STATISTICS IN TRANSITION-new series, December 2012
Vol. 13, No. 3, pp. 473—494

ON CLASSES OF MODIFIED RATIO TYPE AND REGRESSION-CUM-RATIO TYPE ESTIMATORS IN SAMPLE SURVEYS USING TWO AUXILIARY VARIABLES

A.K.P.C. Swain[1]

ABSTRACT

In this paper generalized classes of modified ratio type and regression-cum-ratio type estimators of the finite population mean of the study variable are suggested in the presence of two auxiliary variables in simple random sampling without replacement when the population means of the auxiliary variables are known in advance. Some special cases of the generalized estimators are compared with respect to their biases and efficiencies both theoretically and with the help of some natural populations.

Key words: ratio type estimators, regression-cum-ratio type estimators, simple random sampling, auxiliary variables, bias, efficiency.

1. Introduction

In sample surveys a sampler invariably observes certain auxiliary variables to provide more efficient estimators of the finite population mean of the study variable. The literature on the use of auxiliary information in sample surveys is quite vast and old dating back to early part of the 20th century when the foundation stone of modern sampling theory dealing with stratified random sampling was laid out utilizing the auxiliary information by Bowley (1926) and Neyman (1934, 38). However, the use of auxiliary information in the estimation procedure to improve the precision of estimators was initiated by Watson (1937) and Cochran (1940, 42). Hansen and Hurwitz (1943) were the first to suggest the use of auxiliary information in selecting units with varying probabilities. The customary sources of obtaining auxiliary information on one or more variables having strong correlation with the main variable under study are various census

[1] Former Professor of Statistics Utkal University, Bhubaneswar-751004, India. E-mail: swainakpc@yahoo.co.in.

data, previous surveys, pilot surveys, etc., or may be made available while collecting information on the study variable during the survey operations.

Cochran (1940) was the first to introduce a ratio estimator of the population mean of the study variable by using a single auxiliary variable in the form of a ratio of the sample mean of the study variable to the sample mean of the auxiliary variable, multiplied by the population mean of the auxiliary variable. That is, the classical ratio estimator of the population mean $\bar{Y}$ of the study variable y in simple random sampling is defined as

$$\hat{\bar{Y}}_R = \frac{\bar{y}}{\bar{x}}\bar{X},$$

where $\bar{y}$ and $\bar{x}$ are the sample means of the study variable *y* and auxiliary variable *x* respectively and $\bar{X}$ is the population mean of the auxiliary variable *x*. Ratio estimator is seen to be most efficient when the regression line of *y* on *x* passes through the origin. When the regression line does not pass through the origin, Cochran (1940, 42) suggested a linear regression estimator which was generalized by Hansen et al. (1953) in the form of a difference estimator. The linear regression estimator of the population mean $\bar{Y}$ is defined as

$$\hat{\bar{Y}}_{\mathrm{Re}g} = \bar{y} + b_{yx}(\bar{X} - \bar{x}),$$

where b_{yx} is the sample regression coefficient of *y* on *x* and the difference estimator is defined as

$$\hat{\bar{Y}}_D = \bar{y} + \lambda(\bar{X} - \bar{x}),$$

where λ is a real constant to be suitably chosen.

To estimate the population mean $\bar{Y}$ of the study variable *y* in the presence of p-auxiliary variables x_1, x_2, …, x_p with the advance knowledge of the population means $\bar{X}_1, \bar{X}_2, ..., \bar{X}_p$ respectively, Tripathi (1970, 87) discussed two general classes of estimators for any sampling design, defined by

$$\hat{\bar{Y}}_{MT} = \sum_{i=1}^{p} W_i \left[\hat{\bar{Y}} - t_i \left(\hat{\bar{X}}_i - \bar{X}_i \right) \right],$$

and

$$\hat{Y}^*_{MT} = \hat{\bar{Y}} - \sum t_i^* \left(\hat{\bar{X}}_i - \bar{X}_i \right),$$

where $\hat{\bar{Y}}$ and $\hat{\bar{X}}_i$ are unbiased estimators of $\bar{Y}$ and $\bar{X}_i$ (I = 1, 2, …,p) respectively, t_i and t_i^* are statistics (or real constants) such that their expected values exist, W_i's are non-negative and $\sum_{i=1}^{p} W_i = 1$.

These classes include Olkin's (1958) multivariate ratio estimator, Raj's (1965) multivariate difference estimator, Ghosh's (1947), Srivastava's (1965) and

Shukla's (1965) multivariate regression estimator. Tripathi (1987) also considered other classes of estimators for any sampling design defined by

$$e_1 = \hat{\bar{Y}} - \sum_{i=1}^{p} t_i \left(\hat{\bar{X}}_i^{\alpha_i} - \bar{X}_i^{\alpha_i} \right), \quad e_2 = \hat{\bar{Y}} \prod_{i=1}^{p} \left(\frac{\bar{X}_i}{\hat{\bar{X}}_i} \right)^{\alpha_i}$$

$$e_3 = \hat{\bar{Y}} \sum_{i=1}^{p} W_i \left(\frac{\bar{X}_i}{\hat{\bar{X}}_i} \right)^{\alpha_i} \qquad , e_4 = \hat{\bar{Y}} \frac{\sum_{i=1}^{p} W_i \bar{X}_i^{\alpha_i}}{\sum_{i=1}^{p} W_i \hat{\bar{X}}_i^{\alpha_i}}$$

$$e_5 = \hat{\bar{Y}} \frac{\sum_{i=1}^{p} W_i \hat{\bar{X}}_i^{\alpha_i}}{\sum_{i=1}^{p} W_i \bar{X}_i^{\alpha_i}}$$

where α_i 's are suitably chosen constants and W_i's are non-negative weights such that

$$\sum_{i=1}^{p} W_i = 1 .$$

These classes include estimators proposed by Shukla (1966) and John (1969). Singh (1965,67) suggested a ratio-cum-product estimator where some of the auxiliary variables are positively correlated and others are negatively correlated with the study variable.

Srivastava (1971) suggested a general class of estimator in case of simple random sampling without replacement, defined by

$$e_6 = \bar{y} \left(\frac{\bar{x}_1}{\bar{X}_1}, \frac{\bar{x}_2}{\bar{X}_2}, ..., \frac{\bar{x}_p}{\bar{X}_p} \right) = \bar{y}(u_1, u_2, ..., u_p) ,$$

where $\bar{y}, \bar{x}_1, \bar{x}_2, ...\bar{x}_p$ are sample means of $y, x_1, x_2, ...x_p$ respectively and $u_i = \frac{\bar{x}_i}{\bar{X}_i}$, $i = 1, 2, …, p$ and $\boldsymbol{h}(\cdot)$ is a function of $u_1, u_2, …, u_p$ obeying regularity conditions, such as

(a) The point (u_1, u_2, …, u_p) assumes the value in a closed convex subset R_p of p- dimensional real space containing the point (1, 1, …, 1).

(b) The function $h(u_1, u_2, …, u_p)$ is continuous and bounded in R_p

(c) $h(1,1,…..,1) = 1$

(d) The first and second order partial derivatives of $h(u_1, u_2, …, u_p)$ exist and are continuous and bounded in R_p.

Subsequently, Srivastava and Jhajj (1983) suggested a wider class of estimators defined by

$$e_7 = g\left(\bar{y}, \frac{\bar{x}_1}{\bar{X}_1}, \ldots, \frac{\bar{x}_p}{\bar{X}_p}\right)$$

In what follows we shall consider certain specific classes of modified ratio type, difference-cum-ratio type and regression-cum-ratio type estimators and compare them as regards their biases and efficiencies in the presence of two auxiliary variables having known population means.

2. A generalized class of modified ratio type estimators

Let $U = (U_1, U_2, \ldots, U_N)$ be the finite population of size N. To each unit $U_i\ (i = 1, 2, \ldots, N)$ in the population , the values of the study variable *y* and the auxiliary variables *x* and *z* denoted by the triplet (y_i, x_i, z_i) , $(i = 1, 2, \ldots, N)$ are attached.

Now, define the population means of the study variable *y* and the auxiliary variables *x* and *z* respectively as

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i, \qquad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i, \qquad \bar{Z} = \frac{1}{N}\sum_{i=1}^{N} z_i$$

Further, define the finite population variances of *y*, *x* and *z* and their covariances as

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2, \ S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$$

$$S_z^2 = \frac{1}{N-1}\sum_{i=1}^{N}(z_i - \bar{Z})^2, \ S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})(x_i - \bar{X})$$

$$S_{yz} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})(z_i - \bar{Z}), \text{and } S_{xz} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})(z_i - \bar{Z})$$

Also, the coefficients of variations of y, x and z and their coefficients of covariation are defined by

$$C_y = \frac{S_y}{\bar{Y}}, \qquad C_x = \frac{S_x}{\bar{X}}, \qquad C_z = \frac{S_z}{\bar{Z}}$$

$$C_{yx} = \frac{S_{yx}}{\bar{Y}\bar{X}} = \rho_{yx} C_y C_x, \ C_{yz} = \frac{S_{yz}}{\bar{Y}\bar{Z}} = \rho_{yz} C_y C_z, \text{ and } C_{xz} = \frac{S_{xz}}{\bar{X}\bar{Z}} = \rho_{xz} C_x C_z,$$

where $\rho_{yx,}\ \rho_{yz}\ and\ \rho_{xz}$ are simple correlations between *y* and *x*, *y* and *z* and *x* and *z* respectively.

A simple random sample $'s'$of size n is selected from U without replacement and values (y_i, x_i, z_i) , i = 1, 2, …, *n* are observed on the sampled units.

Define the sample means of y, x and z as

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,\ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ and\ \bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

Let us propose a class of generalized modified ratio type estimator defined by

$$\hat{\bar{Y}}_{gmr} = \bar{y}\left[\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)^{g_1} + (1-\alpha_1)\left(\frac{\bar{x}}{\bar{X}}\right)^{h_1}\right]^{\delta_1}\left[\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)^{g_2} + (1-\alpha_2)\left(\frac{\bar{z}}{\bar{Z}}\right)^{h_2}\right]^{\delta_2}$$

where $\alpha_1, \alpha_2, g_1, g_2, h_1, h_2, \delta_1\ and\ \delta_2$ are real constants to be determined suitably. $0 < \alpha_1, \alpha_2 < 1$.We may fix $g_1, g_2, h_1, h_2, \delta_1\ and\ \delta_2$ determine the optimum values of $\alpha_1\ and\ \alpha_2$ by minimizing the mean square of $\hat{\bar{Y}}_{gmr}$. Now, write

$$\bar{y} = \bar{Y}(1+e_1), \bar{x} = \bar{X}(1+e_2), \bar{z} = \bar{Z}(1+e_3),$$

where $e_1 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}, e_2 = \frac{\bar{x}-\bar{X}}{\bar{X}}$ and $e_3 = \frac{\bar{z}-\bar{Z}}{\bar{Z}}$

We now have $E(e_1) = E(e_2) = E(e_3) = 0$,

$$V(e_1) = \theta\ C_y^2,\ V(e_2) = \theta\ C_x^2,\ V(e_3) = \theta\ C_x^2$$

$$Cov(e_1, e_2) = \theta C_{yx},\ Cov(e_1, e_3) = \theta C_{yz},\ and\ Cov(e_2, e_3) = \theta C_{xz}$$

where $\theta = \left(\frac{1}{n} - \frac{1}{N}\right)$.

Assuming that $\hat{\bar{Y}}_{gmr}$ is a continuous function of $\bar{y}, \bar{x}$ and $\bar{z}$ and the first and second order partial derivatives of $\hat{\bar{Y}}_{gmr}$ exist, we may expand $\hat{\bar{Y}}_{gmr}$ in a Taylor's series at the point $\bar{y} = \bar{Y}, \bar{x} = \bar{X}$ and $\bar{z} = \bar{Z}$ and write

$$\hat{\bar{Y}}_{gmr} - \bar{Y} = \bar{Y}\left[e_1 - \delta_2\mu_1 + \frac{\delta_2(\delta_2 - 1)}{2}\mu_2\right.$$
$$\left. - \delta_1\lambda_1 + \delta_1\delta_2\mu_1\lambda_1 + \frac{\delta_1(\delta_1 - 1)}{2}\lambda_2 - \delta_2\mu_1 e_1 - \delta_1\lambda_1 e_1 + \ldots\right]$$

where

$$\lambda_1 = \alpha_1\left[\left(g_1+h_1\right)e_2+\left(\frac{h_1(h_1-1)}{2}-\frac{g_1(g_1+1)}{2}\right)e_2^2\right]-\left(h_1e_2+\frac{h_1(h_1-1)}{2}e_2^2\right)$$

$$\lambda_2 = \left\{\alpha_1(g_1+h_1)+h_1\right\}^2 e_2^2$$

$$\mu_1 = \alpha_2\left[\left(g_2+h_2\right)e_3+\left(\frac{h_2(h_2-1)}{2}-\frac{g_2(g_2-1)}{2}\right)e_3^2\right]-\left(h_2e_3+\frac{h_2(h_2-1)}{2}e_3^2\right)$$

$$\mu_2 = \left\{\alpha_2(g_2+h_2)+h_2\right\}^2 e_3^2$$

Retaining the first degree terms of e_1, e_2 and e_3 we have

$$\hat{\bar{Y}}_{gmr}-\bar{Y} \cong \bar{Y}\left[e_1-\delta_1\left\{\alpha_1\left(g_1+h_1\right)e_2-h_1e_2\right\}-\delta_2\left\{\alpha_2\left(g_2+h_2\right)e_3-h_2e_3\right\}\right]$$

$$= \bar{Y}\left[e_1-\delta_1e_2\left\{\alpha_1\left(g_1+h_1\right)-h_1\right\}-\delta_2e_3\left\{\alpha_2\left(g_2+h_2\right)-h_2\right\}\right]$$

Thus, to the first order of approximation, i.e. to $0\left(\frac{1}{n}\right)$, the mean square error (MSE) of $\hat{\bar{Y}}_{gmr}$ is given by

$$MSE(\hat{\bar{Y}})_{gmr} = \theta\bar{Y}^2\left[C_y^2+\delta_1^2\left\{\alpha_1\left(g_1+h_1\right)-h_1\right\}^2C_x^2+\delta_2^2\left\{\alpha_2\left(g_2+h_2\right)-h_2\right\}^2C_z^2\right.$$
$$-2\delta_1\left\{\alpha_1\left(g_1+h_1\right)-h_1\right\}C_{yx}$$
$$-2\delta_2\left\{\alpha_2\left(g_2+h_2\right)-h_2\right\}C_{yz}$$
$$\left.+2\delta_1\delta_2\left\{\alpha_1\left(g_1+h_1\right)-h_1\right\}\left\{\alpha_2\left(g_2+h_2\right)-h_2\right\}C_{xz}\right]$$

Now, minimizing MSE ($\hat{\bar{Y}}_{gmr}$) with respect to α_1 *and* α_2, we have

$$\alpha_{1(opt)} = \frac{h_1}{g_1+h_1}+\frac{1}{\delta_1\left(g_1+h_1\right)}\left[\frac{C_x^2C_{yx}-C_{xz}C_{yz}}{C_x^2C_z^2-C_{xz}^2}\right]$$

$$\alpha_{2(opt)} = \frac{h_2}{g_2+h_2}+\frac{1}{\delta_2\left(g_2+h_2\right)}\left[\frac{C_x^2C_{yz}-C_{xz}C_{yx}}{C_x^2C_z^2-C_{xz}^2}\right]$$

Substituting $\alpha_{1(opt)}$ and $\alpha_{2(opt)}$ in the expression for MSE ($\hat{\bar{Y}}_{gmr}$) we have to $0\left(\frac{1}{n}\right)$,

$$MSE\left(\hat{\bar{Y}}_{gmr}\right)_{opt} = \theta\bar{Y}^2C_y^2\left(1-R_{y.xz}^2\right)$$

where $R_{y.xz}$ is the multiple correlation coefficient of *y* on *x* and *z*.

Also, to $0\left(\frac{1}{n}\right)$ the bias of $\hat{\bar{Y}}_{gmr}$ with optimum values of α_1 *and* α_2 is given by

$$\text{Bias}\left(\hat{\bar{Y}}_{gmr}\right)_{opt} = \theta\bar{Y}\left[\left\{\delta_1\frac{h_1(h_1-1)}{2}+\frac{\left(\delta_1h_1+m_1\right)\left(1+g_1-h_1\right)}{2}+\frac{\delta_1-1}{2\delta_1}m_1^2\right\}C_x^2\right.$$
$$+\left\{\delta_2\frac{h_2(h_2-1)}{2}+\frac{\left(\delta_2h_2+m_2\right)\left(1+g_2-h_2\right)}{2}+\frac{\delta_2-1}{2\delta_2}m_2^2\right\}C_z^2$$
$$\left.-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}\right]$$

where $m_1 = \dfrac{C_z^2C_{yx}-C_{xz}C_{yz}}{C_x^2C_z^2-C_{xz}^2} = \dfrac{C_y}{C_x}\left[\dfrac{\rho_{yx}-\rho_{xz}\rho_{yz}}{1-\rho_{zx}^2}\right]$

$$m_2 = \frac{C_x^2C_{yz}-C_{xz}C_{yx}}{C_x^2C_z^2-C_{xz}^2} = \frac{C_y}{C_z}\left[\frac{\rho_{yz}-\rho_{xz}\rho_{yx}}{1-\rho_{xz}^2}\right]$$

2.1. Some special cases of generalized modified ratio type estimators

Consider nine estimators t_1, t_2, …, t_9 in Table 1, which are special cases of $\hat{\bar{Y}}_{gmr}$ by substituting some simple but arbitrary real constants for g_1, g_2, h_1, h_2, δ_1 and δ_2. The optimal asymptotic mean square errors (optimized with respect to α_1 and α_2) of these cases are equal to the optimal asymptotic *MSE* ($\hat{\bar{Y}}_{gmr}$) in the general case, which is independent of free parameters g_1, g_2, h_1, h_2, δ_1 and δ_2.

Thus, *MSE* (t_1) = *MSE* (t_2) = … = *MSE* (t_9) = *MSE* ($\hat{\bar{Y}}_{gmr}$)

$$= \theta\bar{Y}^2C_y^2\left(1-R_{y.xz}^2\right).$$

The biases of estimators to $0\left(\frac{1}{n}\right)$ in the special cases excepting the constant multiplier $\theta\bar{Y}$ are presented in Table 1.

Table 1. Biases of some modified ratio type estimators

Estimator	g_1	h_1	g_2	h_2	δ_1	δ_2	Bias $\left(\hat{\bar{Y}}_{gmr}\right)/\theta\bar{Y}$
$t_1 = \bar{y}\left[\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)+(1-\alpha_1)\left(\frac{\bar{x}}{\bar{X}}\right)\right]$ x $\left[\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)\left(\frac{\bar{z}}{\bar{Z}}\right)\right]$	1	1	1	1	1	1	$\frac{m_1+1}{2}C_x^2+\frac{m_2+1}{2}C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$
$t_2 = \bar{y}\left[\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)+(1-\alpha_1)\right]$ x $\left[\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)\right]$	1	0	1	0	1	1	$m_1C_x^2+m_2C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$
$t_3 = \bar{y}\left[\alpha_1+(1-\alpha_1)\left(\frac{\bar{x}}{\bar{X}}\right)\right]$ x $\left[\alpha_2+(1-\alpha_2)\left(\frac{\bar{z}}{\bar{Z}}\right)\right]$	0	1	0	1	1	1	$m_1m_2C_{xz}\ -m_1C_{yx}-m_2C_{yz}$
$t_4 = \bar{y}/\left[\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)+(1-\alpha_1)\left(\frac{\bar{x}}{\bar{X}}\right)\right]$ x $\left[\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)\left(\frac{\bar{z}}{\bar{Z}}\right)\right]$	1	1	1	1	-1	-1	$\left(\frac{m_1-1}{2}+m_1^2\right)C_x^2$ $+\left(\frac{m_2-1}{2}+m_2^2\right)C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$
$t_5 = \bar{y}/\left[\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)+(1-\alpha_1)\right]$ x $\left[\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)\right]$	1	0	1	0	-1	-1	$m_1(m_1+1)C_x^2+m_2(m_2+1)C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$
$t_6 = \bar{y}/\left[\alpha_1+(1-\alpha_1)\left(\frac{\bar{x}}{\bar{X}}\right)\right]$ x $\left[\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)\left(\frac{\bar{z}}{\bar{Z}}\right)\right]$	0	1	0	1	-1	-1	$m_1^2C_x^2+m_2^2C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$
$t_7 = \bar{y}\dfrac{\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)+(1-\alpha_1)\left(\frac{\bar{x}}{\bar{X}}\right)}{\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)\left(\frac{\bar{z}}{\bar{Z}}\right)}$	1	1	1	1	1	-1	$\frac{m_1+1}{2}C_x^2+\left(\frac{m_2-1}{2}+m_2^2\right)C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$
$t_8 = \bar{y}\dfrac{\alpha_1\left(\frac{\bar{X}}{\bar{x}}\right)+(1-\alpha_1)}{\alpha_2\left(\frac{\bar{Z}}{\bar{z}}\right)+(1-\alpha_2)}$	1	0	1	0	1	-1	$m_1C_x^2+m_2(m_2+1)C_z^2$ $-m_1C_{yx}-m_2C_{yz}+m_1m_2C_{xz}$

Table 1. Biases of some modified ratio type estimators (cont.)

Estimator	g_1	h_1	g_2	h_2	δ_1	δ_2	Bias $\left(\hat{\bar{Y}}_{gmr}\right)/\theta\bar{Y}$
$t_9 = \bar{y}\dfrac{\alpha_1+(1-\alpha_1)\left(\dfrac{\bar{x}}{\bar{X}}\right)}{\alpha_2+(1-\alpha_2)\left(\dfrac{\bar{z}}{\bar{Z}}\right)}$	0	1	0	1	1	-1	$m_2^2C_z^2 - m_1C_{yx} - m_2C_{yz} + m_1m_2C_{xz}$

2.2.Numerical illustrations

To compare the biases of estimators t_1, t_2, … and t_9 empirically, consider the data on Population-1 and Population-2 referred by Perri (2007) as follows:

<u>Population – 1</u>

The data (Perri, 2007) are taken from the survey of Household Income and Wealth conducted by the Bank of Italy in 2002. The survey covers 8011 Italian households composed of 22148 individuals and 13536 income earners. On the target population comprising of 8011 households three variables – y (the household net disposable income), *x* (household consumption) and *z* (the number of household income earners) were observed and the summary statistics are:

$C_y = 0.787$, $C_x = 0.668$, $C_x = 0.4596$

$\rho_{yx} = 0.74$, $\rho_{yz} = 0.458$ and $\rho_{xz} = 0.348$

<u>Population – 2</u>

The data (Perri, 2007) have been collected by a market research company. The population consists of 2376 points of sale for which three variables are surveyed – the sale area (*y*) in square meters, the number of employees (*x*) and the amount of soft drink sales (*z*) in euro x 1000 in a year. The summary statistics are:

$C_y = 1.285$, $C_x = 2.35$, $C_x = 1.651$

$\rho_{yx} = 0.898$, $\rho_{yz} = 0.861$ and $\rho_{xz} = 0.773$

The absolute biases of t_1, t_2, … and t_9 without constant multiplier $\theta\bar{Y}$ are shown in Table 2.

Table 2. Absolute biases of estimators without constant multiplier $\theta\bar{Y}$

Estimator	Absolute bias without constant multiplier	
	Population – 1	Population – 2
t_1	0.20862	4.29576
t_2	0.09479	1.48338

Table 2. Absolute biases of estimators without constant multiplier $\theta\overline{Y}$ (cont.)

Estimator	Absolute bias without constant multiplier	
	Population – 1	Population – 2
t_3	0.33501	1.14016
t_4	0.14630	3.11798
t_5	0.39732	2.31794
t_6	0.3248	0.30560
t_7	0.02961	1.85373
t_8	0.12702	1.76716
t_9	0.30278	0.85639

Comments**:** The computations show that t_7 for Population -1 and t_6 for Population-2 are least biased.

The asymptotic optimal mean square errors of t_1, t_2, and t_9 to $0\left(\frac{1}{n}\right)$ are the same.

For Population-1: *MSE* (t_1) = *MSE* (t_2) = … = *MSE* (t_9) = $\theta\overline{Y}^2\left[0.367490\right]$

For Population-2: *MSE* (t_1) = *MSE* (t_2) = … = *MSE* (t_9) = $\theta\overline{Y}^2\left[1.445764\right]$

3. Difference-cum-ratio estimators

Consider a difference-cum-ratio estimator defined by

$$T_1 = \left[\overline{y} + \lambda\left(\overline{X} - \overline{x}\right)\right]\left(\frac{\overline{Z}}{\overline{z}}\right)$$

where λ is a real constant or a random variable converging in probability to a constant. λ may be selected in an optimum manner by minimizing the mean square error of T_1 with respect to λ.

Now, T_1 may be expanded in a power series with assumption that $\left|\frac{\overline{z}-\overline{Z}}{\overline{Z}}\right| < 1$ for all the ${}^{N}C_n$ samples. In order to derive the bias and mean square error of T_1 to $0\left(\frac{1}{n}\right)$, we retain terms up to and including second degree of the concerned variables and thus

$$T_1 \cong \overline{Y}\left[1 - e_3 + e_3^2 + e_1 - e_1e_3 - \frac{\lambda}{R}e_2 + \frac{\lambda}{R}e_2e_3\right],$$

where $e_1 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}, \quad e_2 = \frac{\bar{x} - \bar{X}}{\bar{X}}, \quad e_3 = \frac{\bar{z} - \bar{Z}}{\bar{Z}} \quad and \quad R = \frac{\bar{Y}}{\bar{x}}$

$$E(T_1) = \theta\bar{Y}\left[C_z^2 + \left(\frac{\lambda}{R}\right)^2 C_x^2 - C_{yz} + \frac{\lambda}{R}C_{xz}\right] + 0\left(\frac{1}{n^2}\right).$$

Thus, T_1 is a biased estimator of $\bar{Y}$, but the bias decreases with increase in sample size.

The mean square error of T_1 to $0\left(\frac{1}{n}\right)$ is given by

$$MSE(T_1) = \theta\bar{Y}^2\left[C_y^2 + \left(\frac{\lambda}{R}\right)^2 C_x^2 + C_z^2 - 2\left(\frac{\lambda}{R}\right)C_{yx} - C_{yz} + \frac{2\lambda}{R}C_{xz}\right]$$

Minimizing $MSE(T_1)$ with respect to λ, the optimum value of λ is given by

$$\lambda_{opt} = R\frac{C_{yx} - C_{xz}}{C_x^2}$$

$$= \beta_{yx} - \beta_{xz}\left(\frac{\bar{Y}}{\bar{Z}}\right), \beta_{yx} \text{ and } \beta_{xz} \text{ being the}$$

population regression coefficients of y on x and x on z respectively.

Substituting the optimum value of λ in the expression for $MSE(T_1)$, the optimum mean square error of T_1 to $0\left(\frac{1}{n}\right)$ reduces to

$$MSE(T_1)_{opt} = \theta\bar{Y}^2\left[\left(C_y^2 + C_z^2 - 2C_{yz}\right) - \left(\rho_{yx}C_y - \rho_{xz}C_z\right)^2\right]$$

Further, the bias of the optimum estimator to $0\left(\frac{1}{n}\right)$ is given by

$$\text{Bias }(T_1)_{opt} = \theta\bar{Y}\left[C_z^2\left(1 - \rho_{xz}^2\right) - C_yC_z\left(\rho_{yz} - \rho_{yx} + \rho_{xz}\right)\right]$$

In practice, a consistent estimator of λ_{opt} may be substituted in place of λ in T_1.

Alternatively, let us consider the regression-cum-ratio estimator.

$$T_1^* = \left[\bar{y} + b_{yx}\left(\bar{X} - \bar{x}\right)\right]\left(\frac{\bar{Z}}{\bar{z}}\right)$$

This estimator was independently proposed by Mohanty (1967) and Swain (1973).

The large sample mean square error of T_1^* to $0\left(\frac{1}{n}\right)$ is given by

$$MSE\left(T_1^*\right) = \theta\bar{Y}^2\left[\left(C_y^2 + C_z^2 - 2C_{yz}\right) - \rho_{yx}^2 C_y^2 + 2\rho_{yx}\rho_{xz}C_yC_z\right]$$

$$= \theta\bar{Y}^2\left[\left(C_y^2 + C_z^2 - 2C_{yz}\right) - \left(\rho_{yx}C_y - \rho_{xz}C_z\right)^2 + \rho_{xz}^2 C_z^2\right]$$

Now, $MSE\left(T_1^*\right) - MSE\left(T_1\right) = \theta\bar{Y}^2\rho_{xz}^2 C_z^2 \geq 0$

Therefore, T_1^* is less efficient than T_1 in large samples.

Consider a class of difference-cum-ratio estimators defined by

$$T_2 = \left[\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)\right]\left(\frac{\bar{Z}}{\bar{z}}\right)^{\alpha},$$

where α and λ are suitably chosen constants and the optimum α and λ may be obtained by minimizing the approximate mean square error of T_2 with respect to α and λ.

Following the usual procedure of obtaining the expected values and mean square errors of non-linear estimators to $0\left(\frac{1}{n}\right)$, we have

$$E\left(T_2\right) = \bar{Y} + \theta\bar{Y}\left[\frac{\alpha\left(\alpha+1\right)}{2}C_z^2 - \alpha C_{yz} + \lambda\alpha\left(\frac{\bar{X}}{\bar{Y}}\right)C_{xz}\right]$$

$$MSE\left(T_2\right) = \theta\bar{Y}^2\left[C_y^2 - \alpha^2 C_z^2 + \left(\frac{\lambda}{R}\right)^2 C_x^2 - 2\alpha C_{yz} - 2\frac{\lambda}{R}C_{yx} + 2\frac{\alpha\lambda}{R}C_{xz}\right]$$

Minimizing $MSE\left(T_2\right)$ with respect to α and λ, we have

$$\alpha_{opt} = \frac{C_y}{C_z}\left[\frac{\rho_{yz} - \rho_{yx}\rho_{xz}}{1-\rho_{xz}^2}\right]$$

$$\lambda_{opt} = \frac{S_y}{S_x}\left[\frac{\rho_{yx} - \rho_{yz}\rho_{xz}}{1-\rho_{xz}^2}\right]$$

As such, $MSE\left(T_2\right)_{opt} = \theta\bar{Y}^2 C_y^2\left(1 - R_{y.xz}^2\right)$.

Considering an alternative to T_2, replacing λ by b_{yx} the sample regression coefficient of *y* or *x*, we have

$$T_2^* = \left[\bar{y} + b_{yx}\left(\bar{X} - \bar{x}\right)\right]\left(\frac{\bar{Z}}{\bar{z}}\right)^{\alpha},$$

suggested by Khare and Srivastava (1981).

Now,

$$MSE\left(T_2^*\right) = \theta\bar{Y}^2\left[C_y^2 + \alpha^2 C_z^2 + \frac{\beta_{yx}^2}{R^2}C_x^2 - 2\alpha C_{yz} - 2\frac{\beta_{yx}}{R}C_{yx} + 2\alpha\frac{\beta_{yx}}{R}C_{xz}\right] + 0\left(\frac{1}{n^2}\right)$$

Minimizing $MSE\left(T_2^*\right)$ to $0\left(\frac{1}{n}\right)$ with respect to α, we have

$$\alpha_{opt} = \frac{C_y}{C_z}\left(\rho_{yz} - \rho_{yx}\rho_{xz}\right)$$

Thus, $MSE\left(T_2^*\right) = \theta\bar{Y}^2 C_y^2\left[\left(1 - \rho_{yx}^2\right) - \left(\rho_{yz} - \rho_{yx}\rho_{xz}\right)^2\right]$

$$MSE\left(T_2^*\right) - MSE\left(T_2\right) = \theta\bar{Y}^2 C_y^2\left[\frac{\rho_{xz}^2\left(\rho_{yz} - \rho_{yx}\rho_{xz}\right)^2}{\left(1 - \rho_{xz}^2\right)}\right] \geq 0$$

This shows that T_2^* is less efficient than T_2.

4. A generalized class of difference-cum-ratio estimator

Define a generalized class of difference-cum-ratio estimator as

$$T_g = \left[\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)\right]\left[\alpha\left(\frac{\bar{Z}}{\bar{z}}\right)^g + (1-\alpha)\left(\frac{\bar{z}}{\bar{Z}}\right)^h\right]^{\delta}$$

where $\lambda, \alpha, g, h, \text{ and } \delta$ are real constants to be determined suitably and $0 < \alpha < 1$.

Considering only first degree terms in e_1, e_2 *and* e_3, T_g may be linearised as

$$T_g \cong e_1 + \delta\left\{(1-\alpha)h - \alpha g\right\}e_3 - \lambda\frac{\bar{X}}{\bar{Y}}e_2$$

$$MSE\left(T_g\right) = \theta\bar{Y}^2\left[C_y^2 + \delta^2\left\{(1-\alpha)h - \alpha g\right\}^2 C_z^2 + \lambda^2\frac{\bar{X}^2}{\bar{Y}^2}C_x^2\right.$$

$$\left. + 2\delta\left\{(1-\alpha)h - \alpha g\right\}C_{yz} - 2\lambda\frac{\bar{X}}{\bar{Y}}C_{yx} - 2\lambda\frac{\bar{X}}{\bar{Y}}\delta\left\{(1-\alpha)h - \alpha g\right\}C_{xz}\right]$$

Minimizing $MSE\left(T_g\right)$ with respect to α and λ, we have

$$\alpha_{opt} = \frac{1}{\delta(g+h)}\left[\frac{\left(C_z^2C_x^2 - C_{xz}^2\right)\delta h + \left(C_{yz}C_x^2 - C_{yx}C_{xz}\right)}{C_z^2C_x^2 - C_{xz}^2}\right]$$

$$= \frac{h}{g+h} + \frac{1}{\delta(g+h)}\left[\frac{C_{yz}C_x^2 - C_{yx}C_{xz}}{C_z^2C_x^2 - C_{xz}^2}\right]$$

$$\lambda_{opt} = \frac{\overline{Y}}{\overline{X}}\left[\frac{C_z^2C_{yx} - C_{yz}C_{xz}}{\left(C_z^2C_x^2 - C_{xz}^2\right)}\right]$$

Thus, $MSE\left(T_g\right)_{opt} = \theta\overline{Y}^2C_y^2\left(1 - R_{y.xz}^2\right)$

Also, Bias $\left(T_g\right)_{opt} = \theta\overline{Y}\left[\frac{\delta h(h-1)}{2}C_z^2 + \delta\left\{\frac{h}{g+h} + \frac{1}{\delta(g+h)}m_2\right\}\left\{\frac{g(g+1)}{2} - \frac{h(h-1)}{2}\right\}C_z^2\right.$

$$\left. + \frac{\delta(\delta-1)}{2}\frac{m_2^2}{\delta^2}C_z^2 - m_2C_{yz} + m_1m_2C_{xz}\right]$$

where $m_1 = \dfrac{C_{yx}C_z^2 - C_{yz}C_{xz}}{C_z^2C_x^2 - C_{xz}^2}$ and $m_2 = \dfrac{C_{yz}C_x^2 - C_{yx}C_{xz}}{C_z^2C_x^2 - C_{xz}^2}$

Let us consider an alternative class of estimators when λ is substituted by the b_{yx}, the sample regression estimate, may be defined as

$$T_g^* = \left[\overline{y} + b_{yx}\left(\overline{X} - \overline{x}\right)\right]\left[\alpha\left(\frac{\overline{Z}}{\overline{z}}\right)^g + (1-\alpha)\left(\frac{\overline{z}}{\overline{Z}}\right)^h\right]^\delta$$

Following usual procedure of finding an approximate mean square error of T_g^*, it may be seen that to terms of $0\left(\frac{1}{n}\right)$,

$$MSE\left(T_g^*\right) = \theta\overline{Y}^2\left[C_y^2 + \delta^2\left\{(1-\alpha)h - \alpha g\right\}^2C_z^2\right.$$

$$+ \left(\frac{\beta_{yx}}{R}\right)^2C_x^2 + 2\delta\left\{(1-\alpha)h - \alpha g\right\}C_{yz}$$

$$\left. -2\left(\frac{\beta_{yx}}{R}\right)C_{yx} - 2\frac{\beta_{yx}}{R}\delta\left\{(1-\alpha)h - \alpha g\right\}C_{xz}\right]$$

Minimizing the $MSE\left(T_g^*\right)$ with respect to α, we have

$$\alpha_{opt} = \frac{h}{g+h} + \frac{1}{\delta(g+h)} \cdot \frac{C_{yz}C_x^2 - C_{yx}C_{xz}}{C_z^2 C_x^2}$$

$$MSE\left(T_g^*\right)_{opt} = \theta \bar{Y}^2 C_y^2 \left[\left(1-\rho_{yx}^2\right) - \left(\rho_{yx}\rho_{xz} - \rho_{yz}\right)^2\right]$$

As $MSE\left(T_g^*\right) > MSE\left(T_g\right)$, T_g is more efficient than T_g^*

4.1. Some special cases of T_g

In the following Table 3, we compare the biases of some special cases of T_g.

Table 3. Biases of some special cases of T_g

Estimator	g	h	δ	Bias
$T_1 = \left[\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)\right] \text{x} \left[\alpha\left(\frac{\bar{Z}}{\bar{z}}\right) + (1-\alpha)\left(\frac{\bar{x}}{\bar{z}}\right)\right]$	1	1	1	$\theta\bar{Y}\left[\frac{m_2+1}{2}C_z^2 - m_2C_{yz} + m_1m_2C_{xz}\right]$
$T_2 = \left[\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)\right] \text{x} \left[\alpha\left(\frac{\bar{Z}}{\bar{z}}\right) + (1-\alpha)\right]$	1	0	1	$\theta\bar{Y}\left[m_2C_z^2 - m_2C_{yz} + m_1m_2C_{xz}\right]$
$T_3 = \left[\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)\right] \text{x} \left[\alpha + (1-\alpha)\left(\frac{\bar{z}}{\bar{Z}}\right)\right]$	0	1	1	$\theta\bar{Y}\left[m_1m_2C_{xz} - m_2C_{yz}\right]$
$T_4 = \frac{\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)}{\alpha\left(\frac{\bar{Z}}{\bar{z}}\right) + (1-\alpha)\left(\frac{\bar{z}}{\bar{Z}}\right)}$	1	1	-1	$\theta\bar{Y}\left[m_2^2C_z^2 + \frac{m_2-1}{2}C_z^2 + m_1m_2C_{xz} - m_2C_{yz}\right]$
$T_5 = \frac{\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)}{\alpha\left(\frac{\bar{Z}}{\bar{z}}\right) + (1-\alpha)}$	1	0	-1	$\theta\bar{Y}\left[m_2\left(m_2+1\right)C_z^2 - m_2C_{yz} + m_1m_2C_{xz}\right]$
$T_6 = \frac{\bar{y} + \lambda\left(\bar{X} - \bar{x}\right)}{\alpha + (1-\alpha)\left(\frac{\bar{z}}{\bar{Z}}\right)}$	0	1	-1	$\theta\bar{Y}\left[m_2^2C_z^2 + m_1m_2C_{xz} - m_2C_{yz}\right] = 0$

4.2. Numerical illustrations

For the Population-1 and Population-2 considered in section 2.2, the absolute biases without constant multiplier are compared in Table 4.

Table 4. Comparison of absolute biases of estimators T_1 - T_6.

Estimator	Absolute bias without constant multiplier	
	Population – 1	Population – 2
T_1	0.11464	1.51887
T_2	0.39060	0.32265
T_3	0.05028	0.59572
T_4	0.06436	0.92315
T_5	0.08251	0.87949
T_6	0	0

5. An alternative class of difference-cum-ratio estimator

Consider a generalized class of estimators suggested by Tripathi (1970, 80)

$$T_{g1} = \frac{\bar{y} + \lambda_1\left(\bar{X} - \bar{x}\right)}{\bar{z} + \lambda_2\left(\bar{X} - \bar{x}\right)}\bar{Z}$$

Following usual procedure of obtaining an approximate mean square error of T_{g1} to $0\left(\frac{1}{n}\right)$, we have

$$MSE\left(T_{g1}\right) = \theta\bar{Y}^2\left[C_y^2 + \left(\frac{\lambda_2}{R_1} - \frac{\lambda_1}{R}\right)^2 C_x^2 + C_z^2 \right.$$

$$\left. +2\left\{\left(\frac{\lambda_2}{R_1}\right) - \left(\frac{\lambda_1}{R}\right)\right\}C_{yx} - 2C_{yz} - 2\left\{\left(\frac{\lambda_2}{R_1}\right) - \left(\frac{\lambda_1}{R}\right)\right\}C_{xz}\right],$$

where $R_1 = \frac{\bar{Z}}{\bar{X}}$.

Minimizing $MSE\left(T_{g1}\right)$ with respect to λ_1 and λ_2 it may be verified that the minimizing equations are not independent and hence λ_1 and λ_2 cannot be solved uniquely. Therefore, fixing λ_1 (or λ_2) to a suitable real constant, we may solve for λ_2 (or λ_1). Thus, the optimum value for λ_2 in terms of λ_1 is given by

$$\lambda_2 = \frac{R_1}{C_x^2}\left(C_{xz} - C_{yx}\right) + \lambda_1\left(\frac{R_1}{R}\right)$$

The optimum asymptotic mean square error is given by

$$MSE\left(T_{g1}\right)_{opt} = \theta \bar{Y}^2 \left[\left(C_y^2 + C_z^2 - 2C_{yz} \right) - \left(\rho_{xz} C_z - \rho_{yx} C_y \right)^2 \right]$$

This shows that there is reduction in the mean square error for T_{g1} by the difference type of adjustment made for $\bar{y}$ *and* $\bar{z}$, using the second auxiliary variable *x*. It may be mentioned here that the optimum value of λ_2 in T_{g1} is not unique because of the presence of free parameter λ_1. A meaningful optimum estimator is obtained by the choice of $\lambda_1 = \beta_{yx}$, which results in obtaining $\lambda_2 = \beta_{zx}$. Thus, one of the optimum estimators may be obtained as

$$T_{g2} = \frac{\bar{y} + \beta_{yx}\left(\bar{X} - \bar{x}\right)}{\bar{z} + \beta_{zx}\left(\bar{X} - \bar{x}\right)} \bar{Z}$$

As β_{yx} *and* β_{zx} are unknown in practice, we may substitute their consistent sample estimates b_{yx} *and* b_{zx} respectively in T_{g2}. This reduced estimator was proposed by Sahoo (1984). It may be verified that

$$MSE\left(T_{g1}\right)_{opt} = MSE\left(T_{g2}\right)$$

Now, consider the generalized estimator, proposed by Khare and Srivastava (1981) as

$$T_{g3} = \frac{\bar{y} + \lambda_1\left(\bar{X} - \bar{x}\right)}{\left[\bar{z} + \lambda_2\left(\bar{X} - \bar{x}\right)\right]^{\alpha}} \bar{Z}^{\alpha}$$

Minimizing the $MSE(T_{g3})$ to $0\left(\frac{1}{n}\right)$ with respect to λ_1, λ_2 *and* α we get

$$\lambda_1 = \beta_{yx}$$

$$\lambda_2 = \beta_{zx}$$

and $$\alpha = \frac{C_y}{C_z} \frac{\rho_{yz} - \rho_{yx}\rho_{zx}}{1 - \rho_{zx}^2}$$

As such,

$$MSE\left(T_{g3}\right)_{opt} = \theta \bar{Y}^2 C_y^2 \left(1 - R_{y.xz}^2\right)$$

Now, MSE(T_{g1}) – MSE(T_{g3})

$$=\theta\bar{Y}^2\left[C_z\sqrt{1-\rho_{zx}^2}-C_y\,\frac{\rho_{yz}-\rho_{yx}\rho_{xz}}{\sqrt{1-\rho_{xz}^2}}\right]\geq 0,$$

showing thereby that T_{g3} with optimum values of the parameters λ_1,λ_2 *and* α is more efficient than T_{g1}.

5.1. Comparison of mean square errors

In the following the optimal asymptotic mean square errors of $\hat{\bar{Y}}_{\text{Re}g}, T_1, T_1^*, T_2, T_2^*, T_g, T_g^*, T_{g1}, T_{g2}$ and T_{g3} are presented and compared.

Now, $MSE\left(\hat{\bar{Y}}_{\text{Re}g}\right)=\theta\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right)$

$$MSE\left(T_1\right)=\theta\bar{Y}^2\left[\left(C_y^2+C_z^2-2C_{yz}\right)-\left(\rho_{yx}C_y-\rho_{xz}C_z\right)^2\right]$$

$$MSE\left(T_1^*\right)=\theta\bar{Y}^2\left[\left(C_y^2+C_z^2-2C_{yz}\right)-\left(\rho_{yx}C_y-\rho_{xz}C_z\right)^2+\rho_{xz}^2C_z^2\right]$$

$$MSE\left(T_2\right)=\theta\bar{Y}^2C_y^2\left(1-R_{y.xz}^2\right)$$

$$MSE\left(T_2^*\right)=\theta\bar{Y}^2C_y^2\left[\left(1-\rho_{yx}^2\right)-\left(\rho_{yz}-\rho_{yx}\rho_{xz}\right)^2\right]$$

$$MSE\left(T_g\right)=\theta\bar{Y}^2C_y^2\left(1-R_{y.xz}^2\right)$$

$$MSE\left(T_g^*\right)=\theta\bar{Y}^2C_y^2\left[\left(1-\rho_{yx}^2\right)-\left(\rho_{yx}\rho_{xz}-\rho_{yz}\right)^2\right]$$

$$MSE\left(T_{g1}\right)=MSE\left(T_{g2}\right)$$

$$=\theta\bar{Y}^2\left[\left(C_y^2+C_z^2-2C_{yz}\right)-\left(\rho_{yx}C_y-\rho_{xz}C_z\right)^2\right]$$

$$MSE\left(T_{g3}\right)=\theta\bar{Y}^2C_y^2\left(1-R_{y.xz}^2\right)$$

Thus, we find

(i) $MSE\left(T_2\right)=MSE\left(T_g\right)=MSE\left(T_{g3}\right)=\theta\bar{Y}^2C_y^2\left(1-R_{y.xz}^2\right)$

(ii) $MSE\left(T_2^*\right)=MSE\left(T_g^*\right)=\theta\bar{Y}^2C_y^2\left[\left(1-\rho_{yx}^2\right)-\left(\rho_{yx}\rho_{xz}-\rho_{yz}\right)^2\right]$

(iii) $MSE\left(T_1\right)=MSE\left(T_{g1}\right)=MSE\left(T_{g2}\right)$

$$= \theta \bar{Y}^2 \left[\left(C_y^2 + C_z^2 - 2C_{yz} \right) - \left(\rho_{yx} C_y - \rho_{xz} C_z \right)^2 \right]$$

(iv) $MSE\left(T_1^*\right) > MSE\left(T_1\right) = MSE\left(T_{g1}\right) = MSE\left(T_{g2}\right)$

5.2.Numerical illustrations

In the following Table 5 we compare the percent relative efficiencies of the difference-cum-ratio estimators/ regression-cum-ratio estimators with respect to the linear regression estimator with x without adjusting for the second auxiliary variable z , using data on populations referred to in section 2.2. The efficiency is defined as the inverse of the optimal asymptotic mean square error.

Table 5. Comparison of Percent Relative Efficiencies

Estimator	Population – 1		Population – 2	
	$MSE/\theta\bar{Y}^2$	Relative Efficiency	$MSE/\theta\bar{Y}^2$	Relative Efficiency
$\hat{\bar{Y}}_{reg(x)}$	0.28020	100	0.31967	100
$T_1 = T_{g1} = T_{g2}$	0.32083	87	0.70878	45
T_1^*	0.34641	81	2.33752	14
$T_2^* = T_g^*$	0.25531	110	0.27370	117
$T_2 = T_g = T_{g3}$	0.25188	111	0.20546	155

Comments**:** Comparison of efficiencies of the estimators under consideration shows that $T_2 = T_g = T_{g3}$ and $T_2^* = T_g^*$ are more efficient than $T_1 = T_{g1} = T_{g2}$, T_1^* and $\hat{\bar{Y}}_{reg(x)}$.

Further, in case of numerical illustrations under consideration, there has been substantial loss in efficiency in using $T_1 = T_{g1} = T_{g2}$ and T_1^* in place of $\hat{\bar{Y}}_{reg(x)}$ using single auxiliary variable x in a linear regression set up.

REFERENCES

BOWLEY, A. L. (1926). Measurements of precision attained in sampling. Bull. Inst. Internat. Statist, 22, 1-62.

COCHRAN, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. J. Agricultural Sc., 30, 262-275.

COCHRAN, W. G. (1942). Sampling theory when the sampling units are unequal sizes. J. Amer. Stat. Assoc., 37, 199-212.

COCHRAN, W. G. (1977). Sampling Techniques. John Wiley and Sons, New York.

HANSEN, M. H. and HURWITZ, W. N. (1943). On the theory of sampling from finite populations. Ann. Math. Stat., 14, 333-362.

HANSEN, M. H., HURWITZ, W. N. and MADOW, W. G. (1953). Sample survey methods and theory, Vol. 2, John Wiley and Sons, New York.

HANIF, M., HAMMD, N. and SHAHBAZ, M. Q. (2010). Some new regression type estimators in two phase sampling. World Applied Sc. Journal, 8(7), 799-803.

JOHN, S. (1969). On multivariate ratio and product estimators, Biometrika, 56, 533-536.

KHARE, B. B. and SRIVASTAVA, S. R. (1980). On an efficient estimator of population mean using two auxiliary variables, Proc. National Acad. Sc. India, 50(A), 209-214.

KHARE, B. B. and SRIVASTAVA, S. R. (1981). A generalized regression ratio estimator for the population mean using two auxiliary variables. Aligarh J. Statist, 1, 43-51.

MOHANTY, S. (1967). Combination of regression and ratio estimate. J. Ind. Statist., 5, 16-19.

MURTHY, M. N. (1964). Product method of estimation. Sankhya, Series A, 26, 294-307.

NEYMAN, J. (1934). On the two different aspects of representative method : The method of stratified. Sampling and the method of purposive selection. J. Roy. Statist. Soc., 97, 558-606.

NEYMAN. J. (1938). Contributions to the theory of sampling human populations. J. Amer. Statist. Assoc., 33, 101-116.

OLKIN, I. (1958). Multivariate ratio estimation for finite populations, Biometrika, 45, 154-165.

PERRI, P. F. (2007). Improved ratio-cum-product type estimators. Statistics in Transition (New Series), 8, 1, 51-69.

RAJ, DES (1965). On a method of using multi-auxiliary information in sample surveys. J. Amer Stat. Assoc., 60, 270-277.

RAO, P. S. R. S. and MUDHOLKAR, G. S. (1967). Generalized multivariate estimator for the mean of finite populations. J. Amer. Stat. Assoc., 62, 1009-12.

ROBSON, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio type estimators. J. Amer. Stat. Assoc., 52, 511-522.

SAHOO, L. N. (1984). A note on estimation of the population mean using two auxiliary variables. Aligarh. J. Statist., 3 and 4, 63-66.

SHUKLA, G. K. (1965). Multivariate regression estimate. J. Ind. Stat. Assoc., 3, 202-211.

SHUKLA, G. K. (1966). An alternative multivariate ratio estimate for finite population. Bull Cal. Stat. Assoc., 15, 127-134.

SINGH, M. P. (1965). On the estimation of ratio and product of population parameters, Sankhya, Series C, 27, 321-328.

SINGH, M. P. (1967). Ratio-cum-product method of estimation, Metrika, 12, 34-43.

SINGH, M. P. (1967). Multivariate product method of estimation for finite population, J. Ind. Soc. Agri. Statist., 19, 1-10.

SRIVASTAVA, S. K. (1965). An estimate of the mean of a finite population using several auxiliary character, J. Ind. Stat. Assoc., 3, 189-194.

SRIVASTAVA, S. K. (1967). An estimator using auxiliary information in sample surveys. Cal. Stat. Assoc., Bull., 16, 121-132.

SRIVASTAVA, S. K. (1971). Generalized estimator for the mean of a finite population using multi-auxiliary information. J. Amer. Stat. Assoc., 66, 404-407.

SRIVASTAVA, S. K. and JHAJJ, H. S. (1983). A class of estimators of the population mean using multi-auxiliary information. Bull. Cal. Stat. Assoc., 32, 47-56.

SWAIN, A. K. P. C. (1973). Some contributions to the theory of sampling, Ph.D. Thesis submitted to the Utkal University, India.

TRIPATHI, T. P. (1970). Contributions to the sampling theory using multivariate information. Ph.D. Thesis submitted to Punjabi University, Patiala, India.

TRIPATHI, T. P. (1980). A general class of estimators of population ratio. Sankhya Series C, 42, 63-75.

TRIPATHI, T. P. (1987). A class of estimators for population means using multivariate auxiliary information under any sampling design. Aligarh. J. Statist., 7, 49-62.

WATSON, D. J. (1937). The estimation of leaf areas. J. Agricultural Sc. 27, 474.