Application of Ratio-type Estimators
for Small Area in Repeated Surveys

Bhim Singh
College of Horticulture and Forestry, Jhalawar
Maharana Pratap University of Agriculture and Technology, Udaipur
bhimsingh1@gmail.com

B.V.S. Sisodia
N.D. University of Agriculture and Technology, Faizabad
bvs@india.com

Abstract

The present paper deals with the application of ratio type estimator to estimate population mean for small area on current occasion in repeated surveys. The estimation of population mean for small area on current occasion has been proposed using separate ratio-type estimator and combined ratio-type estimator. The mean square errors (MSE) of proposed estimators have also been derived and show that the estimate based on combined ratio-type estimator is more efficient than separate ratio-type estimator, under certain conditions. In addition, this theoretical result is supported by an empirical study.

Keywords: Ratio-type estimators, Small area, Repeated surveys, Mean square errors.

1. Introduction

The early work in the small area estimation can be dated back to 1966 and 1968 when Panse et al. (1966) and Singh (1968) examined the feasibility of using double sampling for estimation of yield at the block level. The technique consisted in selecting a large sample of villages and fields from the block for eye estimation of yield prior to the harvest and combining it with results of crop cutting experiments conducted on a sub-sample for obtaining estimates of average yield at the block level. The technique envisaged presence of a strong positive correlation between eye estimate and the actual yield harvested. However, this approach could not succeed due to various physical constraints. It may be remarked that small area statistics techniques are not the substitute for sample survey approach if the survey approach is within the budget constraints. A comprehensive review of small area statistics was given by Purcell and Kish (1979) and Ghosh and Rao (1994). Many studies dealing with the small area estimation problem have been discussed by various authors viz. Sukhatme and Sukhatme (1970), Gonzalez (1973), Cochran (1977), Ghangurde and Singh (1977), Fay and Herriot (1979), Unnikrishan and Kunte (1995), Rao (1999), Upadhyaya and Singh (1999), Kadilar and Cingi (2005), Singh et al. (2005), Jiang and Lahiri (2006), Singh and Vishwakarma (2007), Chandra and Chambers (2009), Salvati et al. (2010), Tzavidis et al. (2010) and Singh and Sisodia (2007, 2011). Information on many characteristics is required over time for planning and making policy decision. In some of the dynamic populations, the characteristics change rapidly over time. Such interests are reflected in periodic censuses of human population, agriculture, livestock etc. Time interval of such censuses is quite long. Census at smaller time interval is also not feasible because of involvement of huge cost and time. The data on many of the characteristics, even on some of the items in

censuses, are required at smaller time interval; for example, agricultural produce, area under different crops, fertilizer consumption, livestock number and its products, prevalence of rate of diseases, sex ratio, composition of families, level of education, unemployment etc. for formulation of development programmes and to make other important policy decisions. Thus, when a sample survey of a population is conducted over different occasions, i.e. over a particular interval of time, the system is referred to as successive sampling over occasions or repeated surveys. In the present paper it has been attempted to propose the estimates of population mean for small areas on current occasion using separate ratio-type estimator and combined ratio-type estimator and study their properties. An empirical study is carried out to demonstrate the performance of the proposed estimators.

2. Methodology

Consider a population of N units. Assume that the number of units remains the same but the value of the unit changes over occasions. Further, let the variability of the population remain the same over occasions. For simplicity, assume that a sample of size n is drawn by the simple random sampling without replacement form the population, on the first occasion. Out of this sample of size n, a portion m, of the sample is retained (matched) on the second occasion and supplemented by an additional sample of size (unmatched) u selected from the remaining (N-n) units such that size of sample on second occasion will also be n.

For developing the relevant theory, the notation x and y will be used for the values of the characteristics under study on the previous and current occasions, respectively. Assuming that there are g's (g = 1, 2, …, G) similar regions for the small areas d's (d = 1, 2, …., D). Further, let

$\bar{x}'_{.g}$ = sample mean for g-th group at the first occasion based on m_g units that are common to both the occasions.

$\bar{y}'_{.g}$ = sample mean for g-th group at the second occasion based on m_g units that are common to both the occasions.

$\bar{y}''_{.g}$ = sample mean for g-th group at the second occasion based on u_g units that are selected on second occasion only.

$\bar{y}_{.g}$ = sample mean for g-th group at the second occasion based on n_g units.

$\bar{x}_{.g}$ = sample mean for g-th group at the first occasion based on n_g units.

$\bar{x}_d$ = sample mean for small area, d, on the first occasion based on n_d units such that $\bar{x}_d = \sum_{g=1}^{G} W_{dg} \, \bar{x}_{.g}$ where $W_{dg} = \dfrac{N_{dg}}{N_d}$.

Also, $\bar{Y}_d = \sum_{g=1}^{G} W_{dg} \, \bar{Y}_{dg}$, population mean on the second occasion for small area; d =1, 2, …, D.

Now, we proposed estimation methods for population mean $\overline{Y}_d$, based separate and combined ratio type estimators for small areas.

2.1 Separate ratio-type method of estimation

Let W'_{dg} denotes the weight of the (d,g)-th cell based on matched sample, i.e. $W'_{dg} = \dfrac{M_{dg}}{M_d}$, the ratio $R_{.g} = \dfrac{\overline{Y}_{.g}}{\overline{X}_{.g}}$, for g-th group and also Z_{dg} denotes the weight of the (d,g)-th cell based on unmatched sample, i.e. $Z_{dg} = \dfrac{U_{dg}}{U_d}$. However, we will use weight W_{dg} in developing the estimator for $\overline{Y}_d$ based on matched and unmatched samples because $\overline{Y}_d = \sum_{g=1}^{G} W_{dg}\,\overline{Y}_{dg}$.

Thus, a separate ratio-type estimator of the population mean on the second occasion, $\overline{Y}_d$, based on matched samples is developed as

$$\overline{y}_{m_d} = \sum_{g=1}^{G} W_{dg}\,\frac{\overline{y}'_{.g}}{\overline{x}'_{.g}}\cdot \overline{x}_{.g} \tag{2.1.1}$$

To obtain the bias and MSE, let us define $\overline{y}'_{.g} = \overline{Y}_{.g} + \varepsilon_1$, $\overline{x}'_{.g} = \overline{X}_{.g} + \varepsilon_2$, $\overline{x}_{.g} = \overline{X}_{.g} + \varepsilon_3$, using these notations, we have $E(\varepsilon_1) = E(\varepsilon_2) = E(\varepsilon_3) = 0$,

$E(\varepsilon_1^2) = (1/m_g - 1/N_g)S_{y.g}^2$, $E(\varepsilon_2^2) = (1/m_g - 1/N_g)\,S_{x.g}^2$, $E(\varepsilon_3^2) = (1/n_g - 1/\text{-}N_g)S_{x.g}^2$ $E(\varepsilon_1\varepsilon_2) = (1/m_g - 1/N_g)S_{y.g\,x.g}$, $E(\varepsilon_1\varepsilon_3) = (1/n_g - 1/N_g)S_{y.g\,x.g}$, $E(\varepsilon_2\varepsilon_3) = (1/n_g - 1/N_g)\,S_{x.g}^2$

where,

$$S_{x.g}^2 = \frac{1}{N_g - 1}\sum_{d=1}^{D}\sum_{k=1}^{N_g}(X_{dgk} - \overline{X}_{.g})^2\;,\; S_{y.g}^2 = \frac{1}{N_g - 1}\sum_{d=1}^{D}\sum_{k=1}^{N_g}(Y_{dgk} - \overline{Y}_{.g})^2$$

and $\quad S_{y.g\,x.g} = \dfrac{1}{N_g - 1}\sum_{d=1}^{D}\sum_{k=1}^{N_g}(Y_{dgk} - \overline{Y}_{.g})(X_{dgk} - \overline{X}_{.g})$

The bias and MSE of the estimator in equation (2.1.1) are respectively given by

$$B(\overline{y}_{m_d}) = \sum_{g=1}^{G} W_{dg}\,\overline{Y}_{.g}\left(\frac{1}{m_g} - \frac{1}{n_g}\right)\!\left(C_{x.g}^2 - \rho_{.g}C_{y.g}C_{x.g}\right) + \sum_{g=1}^{G} W_{dg}(\overline{Y}_{.g} - \overline{Y}_{dg}) \tag{2.1.2}$$

However, if $\overline{Y}_{.g} \cong \overline{Y}_{dg}$ for all g = 1, 2, …, G, then second term of (2.1.2) will be approximately zero, and (2.1.2) will reduce to

$$B(\overline{y}_{m_d}) = \sum_{g=1}^{G} W_{dg}\,\overline{Y}_{.g}\left(\frac{1}{m_g} - \frac{1}{n_g}\right)\!\left(C_{x.g}^2 - \rho_{.g}C_{y.g}C_{x.g}\right) \tag{2.1.3}$$

$$MSE(\overline{y}_{m_d}) = \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{n_g} - \frac{1}{N_g} \right) S_{y.g}^2 + \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{m_g} - \frac{1}{n_g} \right) \left(S_{y.g}^2 - 2\rho_{.g} R_{.g} S_{y.g} S_{x.g} + R_{.g}^2 S_{x.g}^2 \right) \quad (2.1.4)$$

where $R_{.g} = \dfrac{\overline{Y}_{.g}}{\overline{X}_{.g}}$, estimates of population ratio for the group g; g = 1, 2, …, G.; $S_{x.g}^2$ and $S_{y.g}^2$ are the mean square error for g-th group of values of x and y respectively and also $C_{x.g}$ and $C_{y.g}$ are the coefficient of variation of $x_{.g}$ and $y_{.g}$ respectively and $\rho_{.g}$ is the correlation coefficient between $x_{.g}$ and $y_{.g}$.

Another estimator of the population mean on the second occasion, $\overline{Y}_d$, independent of $\overline{y}_{m_d}$, can be made on the basis of u units selected on the second occasion i.e.

$$\overline{y}_{u_d} = \sum_{g=1}^{G} W_{dg} \, \overline{y}_{.g}'' \quad (2.1.5)$$

Since, $\overline{y}_{u_d}$ is also a biased estimator of $\overline{Y}_d$. However, $E(\overline{y}_{u_d}) \cong \overline{Y}_d$ if $\overline{Y}_{.g} \cong \overline{Y}_{dg}$ for all g = 1, 2, …, G. If it is so, then

The MSE of $\overline{y}_{u_d}$ is given by

$$MSE(\overline{y}_{u_d}) = \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{u_g} - \frac{1}{N_g} \right) S_{y.g}^2 \quad (2.1.6)$$

Therefore, an improved estimator of population mean $\overline{Y}_d$, for small area, d, d = 1, 2, …, D. on current occasion can be obtained by combining (2.1.1) and (2.1.5) as given below

$$\overline{y}_d = \phi_d \, \overline{y}_{m_d} + (1 - \phi_d) \, \overline{y}_{u_d} \quad (2.1.7)$$

where ϕ_d is an arbitrary constant

Since, $\overline{y}_d$ is also a biased estimator and assuming that $\overline{Y}_{.g} = \overline{Y}_{dg}$ for all g = 1, 2, …, G, bias, therefore, is given as

$$B(\overline{y}_d) = \phi_d \sum_{g=1}^{G} W_{dg} \, \overline{Y}_{.g} \left(\frac{1}{m_g} - \frac{1}{n_g} \right) (C_{x.g}^2 - \rho_{.g} C_{y.g} C_{x.g}) \quad (2.1.8)$$

The mean square error (MSE) of $\overline{y}_d$, using (2.6) is given by

$$MSE(\overline{y}_d)_s = \phi_d^2 \, MSE(\overline{y}_{m_d}) + (1 - \phi_d)^2 \, MSE(\overline{y}_{u_d}) \quad (2.1.9)$$

To obtain optimum value of ϕ_d, differentiating (2.1.9) with respect to ϕ_d and equating it to zero as follows:

$$\frac{\partial}{\partial \phi_d} \left[MSE(\overline{y}_d)_s \right] = 2 \, \phi_d \, MSE(\overline{y}_{m_d}) - 2(1 - \phi_d) \, MSE(\overline{y}_{u_d}) = 0$$

From this equation, we obtain

$$\phi_d = \frac{MSE(\bar{y}_{u_d})}{MSE(\bar{y}_{m_d}) + MSE(\bar{y}_{u_d})}$$ (2.1.10)

From (2.8) and (2.10), we have

$$MSE(\bar{y}_d)_s = \frac{MSE(\bar{y}_{m_d})MSE(\bar{y}_{u_d})}{MSE(\bar{y}_{m_d}) + MSE(\bar{y}_{u_d})}$$ (2.1.11)

From (2.1.4), (2.1.6) and (2.1.11), we have

$$MSE(\bar{y}_d)_s =$$

$$\frac{\left\{\sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{n_g} - \frac{1}{N_g}\right)S_{y.g}^2 + \sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{m_g} - \frac{1}{n_g}\right)\left(S_{y.g}^2 - 2\rho_{.g}R_{.g}S_{y.g}S_{x.g} + R_{.g}^2 S_{x.g}^2\right)\right\}\sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{u_g} - \frac{1}{N_g}\right)S_{y.g}^2}{\sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{n_g} - \frac{1}{N_g}\right)S_{y.g}^2 + \sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{m_g} - \frac{1}{n_g}\right)\left(S_{y.g}^2 - 2\rho_{.g}R_{.g}S_{y.g}S_{x.g} + R_{.g}^2 S_{x.g}^2\right) + \sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{u_g} - \frac{1}{N_g}\right)S_{y.g}^2}$$ (2.1.12)

For large N_g, $\dfrac{1}{N_g}$ are negligible, then we have

$$MSE(\bar{y}_d)_S = \frac{\left[\sum_{g=1}^{G} W_{dg}^2\left\{\frac{n_g S_{y.g}^2 + (n_g - m_g)\left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g}R_{.g}S_{y.g}S_{x.g}\right)}{m_g n_g}\right\}\right].\sum_{g=1}^{G} \frac{W_{dg}^2 S_{y.g}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2\left\{\frac{n_g u_g S_{y.g}^2 + u_g(n_g - m_g)\left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g}R_{.g}S_{y.g}S_{x.g}\right) + m_g n_g S_{y.g}^2}{m_g n_g u_g}\right\}}$$

On simplification, we get

$$MSE(\bar{y}_d)_s = \frac{\left[\sum_{g=1}^{G} W_{dg}^2\left\{\frac{n_g S_{y.g}^2 + u_g\left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g}R_{.g}S_{y.g}S_{x.g}\right)}{n_g(n_g - u_g)}\right\}\right].\sum_{g=1}^{G} \frac{W_{dg}^2 S_{y.g}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2\left\{\frac{n_g^2 S_{y.g}^2 + u_g^2\left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g}R_{.g}S_{y.g}S_{x.g}\right)}{n_g u_g(n_g - u_g)}\right\}}$$ (2.1.13)

If $S_{y.g}^2 = S_{x.g}^2 = S^2$, (say) for all g, then

$$MSE(\bar{y}_d)_s = \frac{\left[\sum_{g=1}^{G} W_{dg}^2\left\{\frac{n_g - u_g\rho_{.g}^2 + u_g\left(R_{.g} - \rho_{.g}\right)^2}{n_g(n_g - u_g)}\right\}\right].\sum_{g=1}^{G} \frac{W_{dg}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2\left\{\frac{n_g^2 - u_g^2\rho_{.g}^2 + u_g^2\left(R_{.g} - \rho_{.g}\right)^2}{n_g u_g(n_g - u_g)}\right\}} S^2$$ (2.1.14)

If $R_{.g} = \rho_{.g}$, when $\sigma^2_{y_{.g}} = \sigma^2_{x_{.g}}$, then (2.1.14) will reduce to

$$\text{MSE}(\bar{y}_d)_s = \frac{\left[\sum_{g=1}^{G} W_{dg}^2 \left\{\frac{n_g - u_g \rho_{.g}^2}{n_g(n_g - u_g)}\right\}\right] \cdot \sum_{g=1}^{G} \frac{W_{dg}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2 \left\{\frac{n_g^2 - u_g^2 \rho_{.g}^2}{n_g u_g (n_g - u_g)}\right\}} S^2 \tag{2.1.15}$$

2.2 Combined ratio-type method of estimation

In line with Hansen *et al.* (1946) combined ratio-type estimator, we now proposed combined ratio-type estimator for matched samples on current occasion as:

$$\bar{y}'_{m_d} = \frac{\sum_{g=1}^{G} W_{dg} \bar{y}'_{.g}}{\sum_{g=1}^{G} W_{dg} \bar{x}'_{.g}} . \bar{x}_d \tag{2.2.1}$$

and the estimate for unmatched samples on current occasion is defined earlier in (2.1.5) as

$$\bar{y}_{u_d} = \sum_{g=1}^{G} W_{dg} \bar{y}''_{.g} \tag{2.2.2}$$

An improved estimator of population mean $\bar{Y}_d$, for small area d; d = 1, 2, …, D. on current occasion can be obtained by combining (2.2.1) and (2.2.2) given as:

$$\bar{y}_d = \varphi_d \bar{y}'_{m_d} + (1 - \varphi_d) \bar{y}_{u_d} \tag{2.2.3}$$

where φ_d is an arbitrary constant to be obtained.

To obtain the bias and MSE, let us define $\sum_{g=1}^{G} W_{dg} \bar{y}'_{.g} = \bar{Y}_d + \eta_1$, $\sum_{g=1}^{G} W_{dg} \bar{x}'_{.g} = \bar{X}_d + \eta_2$, $\bar{x}_d = \bar{X}_d + \eta_3$, using these notations, we have E(η_1) = E(η_2) = E(η_3) =0,

$$E(\eta_1\eta_2) = \sum_{g=1}^{G} W_{dg}^2 (1/m_g - 1/N_g) S_{y_{.g} x_{.g}}, E(\eta_1\eta_3) = \sum_{g=1}^{G} W_{dg}^2 (1/n_g - 1/N_g) S_{y_{.g} x_{.g}},$$

$$E(\eta_2\eta_3) = \sum_{g=1}^{G} W_{dg}^2 (1/n_g - 1/N_g) S_{x_{.g}}^2, E(\eta_1^2) = \sum_{g=1}^{G} W_{dg}^2 (1/m_g - 1/N_g) S_{y_{.g}}^2,$$

$$E(\eta_2^2) = \sum_{g=1}^{G} W_{dg}^2 (1/m_g - 1/N_g) S_{x_{.g}}^2, E(\eta_3^2) = \sum_{g=1}^{G} W_{dg}^2 (1/n_g - 1/N_g) S_{x_{.g}}^2,$$

The bias and MSE of the estimator in equation (2.2.1) is given by

$$B(\bar{y}\,'_{m_d}) = \bar{Y}_d \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{m_g} - \frac{1}{n_g} \right) \left(C_{x_{.g}}^2 - \rho_{.g} C_{y_{.g}} C_{x_{.g}} \right) \tag{2.2.4}$$

where $C_{x_{.g}}$ and $C_{y_{.g}}$ are the coefficient of variation of $x_{.g}$ and $y_{.g}$ and $\rho_{.g}$ is the correlation coefficient between $x_{.g}$ and $y_{.g}$.

$$MSE(\bar{y}\,'_{m_d}) = \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{n_g} - \frac{1}{N_g} \right) S_{y_{.g}}^2 + \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{m_g} - \frac{1}{n_g} \right) \left(S_{y_{.g}}^2 - 2\rho_{.g} R_d S_{y_{.g}} S_{x_{.g}} + R_d^2 S_{x_{.g}}^2 \right) \tag{2.2.5}$$

where $R_d = \dfrac{\bar{Y}_d}{\bar{X}_d}$, estimates of population mean ratio for small area d.

also MSE of $\bar{y}_{u_d}$ is given by

$$MSE(\bar{y}_{u_d}) = \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{u_g} - \frac{1}{N_g} \right) S_{y_{.g}}^2 \tag{2.2.6}$$

and applying the same procedure for (2.2.2), the bias and MSE are respectively found as follows:

$$B(\bar{y}_d) = \varphi_d \bar{Y}_d \sum_{g=1}^{G} W_{dg}^2 \left(\frac{1}{m_g} - \frac{1}{n_g} \right) \left(C_{x_{.g}}^2 - \rho_{.g} C_{y_{.g}} C_{x_{.g}} \right) \tag{2.2.7}$$

MSE of $\bar{y}_d$, using (2.2.3) is given by

$$MSE(\bar{y}_d)_c = \varphi_d^2 \, MSE(\bar{y}\,'_{m_d}) + (1 - \varphi_d)^2 \, MSE(\bar{y}_{u_d}) \tag{2.2.8}$$

In order to find the equation of φ_d which makes the MSE minimum, we should take the derivative of the MSE with respect to φ_d and equating this equation to zero as follows:

$$\frac{\partial}{\partial \varphi_d} \left[MSE(\bar{y}_d)_c \right] = 2\varphi_d MSE(\bar{y}\,'_{m_d}) - 2(1 - \varphi_d) MSE(\bar{y}_{u_d}) = 0$$

From this equation, we obtain

$$\varphi_d = \frac{MSE(\bar{y}_{u_d})}{MSE(\bar{y}\,'_{m_d}) + MSE(\bar{y}_{u_d})} \tag{2.2.9}$$

From (2.2.3) and (2.2.8), we have

$$MSE(\bar{y}_d)_c = \frac{MSE(\bar{y}\,'_{m_d}) MSE(\bar{y}_{u_d})}{MSE(\bar{y}\,'_{m_d}) + MSE(\bar{y}_{u_d})} \tag{2.2.10}$$

Putting (2.2.6), (2.2.8) in (2.2.10), we obtain

$$
MSE(\bar{y}_d)_c = \frac{\left\{\sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{n_g}-\frac{1}{N_g}\right)S_{y.g}^2 + \sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{m_g}-\frac{1}{n_g}\right)\left(S_{y.g}^2 - 2\rho_{.g}R_d S_{y.g} S_{x.g} + R_d^2 S_{x.g}^2\right)\right\}\sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{u_g}-\frac{1}{N_g}\right)S_{y.g}^2}{\sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{n_g}-\frac{1}{N_g}\right)S_{y.g}^2 + \sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{m_g}-\frac{1}{n_g}\right)\left(S_{y.g}^2 - 2\rho_{.g}R_d S_{y.g} S_{x.g} + R_d^2 S_{x.g}^2\right) + \sum_{g=1}^{G} W_{dg}^2\left(\frac{1}{u_g}-\frac{1}{N_g}\right)S_{y.g}^2}
$$

$$(2.2.11)$$

For large N_g, $\dfrac{1}{N_g}$ are negligible, then after simplification, the MSE of combined ratio type estimator is given by

$$
MSE(\bar{y}_d)_c = \frac{\left[\sum_{g=1}^{G} W_{dg}^2\left\{\dfrac{n_g S_{y.g}^2 + u_g\left(R_d^2 S_{x.g}^2 - 2\rho_{.g}R_d S_{y.g} S_{x.g}\right)}{n_g(n_g-u_g)}\right\}\right].\sum_{g=1}^{G}\dfrac{W_{dg}^2 S_{y.g}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2\left\{\dfrac{n_g^2 S_{y.g}^2 + u_g^2\left(R_d^2 S_{x.g}^2 - 2\rho_{.g}R_d S_{y.g} S_{x.g}\right)}{n_g u_g(n_g-u_g)}\right\}}
$$

$$(2.2.12)$$

If $S_{y.g}^2 = S_{x.g}^2 = S^2$, (say) for all g, then

$$
MSE(\bar{y}_d)_c = \frac{\left[\sum_{g=1}^{G} W_{dg}^2\left\{\dfrac{n_g - u_g\rho_{.g}^2 + u_g\left(R_d - \rho_{.g}\right)^2}{n_g(n_g-u_g)}\right\}\right].\sum_{g=1}^{G}\dfrac{W_{dg}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2\left\{\dfrac{n_g^2 - u_g^2\rho_{.g}^2 + u_g^2\left(R_d - \rho_{.g}\right)^2}{n_g u_g(n_g-u_g)}\right\}} S^2
$$

$$(2.2.13)$$

If $R_d = \rho_{.g}$, when $\sigma_{y.g}^2 = \sigma_{x.g}^2$, then

$$
MSE(\bar{y}_d)_c = \frac{\left[\sum_{g=1}^{G} W_{dg}^2\left\{\dfrac{n_g - u_g\rho_{.g}^2}{n_g(n_g-u_g)}\right\}\right].\sum_{g=1}^{G}\dfrac{W_{dg}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2\left\{\dfrac{n_g^2 - u_g^2\rho_{.g}^2}{n_g u_g(n_g-u_g)}\right\}} S^2
$$

$$(2.2.14)$$

The equations (2.1.15) and (2.2.14) shows that these are similar under the following conditions $R_{.g} = \rho_{.g}$, $R_d = \rho_{.g}$ and $\sigma_{y.g}^2 = \sigma_{x.g}^2$.

3. Efficiency comparisons

It is interesting to compare (2.1.13) and (2.2.12) and to note that the MSE of the combined ratio estimate has the same form as that of the ratio estimate based on separate

group except that there is now a single ratio R_d in place of $R_{.g}$. The difference between the mean square errors of (2.1.13) and (2.2.12) can be expressed as

$$\text{MSE}\,(\overline{y}_d)_c - \text{MSE}\,(\overline{y}_d)_s =$$

$$\frac{\left[\sum_{g=1}^{G} W_{dg}^2 \left\{ \frac{n_g S_{y.g}^2 + u_g \left(R_d^2 S_{x.g}^2 - 2\rho_{.g} R_d S_{y.g} S_{x.g}\right)}{n_g(n_g - u_g)} \right\}\right] \cdot \sum_{g=1}^{G} \frac{W_{dg}^2 S_{y.g}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2 \left\{ \frac{n_g^2 S_{y.g}^2 + u_g^2 \left(R_d^2 S_{x.g}^2 - 2\rho_{.g} R_d S_{y.g} S_{x.g}\right)}{n_g u_g (n_g - u_g)} \right\}}$$

$$- \frac{\left[\sum_{g=1}^{G} W_{dg}^2 \left\{ \frac{n_g S_{y.g}^2 + u_g \left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g} R_{.g} S_{y.g} S_{x.g}\right)}{n_g(n_g - u_g)} \right\}\right] \cdot \sum_{g=1}^{G} \frac{W_{dg}^2 S_{y.g}^2}{u_g}}{\sum_{g=1}^{G} W_{dg}^2 \left\{ \frac{n_g^2 S_{y.g}^2 + u_g^2 \left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g} R_{.g} S_{y.g} S_{x.g}\right)}{n_g u_g (n_g - u_g)} \right\}}$$

On simplification, we get

$$\text{MSE}\,(\overline{y}_d)_c - \text{MSE}\,(\overline{y}_d)_s =$$

$$\frac{\left[\sum_{g=1}^{G} W_{dg}^4 S_{y.g}^2 \left\{ \frac{(R_d - R_{.g})^2 S_{x.g}^2 + 2(R_d - R_{.g})(R_{.g} S_{x.g}^2 - \rho_{.g} S_{y.g} S_{y.g})}{n_g(n_g - u_g)} \right\}\right] \cdot \sum_{g=1}^{G} \frac{W_{dg}^2 S_{y.g}^2}{u_g}}{A.B}$$

(3.1)

where,

$$A = \sum_{g=1}^{G} W_{dg}^2 \left\{ \frac{n_g^2 S_{y.g}^2 + u_g^2 \left(R_d^2 S_{x.g}^2 - 2\rho_{.g} R_d S_{y.g} S_{x.g}\right)}{n_g u_g (n_g - u_g)} \right\}$$

(3.2)

$$B = \sum_{g=1}^{G} W_{dg}^2 \left\{ \frac{n_g^2 S_{y.g}^2 + u_g^2 \left(R_{.g}^2 S_{x.g}^2 - 2\rho_{.g} R_{.g} S_{y.g} S_{x.g}\right)}{n_g u_g (n_g - u_g)} \right\}$$

(3.3)

It will be seen that (3.1) depends upon the magnitude of the variation between the group ratio and the value of $\left(R_{.g} S_{x.g}^2 - \rho_{.g} S_{y.g} S_{x.g}\right)$. The latter will be small, vanishing in fact when the regression of the y on x is a straight line through the origin within each group. It follows, therefore, that the combined estimate will have a lower precision than that based on separate group. On the other hand, the bias in the former estimate will be smaller than in the latter. Unless, therefore, the population ratios in the different group vary considerably, the use of a combined ratio would provide an estimate which has a negligible bias and whose precision is almost as high as that of the estimate based on separate ratios.

4. Application

The study was confined to Jorium Gram Panchayat of Faizabad district of state of Uttar Pradesh, INDIA. A survey is conducted of the all households/farmers having milk cattle. These households/farmers were further classified according to land holding size (category of the households/farmers) and breeds of cattle being reared by the households. The objectives of the study are to estimate the population mean (milk yield, litre/day) on current occasion (2nd occasion) in different categories of farmers/households. Therefore, socio-economic classes, i.e. categories of farmers are considered as small areas. The breeds of the cattle are considered as groups. Therefore, target population is divided into non-overlapping small areas (categories of farmers). It is also divided into non-overlapping groups according to breeds of the cattle. The population, however, could be classified into different non-overlapping groups according to lactation number of cattle, date of calving etc in order to carry out cohort studies. But in the present study, it is confined to only breed-wise grouping i.e. indigenous cattle and exotic cattle. In small area estimation with two-way classification of the population, it is essential assumption that small areas must be of similar characteristics across the groups. It is expected that categories of the farmers would be having similar characteristics across the groups, and that is why categories of the farmers are considered as small areas in the present empirical study. The estimates of population means (milk yield, litre/day) for small areas (marginal farmers and other farmer's category) on current occasion (2nd occasion) have been worked out using two ratio-type estimators as described in Section 2. The summary statistics of the data are given in Table 1. From Table 1 the average milk yields (litre/day), using separate ratio-type estimation method, on current occasion (likely end of monsoon season) are 0.554 and 1.174 for marginal farmers and other farmer's category respectively. In the similar ways, the average milk yields (litre/day) using combined ratio-type estimation method on current occasion (likely end of monsoon season) are 0.753 and 0.859 for marginal farmers and other farmers category respectively. The MSE values of separate and combined ratio estimators have also been computed and are given in Table 2. This indicates that the MSE(s) of separate ratio-type estimator is found to be greater than the combined ratio-type estimator in case of marginal farmer's category as well as other farmer's category. Also, the variation within group is considerably large and the bias of the separate ratio estimator is greater than the combined ratio estimator. Therefore, combined ratio-type estimator is more appropriate than the separate ratio-type estimator.

5. Conclusion

We have derived the method of estimation of population mean using separate and combined ratio type estimators for small area on current occasion in repeated surveys and obtained its MSE equation. By these equations, the MSE of the proposed estimators have been compared and by this comparison it has been found that the combined ratio type estimator has smaller MSE than the separate ratio type estimator, a small sample in each group is present and also bias of the separate ratio estimator is greater than the combined ratio estimator. Therefore, combined ratio-type estimator is more efficient that separate ratio type estimator. In the fourth coming studies, we hope to improve the proposed estimators.

References

1. Chandra, H. and Chambers, R. (2009). Multipurpose weighting for small area estimation. *J. Official Stat.* 25(3), 379-395.

2. Cochran, W.G. (1963). *Sampling techniques*. Second Edition, John Wiley and Sons, New York, London.

3. Fay, R.E. and Herriot, R.A. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *J. Amer. Statist. Assoc.* 74, 269-277.

4. Ghangurde, P.D. and Singh, M.P. (1977). Synthetic estimates in periodic household surveys. *Survey Meth.* 3, 152-181.

5. Ghose, M. and Rao, J.N.K. (1994). Small area estimation: an appraisal. *Stat. Sci.* 9(1), 55-93.

6. Gonzalez, M.E. (1973). Use and evaluation of synthetic estimates. Proceedings of the Social Statistics Section. *American Statistical Association*, 33-36.

7. Hansen, Hurwitz and Gurney (1946). Problem and methods of the sample survey of business. *Journal of American Statistical Association*, 41, 174-189.

8. Jiang, J. and Lahiri, P. (2006). Mixed model prediction and small area estimation (with discussions). Test, 15(1), 1–96.

9. Kadilar, C. and Cingi, H. (2005). A New Ratio Estimator in Stratified Random Sampling. *Commun. in Statist. Theo. and Meth.*, 34(3), 597–602.

10. Panse, V.G., Rajagopalan, M. and Pillai, S. (1966). Estimation of crop yields for small areas. *Biometrics*, 66, 374-388.

11. Purcell, N.J. and Kish, L. (1979). Estimation of small domains. *Biometrics*, 35, 365-384.

12. Rao, J.N.K. (1999). Some recent advances in model-based small area estimation. *Survey Meth.*, 25, 175–186.

13. Salvati, N., Chandra, H., Ranalli, M.G. and Chambers, R. (2010). Small area estimation using a nonparametric model-based direct estimator. *Computational Statistics and Data Analysis*, 54, 2159-2171.

14. Singh, B. (2007). On Small Area Estimation in Repeated Surveys. PhD Thesis Submitted to N.D.U.A. & T., Faizabad, U.P., India.

15. Singh, B. and Sisodia, B.V.S. (2011). Methods of Estimation of Change for Small Area in Repeated Surveys. *Internat. J. of Statist. and Syst.*, 6(1), 25–33.

16. Singh, B.B., Shukla, G.K. and Kundu, D. (2005). Spatio-temporal models in small area estimation. *Survey Meth.*, 31, 183-195.

17. Singh, D. (1968). Estimates in successive sampling using a multistage design. *Journal of American Statistical Association*, 63, 99-112.

18. Singh, H.P. and G.K. Vishwakarma (2007). Modified exponential ratio and product estimators for finite population mean in double sampling. *Austrian J. Statist.*, 36(3), 217-225.

19. Sukhatme, P.V. and Sukhatme, B.V. (1970). Sampling Theory of surveys with Applications, Iowa State University Press, Ames, IA.

20. Tzavidis, N., Marchetti, S., and Chambers, R. (2010). Robust prediction of small area means and quantiles. *Australian and New Zealand Journal of Statistics*, 52, 167-186.

21. Unnikrishan, N. K. and S. Kunte (1995). Optimality of analogue of Basu's estimator under double sampling design. *Sankhya* B 57, 103-111.

22. Upadhyaya, L. N. and H. P. Singh (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometr. J.*, 41, 627-636.

Table 1: Data statistics of population parameters

Marginal Farmer	Other Farmer
$\bar{x}_{.1} = 2.355$	$\bar{x}_{.2} = 3.550$
$\bar{y}_{.1} = 0.917$	$\bar{y}_{.2} = 1.068$
$R_{.1} = 0.3893843$	$R_{.2} = 0.3008451$
$s_{x_{.1}}^2 = 1.6271930$	$s_{x_{.2}}^2 = 14.8722222$
$s_{y_{.1}}^2 = 0.7279412$	$s_{y_{.2}}^2 = 3.7261364$
$\rho_{.1} = 0.6469467$	$\rho_{.2} = 0.9880312$
$\bar{X}_1 = 2.594$	$\bar{X}_2 = 3.073$
$\bar{Y}_1 = 0.947$	$\bar{Y}_2 = 1.008$
$R_1 = 0.3650732$	$R_2 = 0.3281250$
$\bar{y}_{m_1} = 0.6804068$	$\bar{y}_{m_2} = 0.9584618$
$\bar{y}_{u_1} = 0.9561250$	$\bar{y}_{u_2} = 1.3092430$
$\phi_1 = 0.7837673$	$\phi_2 = 0.3866321$
$\bar{y}_{1S} = 0.5544385$	$\bar{y}_{2S} = 1.1734707$
$\bar{y}'_{m_1} = 0.7002814$	$\bar{y}'_{m_2} = 0.3192090$
$\varphi_1 = 0.7951316$	$\varphi_2 = 0.4551932$
$\bar{y}_{1C} = 0.7526701$	$\bar{y}_{2C} = 0.8584539$
$B(\bar{y}_1)_S = 0.0026735$	$B(\bar{y}_2)_S = 0.0047015$
$B(\bar{y}_1)_C = 0.0000833$	$B(\bar{y}_2)_C = 0.0031568$

Table 2: MSE value of proposed estimators

Separate estimator		Combined estimator	
Marginal Farmer	Other Farmer	Marginal Farmer	Other Farmer
$MSE(\bar{y}_{m_1}) = 0.0269951$	$MSE(\bar{y}_{m_2}) = 0.0606841$	$MSE(\bar{y}'_{m_1}) = 0.0252108$	$MSE(\bar{y}'_{m_2}) = 0.0457824$
$MSE(\bar{y}_1)_s = 0.0211579$	$MSE(\bar{y}_2)_s = 0.0234624$	$MSE(\bar{y}_1)_c = 0.0200459$	$MSE(\bar{y}_2)_c = 0.0208398$