 0.2714)$ E = 121.6969	$\hat{y}^*_{D1}$
5	0.1880	$(0, 0.0545)$ E = 100	$(0.9630, 1.0857)$ E = 65.3910	$(0.4815, 0.7130)$ E = 95.0070	$(0.7130, 0.9630)$ E = 69.8385	$(0.1634, 0.3404)$ **E = 103.0125**	$(0.3404, 0.4815)$ E = 96.7010	$(0.0544, 0.1634)$ E = 102.5313	t_{MS2}
6	0.2030	$(0, 0.1180)$ E = 100	$(0.7688, 2.9000)$ E = 93.4030	$(0.4858, 0.5188)$ E = 99.4442	$(0.5188, 0.7688)$ E = 99.1655	$(0.3702, 0.4858)$ E = 99.6329	$(0.2523, 0.3702)$ E = 100.4404	$(0.1180, 0.2523)$ **E = 100.4795**	$\hat{y}^*_{D2}$
7	0.0770	$(0, 0.1565)$ **E = 100**	$(0.8333, 0.9243)$ E = 50.2438	$(0.5937, 0.8333)$ E = 71.4247	$(0.4951, 0.5937)$ E = 81.7226	$(0.4083, 0.4951)$ E = 85.2237	$(0.3301, 0.4083)$ E = 92.7180	$(0.1565, 0.3301)$ E = 94.4953	$\bar{y}$

regions of all the estimators in which one estimator is superior to all others and the efficiencies of all the estimators with respect to the estimator $\bar{y}$ are given in Table 4.

Note.
(1) Bold figure in Table 3 indicates the minimum bias for the given population.
(2) Bold figure in Table 4 indicates the maximum efficiency for the given population.
(3) For Table 4, figure in bracket in each box indicates the preference region of the corresponding estimator for the given population.

From Table 4, we see that in all the populations the value of K for the most efficient estimator really lies in the corresponding preference region of the same estimator. Hence it may be concluded that for the known value of K in any population, we can choose the most efficient estimator among the estimators namely $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_{RD}$, t_{MS1}, t_{MS2}, $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$.

Appendix A

Up to the terms of order n^{-1}, we have

$$B(\bar{y}_R) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}(C_x^2 - \rho_{yx} C_y C_x) \tag{A.1}$$

$$B(t_{MS1}) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y} \frac{1}{X_m + \bar{X}} \left(\frac{\bar{X}^2 C_x^2}{X_m + \bar{X}} - \rho_{yx} C_y \bar{X} C_x\right) \tag{A.2}$$

$$B(t_{MS2}) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y} \frac{1}{X_M + \bar{X}} \left(\frac{\bar{X}^2 C_x^2}{X_M + \bar{X}} - \rho_{yx} C_y \bar{X} C_x\right). \tag{A.3}$$

Acknowledgements

Authors are very thankful to the referees for their valuable suggestions given for improving the original version of the manuscript.

References

Chaudhuri, A. and Adhikari, A. K. (1979). Improving on the efficiencies of standard sampling strategies through variate transformations in estimators for finite population means, Tech. Report, ISI Calcutta.

Cochran, W. G. (1977). *Sampling Techniques*, Third Edition, Wiley Eastern Limited, New Delhi, India.

Mohanty, S. and Das, M. N. (1971). Use of transformation in sampling, *J. Ind. Soc. Agr. Stat.*, **23**(2), 83–87.

Mohanty, S. and Sahoo, J. (1995). A note on improving the ratio method of estimators through linear transformation using certain known population parameters, *Sankhya B*, **57**(1), 93–102.

Panse, V. G. and Sukhatme, P. V. (1967). *Statistical Methods for Agricultural Workers*, Indian Council of Agricultural Research, New Delhi, India.

Reddy, V. N. (1974). On a transformed ratio method of estimation, *Sankhya C*, **36**, 59–70.

Reddy, V. N. (1978). A study on the use of prior knowledge on certain population parameters in estimation, *Sankhya C*, **40**, 29–37.

Reddy, V. N. and Rao, T. J. (1977). Modified pps method of estimation, *Sankhya C*, **39**, 185–197.

Sampford, M. R. (1962). *An Introduction to Sampling Theory*, Oliver and Boyd.

Singh, D. and Chaudhary, F. S. (1986). Theory and analysis of sample survey designs, First Edition, Wiley Eastern Limited, New Delhi, India.

Srivenkataramana, T. (1978). Change of origin and scale in ratio and difference methods of estimation in sampling, *Canad. J. Stat.*, **6**, 79–86.

Srivenkataramana, T. (1980). A dual to ratio estimator in sample surveys, *Biometrika*, **67**(1), 199–204.