

# Estimation of Finite Population Mean using Random Non-Response in Survey Sampling

Housila P. Singh  
School of Studies in Statistics  
Vikram University, Ujjain-456010  
India  
hpsujn@gmail.com

Ramkrishna S. Solanki  
School of Studies in Statistics  
Vikram University, Ujjain-456010  
India  
ramkssolanki@gmail.com

## Abstract

This paper consider the problem of estimating the population mean under three different situations of random non-response envisaged by Singh et al (2000). Some ratio and product type estimators have been proposed and their properties are studied under an assumption that the number of sampling units on which information can not be obtained owing to random non-response follows some distribution. The suggested estimators are compared with the usual ratio and product estimators. An empirical study is carried out to show the performance of the suggested estimators over usual unbiased estimator, ratio and product estimators. A generalized version of the proposed ratio and product estimators is also given.

**Keywords:** Auxiliary variable, Study variate, Random non-response, Bias and Variance.

## 1. Introduction

It is well recognized that the suitable use of auxiliary in probability sampling results in considerable reduction in the variance of the estimators of the population mean. Ratio, product and regression methods of estimation are good examples in this context. When the correlation between the study variate and the auxiliary variate is positive (high), the ratio method of estimation envisaged by Cochran (1940) is effectively employed. On the other hand if this correlation is negative high the product method of estimation propounded by Robson (1957) and rediscovered by Murthy (1964) is used. Later various ramifications of classical ratio and product estimators have been presented with their properties by different authors assuming that all the observations on selected units in the sample are available for instance, see Cochran (1977), Sukhatme et al (1984). Krishnaiah and Rao (1988), Chaudhuri and Vos (1988) and Mangat (1996). However in many practical situations encountered in sample surveys observations are not available for all the selected units in the sample i.e. some observations may be missing for various reasons such as unwillingness of some selected units to supply the desired information, accidental loss of information caused by unknown factors, failure on the part of investigator to gather correct information, see Toutenburg and Srivastava (1988). In fact, missingness of

observation are usually encountered in opinion polls, market research surveys, mail enquiries, socio economic investigation, medical studies and many other experiments. Statisticians have identified for some time that failure to account for the stochastic nature of incompleteness can spoil inference. An obvious question arises what one needs to assume to justify ignoring the incomplete mechanism. Rubin (1976) advocated three concepts: missing at random (MAR), observed at random (OAR) and parameter distribution (PD). Rubin defined “the data are MAR if the probability of the observed missing ness pattern, given the observed and unobserved data, does not depend on the value of the unobserved data”. Heitzan and Basu (1996) have distinguished the meaning of missing at random (MAR) and missing completely at random (MCAR) in a very nice way, see Singh et al (2000). Tracy and Osahan (1994) studied the effect of random non-response on the conventional ratio estimator of the population mean in two situations: (i) non-response on the study as well as the auxiliary variables and (ii) non-response on the study variable only. Singh et al (2000) suggested three regression type estimators, which are further generalized by Singh and Tracy (2001) and Singh et al (2007), in the presence of random non-response in different situations under an assumption that the number of sampling units on which information can not be obtained due to random non response follow some distribution. Singh and Joarder (1998), Singh et al (2003) and Ahmed et al (2005) have studied the effect of random non-response on different estimators of population variance.

In this paper we have studied the effect of random non-response on the Tuteja and Bahl (1991) ratio and product estimators and on their generalized version in different situations under an assumption that the number of sampling units on which information can not be obtained owing to random non-response follows some distribution.

## 2. Distribution of Random Non-response and Some Expected values

Let  $U = (U_1, U_2, \dots, U_N)$  denote the population of units from which a sample random sample of size  $n$  is drawn without replacement. If  $r \{r = 0, 1, 2, \dots, (n-2)\}$  denotes the number of sampling units on which information could not be obtained due to random non-response, and then the remaining  $(n-r)$  units in the sample can be treated as simple random sampling without replacement (SRSWOR) sample from  $U$ . We assume that  $r$  lies between 0 and  $(n-2)$  i.e.  $0 \leq r \leq (n-2)$ . We assume that if  $p$  denotes the probability of non-response among the  $(n-2)$  possible values of responses, then  $r$  has the discrete distribution:

$$P(r) = \frac{(n-r)}{(np+2p)} {}^{n-2}C_r p^r q^{n-2-r} \quad (2.1)$$

where  $q = (1-p)$  and  $r = 0, 1, 2, \dots, (n-2)$ .

Let us define,

$$\varepsilon_{o(n-r)} = \frac{\bar{y}_{(n-r)}}{\bar{Y}} - 1, \quad \varepsilon_{l(n-r)} = \frac{\bar{x}_{(n-r)}}{\bar{X}} - 1, \quad \varepsilon_1 = \frac{\bar{x}}{\bar{X}} - 1, \quad \bar{y}_{(n-r)} = (n-r)^{-1} \sum_{i=1}^{n-r} y_i,$$

$$\begin{aligned}\bar{X}_{(n-r)} &= (n-r)^{-1} \sum_{i=1}^{n-r} x_i, \quad \bar{x} = n^{-1} \sum_{i=1}^n x_i, \quad s_{y(n-r)}^2 = (n-r-1)^{-1} \sum_{i=1}^{n-r} \{y_i - \bar{y}_{(n-r)}\}^2, \\ s_{x(n-r)}^2 &= (n-r-1)^{-1} \sum_{i=1}^{n-r} \{x_i - \bar{x}_{(n-r)}\}^2, \quad s_x^2 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2, \\ s_{xy(n-r)} &= (n-r-1)^{-1} \sum_{i=1}^{n-r} \{x_i - \bar{x}_{(n-r)}\} \{y_i - \bar{y}_{(n-r)}\}, \quad S_y^2 = (N-1)^{-1} \sum_{i=1}^N (y_i - \bar{Y})^2, \\ S_x^2 &= (N-1)^{-1} \sum_{i=1}^N (x_i - \bar{X})^2, \quad S_{xy} = (N-1)^{-1} \sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y}), \quad \bar{X} = N^{-1} \sum_{i=1}^N x_i \text{ and} \\ \bar{Y} &= N^{-1} \sum_{i=1}^N y_i.\end{aligned}$$

The probability model defined at (2.1) is free from actual data values, hence it can be considered as a model suitable for MAR situation. Then under the probability model given by (2.1), we have following results:

$$\begin{aligned}E\{\varepsilon_{o(n-r)}\} &= E\{\varepsilon_{l(n-r)}\} = E(\varepsilon_1) = 0, \\ E\{\varepsilon_{o(n-r)}^2\} &= \left[\frac{1}{(nq+2p)} - \frac{1}{N}\right] C_y^2, \quad E\{\varepsilon_{l(n-r)}^2\} = \left[\frac{1}{(nq+2p)} - \frac{1}{N}\right] C_x^2, \quad E(\varepsilon_1^2) = \left(\frac{1}{n} - \frac{1}{N}\right) C_x^2, \\ E\{\varepsilon_{o(n-r)} \varepsilon_{l(n-r)}\} &= \left[\frac{1}{(nq+2p)} - \frac{1}{N}\right] \rho C_y C_x, \quad E\{\varepsilon_{o(n-r)} \varepsilon_1\} = \left(\frac{1}{n} - \frac{1}{N}\right) \rho C_y C_x, \\ E\{\varepsilon_{l(n-r)} \varepsilon_1\} &= \left(\frac{1}{n} - \frac{1}{N}\right) C_x^2\end{aligned}$$

where,

$$C_y = S_y / \bar{Y}, \quad C_x = S_x / \bar{X}, \quad \rho = S_{xy} / (S_x S_y),$$

It may be noted that if  $p = 0$  i.e. there is no non-response, the above expected values coincides with the usual results.

### 3. Proposed Strategies for Finite Population Mean

#### 3.1 Strategy – I:

We are considering the situation when random non-response exists on both the study variable  $y$  and the auxiliary variable  $x$  and population mean  $\bar{X}$  of the auxiliary variable  $x$  is known. The classical ratio and product estimators in this situation are respectively defined by,

$$t_{IR} = \bar{y}_{(n-r)} \left\{ \frac{\bar{X}}{\bar{X}_{(n-r)}} \right\} \quad (3.1)$$

$$t_{IP} = \bar{y}_{(n-r)} \left\{ \frac{\bar{X}_{(n-r)}}{\bar{X}} \right\} \quad (3.2)$$

To the first degree of approximation, the biases and variances of  $t_{1R}$  and  $t_{1P}$  are respectively given by,

$$B(t_{1R}) = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] \left( \frac{S_x^2}{\bar{X}} \right) (R - \beta) \quad (3.3)$$

$$B(t_{1P}) = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] \left( \frac{S_x^2 \beta}{\bar{X}} \right) \quad (3.4)$$

$$\text{Var}(t_{1R}) = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] [S_y^2 + R S_x^2 (R - 2\beta)] \quad (3.5)$$

$$\text{Var}(t_{1P}) = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] [S_y^2 + R S_x^2 (R + 2\beta)] \quad (3.6)$$

where  $\beta = S_{xy} / S_x^2$  is the population regression coefficient of  $y$  on  $x$  and

$\sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y}) / (N - 1)$  and  $S_x^2$  is defined earlier.

We suggest the modified ratio and product estimators on the lines of Tuteja and Bahl (1991) respectively as,

$$t_{1Re} = \bar{y}_{(n-r)} \exp \left\{ \frac{\bar{X} - \bar{x}_{(n-r)}}{\bar{X} + \bar{x}_{(n-r)}} \right\} \quad (3.7)$$

$$t_{1Pe} = \bar{y}_{(n-r)} \exp \left\{ \frac{\bar{x}_{(n-r)} - \bar{X}}{\bar{x}_{(n-r)} + \bar{X}} \right\} \quad (3.8)$$

The generalized version of  $t_{1Re}$  and  $t_{1Pe}$  is given by ,

$$t_{1(a_1)} = \bar{y}_{(n-r)} \exp \left\{ \frac{a_1 (\bar{x}_{(n-r)} - \bar{X})}{(\bar{x}_{(n-r)} + \bar{X})} \right\} \quad (3.9)$$

where ' $a_1$ ' is a constant to be chosen suitably. For  $a_1 = -1$ ,  $t_{1(a_1)}$  reduces to the modified ratio estimator  $t_{1Re}$  while for  $a_1 = 1$ , it reduces to the modified product estimator  $t_{1Pe}$ . If we set  $a_1 = 0$ , the estimator  $t_{1(a_1)}$  bowls down to the usual unbiased estimator  $\bar{y}_{(n-r)}$ .

To obtain the biases and mean squared errors of  $t_{1Re}$ ,  $t_{1Pe}$  and  $t_{1(a_1)}$ , we express

$t_{1Re}$ ,  $t_{1Pe}$  and  $t_{1(a_1)}$  in terms of  $\epsilon$ 's we have,

$$t_{1Re} = \bar{Y} (1 + \epsilon_{o(n-r)}) \exp \left[ -\frac{\epsilon_{l(n-r)}}{2 + \epsilon_{l(n-r)}} \right] \quad (3.10)$$

$$t_{1Pe} = \bar{Y} (1 + \epsilon_{o(n-r)}) \exp \left[ \frac{\epsilon_{l(n-r)}}{2 + \epsilon_{l(n-r)}} \right] \quad (3.11)$$

$$t_{l(a_1)} = \bar{Y} (1 + \varepsilon_{o(n-r)}) \exp\left[\frac{a_1 \varepsilon_{l(n-r)}}{2 + \varepsilon_{l(n-r)}}\right] \quad (3.12)$$

Expanding the right hand sides of (3.10), 3.11) and (3.12) in terms of  $\varepsilon$ 's and neglecting terms of  $\varepsilon$ 's having power greater than two we have,

$$t_{lRe} \cong \bar{Y} \left[ 1 + \varepsilon_{o(n-r)} - \frac{1}{2} \varepsilon_{l(n-r)} + \frac{1}{8} (3 \varepsilon_{l(n-r)}^2 - 4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)}) \right]$$

or,

$$(t_{lRe} - \bar{Y}) = \bar{Y} \left[ \varepsilon_{o(n-r)} - \frac{1}{2} \varepsilon_{l(n-r)} + \frac{1}{8} (3 \varepsilon_{l(n-r)}^2 - 4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)}) \right] \quad (3.13)$$

$$t_{lPe} \cong \bar{Y} \left[ 1 + \varepsilon_{o(n-r)} + \frac{1}{2} \varepsilon_{l(n-r)} + \frac{1}{8} (4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)} - \varepsilon_{l(n-r)}^2) \right]$$

or,

$$(t_{lPe} - \bar{Y}) = \bar{Y} \left[ \varepsilon_{o(n-r)} + \frac{1}{2} \varepsilon_{l(n-r)} + \frac{1}{8} (4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)} - \varepsilon_{l(n-r)}^2) \right] \quad (3.14)$$

$$t_{l(a_1)} \cong \bar{Y} \left[ 1 + \varepsilon_{o(n-r)} + \frac{a_1}{2} \varepsilon_{l(n-r)} + \frac{a_1(a_1-2)}{8} \varepsilon_{l(n-r)}^2 + \frac{a_1}{2} \varepsilon_{o(n-r)} \varepsilon_{l(n-r)} \right]$$

or,

$$(t_{l(a_1)} - \bar{Y}) = \bar{Y} \left[ \varepsilon_{o(n-r)} + \frac{a_1}{2} \varepsilon_{l(n-r)} + \frac{a_1}{8} \{ (a_1 - 2) \varepsilon_{l(n-r)}^2 + 4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)} \} \right] \quad (3.15)$$

Taking expectation of both sides of (3.13), (3.14) and (3.15) we get the biases of  $t_{lRe}$ ,  $t_{lPe}$  and  $t_{l(a_1)}$  to the first degree of approximation, respectively as:

$$B(t_{lRe}) = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] \left( \frac{S_x^2}{8\bar{X}} \right) (3R - 4\beta) \quad (3.16)$$

$$B(t_{lPe}) = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] \left( \frac{S_x^2}{8\bar{X}} \right) (4\beta - R) \quad (3.17)$$

$$B\{t_{l(a_1)}\} = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] \left( \frac{a_1 S_x^2}{8\bar{X}} \right) \{ (a_1 - 2)R + 4\beta \} \quad (3.18)$$

It is to be noted that for  $a_1 = -1$  and  $a_1 = 1$  in (3.18) we get the biases of  $t_{lRe}$  and  $t_{lPe}$  respectively given by (3.16) and (3.17). It is also interesting to note that if we set  $a_1 = -2$  and  $a_1 = 2$  in (3.18) we get the biases of the classical ratio ( $t_{lR}$ ) and product ( $t_{lP}$ ) estimators given by (3.3) and (3.4) respectively.

Squaring both sides of (3.13), (3.14) and (3.15) and neglecting terms of  $\varepsilon$ 's having power greater than two we have,

$$(t_{lRe} - \bar{Y})^2 = \bar{Y}^2 \left[ \varepsilon_{o(n-r)}^2 + \frac{1}{4} \{ \varepsilon_{l(n-r)}^2 - 4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)} \} \right] \quad (3.19)$$

$$(t_{lPe} - \bar{Y})^2 = \bar{Y}^2 \left[ \varepsilon_{o(n-r)}^2 + \frac{1}{4} \{ \varepsilon_{l(n-r)}^2 + 4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)} \} \right] \quad (3.20)$$

$$\{t_{l(a_1)} - \bar{Y}\}^2 = \bar{Y}^2 [\varepsilon_{o(n-r)}^2 + (\frac{a_1}{4}) \{a_1 \varepsilon_{l(n-r)}^2 + 4 \varepsilon_{o(n-r)} \varepsilon_{l(n-r)}\}] \quad (3.21)$$

Taking expectation of both sides of (3.19), (3.20) and (3.21) we get the variances of  $t_{lRe}$ ,  $t_{lPe}$  and  $t_{l(a_1)}$  to the first degree of approximation respectively as

$$\text{Var}(t_{lRe}) = [\frac{1}{(nq+2p)} - \frac{1}{N}] [S_y^2 + (\frac{RS_x^2}{4})(R-4\beta)] \quad (3.22)$$

$$\text{Var}(t_{lPe}) = [\frac{1}{(nq+2p)} - \frac{1}{N}] [S_y^2 + (\frac{RS_x^2}{4})(R+4\beta)] \quad (3.23)$$

$$\text{Var}\{t_{l(a_1)}\} = [\frac{1}{(nq+2p)} - \frac{1}{N}] [S_y^2 + (\frac{a_1 RS_x^2}{4})(a_1 R + 4\beta)] \quad (3.24)$$

$$\text{Var}\{\bar{y}_{(n-r)}\} = [\frac{1}{(nq+2p)} - \frac{1}{N}] S_y^2 \quad (3.25)$$

where  $\text{Var}(t_{lR})$ ,  $\text{Var}(t_{lP})$ ,  $\text{Var}(t_{lRe})$ ,  $\text{Var}(t_{lPe})$  are respectively given by (3.5), (3.6), (3.22), and (3.23).

### 3.2 Choice of the Optimum Value of the Constant 'a<sub>1</sub>' and Thus the Optimum Estimator

The variance of the proposed class of estimators  $t_{l(a_1)}$  is minimized for

$$a_1 = -\frac{2\beta}{R} \quad (3.26)$$

Thus the resulting minimum variance of  $t_{l(a_1)}$  is given by

$$\min. \text{Var}\{t_{l(a_1)}\} = [\frac{1}{(nq+2p)} - \frac{1}{N}] S_y^2 (1 - \rho^2) \quad (3.27)$$

If  $p=0$ , then the variance in (3.27) reduces to the variance of the usual regression estimator. In this situation, we can estimate 'a<sub>1</sub>' by

$$\hat{a}_1 = -2 \left\{ \frac{\hat{\beta}_{(n-r)}}{\hat{R}_{(n-r)}} \right\}, \quad (3.28)$$

where  $\hat{\beta}_{(n-r)} = \frac{S_{xy(n-r)}}{S_{x(n-r)}^2}$ , and  $\hat{R}_{(n-r)} = \frac{\bar{y}_{(n-r)}}{\bar{x}_{(n-r)}}$ . Thus we have the following theorems:

**Theorem 3.1** :- The estimator based on estimated optimum value

$\hat{a}_1 = -2 \left\{ \frac{\hat{\beta}_{(n-r)}}{\hat{R}_{(n-r)}} \right\}$ , if random non-response exists on both the study and auxiliary

variables, is given by

$$t_{l(\hat{a}_1)} = \bar{y}_{(n-r)} \exp \left[ \frac{2\hat{\beta}_{(n-r)} \{\bar{X} - \bar{x}_{(n-r)}\}}{\hat{R}_{(n-r)} \{\bar{X} + \bar{x}_{(n-r)}\}} \right] \quad (3.29)$$

It can be easily shown to the first degree of approximation that

$$\text{Var}\{t_{l(a_1)}\} = \left[ \frac{1}{(nq + 2p)} - \frac{1}{N} \right] S_y^2 (1 - \rho^2) = \min. \text{Var}\{t_{l(a_1)}\} \quad (3.30)$$

### 3.3 Strategy – II:

We are considering the situation when information on the study variable  $y$  could not be obtained for  $r$  units while information on the auxiliary variable  $x$  is available and the population mean  $\bar{X}$  of the auxiliary variable  $x$  is known. For example in an industrial survey, the enumerator may not be able to obtain data on the output of some factories while data on the number of workers or the size of machinery may be obtained readily from the records, see Tracy and Osahan (1994).

In such circumstances, the usual ratio and product estimators are defined by,

$$t_{2R} = \bar{y}_{(n-r)} \left( \frac{\bar{X}}{\bar{x}} \right) \quad (3.31)$$

$$t_{2P} = \bar{y}_{(n-r)} \left( \frac{\bar{x}}{\bar{X}} \right) \quad (3.32)$$

To the first degree of approximation, the biases and variance of the estimators  $t_{2R}$  and  $t_{2P}$  are respectively given by,

$$B(t_{2R}) = \frac{1}{\bar{X}} \left( \frac{1}{n} - \frac{1}{N} \right) S_x^2 (R - \beta) \quad (3.33)$$

$$B(t_{2P}) = \left( \frac{1}{n} - \frac{1}{N} \right) \frac{S_x^2}{\bar{X}} \beta \quad (3.34)$$

$$\text{Var}(t_{2R}) = \left( \frac{1}{n} - \frac{1}{N} \right) [S_y^2 + RS_x^2 (R - 2\beta)] + \left[ \frac{1}{(nq + 2p)} - \frac{1}{n} \right] S_y^2 \quad (3.35)$$

$$\text{Var}(t_{2P}) = \left( \frac{1}{n} - \frac{1}{N} \right) [S_y^2 + RS_x^2 (R + 2\beta)] + \left[ \frac{1}{(nq + 2p)} - \frac{1}{n} \right] S_y^2 \quad (3.36)$$

We suggest the modified ratio and product estimators for population mean  $\bar{Y}$  respectively as,

$$t_{2Re} = \bar{y}_{(n-r)} \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \quad (3.37)$$

$$t_{2Pe} = \bar{y}_{(n-r)} \exp \left( \frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}} \right) \quad (3.38)$$

and the generalized version of  $t_{2Re}$  and  $t_{2Pe}$  is given by,

$$t_{2(a_2)} = \bar{y}_{(n-r)} \exp \left\{ \frac{a_2 (\bar{x} - \bar{X})}{(\bar{x} + \bar{X})} \right\} \quad (3.39)$$

To the first degree of approximation, the biases and variances of the estimators  $t_{2Re}$ ,  $t_{2Pe}$  and  $t_{2(a_2)}$  are respectively given by,

$$B(t_{2Re}) = \left( \frac{1}{n} - \frac{1}{N} \right) \left( \frac{S_x^2}{8\bar{X}} \right) (3R - 4\beta) \quad (3.40)$$

$$B(t_{2Pe}) = \left(\frac{1}{n} - \frac{1}{N}\right) \left(\frac{S_x^2}{8\bar{X}}\right) (4\beta - R) \quad (3.41)$$

$$B\{t_{2(a_2)}\} = \left(\frac{1}{n} - \frac{1}{N}\right) \left(\frac{a_2 S_x^2}{8\bar{X}}\right) (a_2 - 2)R + 4\beta \quad (3.42)$$

$$\text{Var}(t_{2Re}) = \left(\frac{1}{n} - \frac{1}{N}\right) \left[S_y^2 + \left(\frac{RS_x^2}{4}\right) (R - 4\beta)\right] + \left[\frac{1}{(nq + 2p)} - \frac{1}{n}\right] S_y^2 \quad (3.43)$$

$$\text{Var}(t_{2Pe}) = \left(\frac{1}{n} - \frac{1}{N}\right) \left[S_y^2 + \left(\frac{RS_x^2}{4}\right) (R + 4\beta)\right] + \left[\frac{1}{(nq + 2p)} - \frac{1}{n}\right] S_y^2 \quad (3.44)$$

$$\text{Var}\{t_{2(a_2)}\} = \left(\frac{1}{n} - \frac{1}{N}\right) \left[S_y^2 + \left(\frac{a_2 RS_x^2}{4}\right) (a_2 R + 4\beta)\right] + \left[\frac{1}{(nq + 2p)} - \frac{1}{n}\right] S_y^2 \quad (3.45)$$

The variance of the estimator  $t_{2(a_2)}$  at (3.45) is minimized for

$$a_2 = -\frac{2\beta}{R} \quad (3.46)$$

Substitution of (3.46) in (3.45) yields the minimum variance of the estimator

$t_{2(a_2)}$  as

$$\text{min. Var}\{t_{2(a_2)}\} = \left[\frac{1}{(nq + 2p)} - \frac{1}{n}\right] S_y^2 + \left(\frac{1}{n} - \frac{1}{N}\right) S_y^2 (1 - \rho^2) \quad (3.47)$$

If  $p = 0$  then the variance in (3.47) reduces to the variance of the usual

regression estimator. We can estimate  $a_2$  by putting  $\hat{a}_2 = \frac{-2\hat{\beta}_{(n-r)}^*}{\hat{R}_{(n-r)}^*}$ , where

$$\hat{\beta}_{(n-r)}^* = \frac{S_{yx(n-r)}}{S_x^2} \text{ and } \hat{R}_{(n-r)}^* = \frac{\bar{y}_{(n-r)}}{\bar{x}}, \text{ in (3.39) we get an estimator based on}$$

estimated optimum value of  $a_2$  for population mean  $\bar{Y}$  as

$$t_{2(\hat{a}_2)} = \bar{y}_{(n-r)} \exp\left\{\frac{2\hat{\beta}_{(n-r)}^* (\bar{X} - \bar{x})}{\hat{R}_{(n-r)}^* (\bar{X} + \bar{x})}\right\}. \quad (3.48)$$

It can be shown to the first degree of approximation that

$$\text{Var}\{t_{2(\hat{a}_2)}\} = \text{min. Var}\{t_{2(a_2)}\}.$$

### 3.4 Strategy – III

Here we again consider the situation when information on the study variate  $y$  could not be obtained for  $r$  units while information on the auxiliary variable  $x$  is obtained for all the sample units. But the difference is that the population mean  $\bar{X}$  of the auxiliary variable  $x$  is not known. In this situation the classical ratio and product estimators of the population mean  $\bar{Y}$  are respectively defined by

$$t_{3R} = \bar{y}_{(n-r)} \left\{ \frac{\bar{X}}{\bar{x}_{(n-r)}} \right\} \quad (3.49)$$

$$t_{3P} = \bar{y}_{(n-r)} \left\{ \frac{\bar{x}_{(n-r)}}{\bar{x}} \right\} \quad (3.50)$$



To the first degree of approximation the biases and variances of the estimators  $t_{3R}$  and  $t_{3P}$  are respectively given by

$$B(t_{3R}) = \left(\frac{S_x^2}{\bar{X}}\right) \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} (R - \beta) \quad (3.51)$$

$$B(t_{3P}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} \frac{\beta S_x^2}{\bar{X}} \quad (3.52)$$

$$\text{Var}(t_{3R}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} [S_y^2 + R S_x^2 (R - 2\beta)] + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.53)$$

$$\text{Var}(t_{3P}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} [S_y^2 + R S_x^2 (R + 2\beta)] + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.54)$$

We further define the modified ratio and product estimators for population mean  $\bar{Y}$  as

$$t_{3Re} = \bar{y}_{(n-r)} \exp\left(\frac{\bar{X} - \bar{X}_{(n-r)}}{\bar{X} + \bar{X}_{(n-r)}}\right) \quad (3.55)$$

$$t_{3Pe} = \bar{y}_{(n-r)} \exp\left(\frac{\bar{X}_{(n-r)} - \bar{X}}{\bar{X}_{(n-r)} + \bar{X}}\right) \quad (3.56)$$

and the generalized version of  $t_{3Re}$  and  $t_{3Pe}$  is given by

$$t_{3(a_3)} = \bar{y}_{(n-r)} \exp\left[\frac{a_3 (\bar{X}_{(n-r)} - \bar{X})}{(\bar{X}_{(n-r)} + \bar{X})}\right], \quad (3.57)$$

where  $a_3$  is a constant to be chosen suitably. To the first degree of approximation, the biases and variances of the estimators  $t_{3Re}$ ,  $t_{3Pe}$  and  $t_{3(a_3)}$  are respectively given by

$$B(t_{3Re}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} \left(\frac{S_x^2}{8\bar{X}}\right) (3R - 4\beta) \quad (3.58)$$

$$B(t_{3Pe}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} \left(\frac{S_x^2}{8\bar{X}}\right) (4\beta - R) \quad (3.59)$$

$$B\{t_{3(a_3)}\} = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} \left(\frac{a_3 S_x^2}{8\bar{X}}\right) \{R(a_3 - 2) + 4\beta\} \quad (3.60)$$

$$\text{Var}(t_{3Re}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} [S_y^2 + \left(\frac{R S_x^2}{4}\right) (R - 4\beta)] + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.61)$$

$$\text{Var}(t_{3Pe}) = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} [S_y^2 + \left(\frac{R S_x^2}{4}\right) (R + 4\beta)] + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.62)$$

$$\text{Var}\{t_{3(a_3)}\} = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} [S_y^2 + \left(\frac{a_3 R S_x^2}{4}\right) (a_3 R + 4\beta)] + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.63)$$

the variance of  $t_{3(a_3)}$  is minimized for

$$a_3 = -\frac{2\beta}{R} \quad (3.64)$$

Thus the resulting minimum variance of  $t_{3(a_3)}$  is given by

$$\text{min. Var}\{t_{3(a_3)}\} = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} S_y^2 (1 - \rho^2) + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.65)$$

The estimate of  $a_3$  based on the sample data at hand is given by

$$\hat{a}_3 = \hat{a}_2 = \frac{-2\hat{\beta}_{(n-r)}^*}{\hat{R}_{(n-r)}^*} \quad (3.66)$$

Thus the resulting estimator based on estimated optimum value  $\hat{a}_3$  is given by

$$t_{3(\hat{a}_3)} = \bar{y}_{(n-r)} \exp\left[\frac{2\hat{\beta}_{(n-r)}^* \{\bar{x} - \bar{x}_{(n-r)}\}}{\hat{R}_{(n-r)}^* \{\bar{x} + \bar{x}_{(n-r)}\}}\right] \quad (3.67)$$

It can be shown to the first degree of approximation that

$$\text{Var} \{t_{3(\hat{a}_3)}\} = \min. \text{Var} \{t_{3(a_3)}\} = \left\{ \frac{1}{(nq+2p)} - \frac{1}{n} \right\} S_y^2 (1-\rho^2) + \left( \frac{1}{n} - \frac{1}{N} \right) S_y^2 \quad (3.68)$$

### 3.5 Efficiency Comparison

From (3.5){(3.35),(3.53)}, (3.6){(3.36),(3.54)}, (3.22){(3.43),(3.61)}, (3.23){(3.44),(3.62)} (3.24){(3.45), (3.63)} and (3.25) we have

$$(i) \quad \text{Var} (t_{jRe}) < \text{Var} \{\bar{y}_{(n-r)}\}, \quad (j = 1, 2, 3) \quad \text{if} \quad \frac{\beta}{R} > \frac{1}{4} \quad (3.69)$$

$$(ii) \quad \text{Var} (t_{jRe}) < \text{Var} (t_{jR}), \quad (j = 1, 2, 3) \quad \text{if} \quad \frac{\beta}{R} < \frac{3}{4} \quad (3.70)$$

$$(iii) \quad \text{Var} (t_{jPe}) < \text{Var} \{\bar{y}_{(n-r)}\}, \quad (j = 1, 2, 3) \quad \text{if} \quad \frac{\beta}{R} < -\frac{1}{4} \quad (3.71)$$

$$(iv) \quad \text{Var} (t_{jPe}) < \text{Var} (t_{jP}), \quad (j = 1, 2, 3) \quad \text{if} \quad \frac{\beta}{R} > -\frac{3}{4} \quad (3.72)$$

$$(v) \quad \text{Var} \{t_{j(a_j)}\} < \text{Var} \{\bar{y}_{(n-r)}\}, \quad (j = 1, 2, 3) \quad \text{if} \quad \min. \{0, (-\frac{4\beta}{R})\} < a_1 < \max. \{0, (-\frac{4\beta}{R})\} \quad (3.73)$$

$$(vi) \quad \text{Var} \{t_{j(a_1)}\} < \text{Var} (t_{jR}), \quad (j = 1, 2, 3) \quad \text{if} \quad \min. \{-2, 2(1 - \frac{2\beta}{R})\} < a_1 < \max. \{-2, 2(1 - \frac{2\beta}{R})\} \quad (3.74)$$

$$(vii) \quad \text{Var} \{t_{j(a_j)}\} < \text{Var} (t_{jRe}), \quad (j = 1, 2, 3) \quad \text{if} \quad \min. \{-1, (1 - \frac{4\beta}{R})\} < a_1 < \max. \{-1, (1 - \frac{4\beta}{R})\} \quad (3.75)$$

$$\begin{aligned} \text{(viii)} \quad \text{Var} \{t_{j(a_j)}\} &< \text{Var} (t_{jp}), \quad (j = 1, 2, 3) \quad \text{if} \\ \min. \{2, -2(1 + \frac{2\beta}{R})\} &< a_1 < \max. \{2, -2(1 + \frac{2\beta}{R})\} \end{aligned} \quad (3.76)$$

$$\begin{aligned} \text{(ix)} \quad \text{Var} \{t_{j(a_j)}\} &< \text{Var} (t_{jpe}), \quad (j = 1, 2, 3) \quad \text{if} \\ \min. \{1, -(1 + \frac{4\beta}{R})\} &< a_1 < \max. \{1, -(1 + \frac{4\beta}{R})\} \end{aligned} \quad (3.77)$$

Combining (3.69) and (3.70) we find that the proposed estimator  $t_{jRe}$  ( $j = 1, 2, 3$ ) is better than usual unbiased estimator  $\bar{y}_{(n-r)}$  and ratio estimator  $t_{jR}$  ( $j = 1, 2, 3$ ) if

$$\frac{1}{4} < \frac{\beta}{R} < \frac{3}{4} \quad (3.78)$$

Further combining (3.71) and (3.72) we obtain that the proposed estimator  $t_{jpe}$  ( $j = 1, 2, 3$ ) is more efficient than usual unbiased estimator  $\bar{y}_{(n-r)}$  and the product estimator  $t_{jp}$  ( $j = 1, 2, 3$ ) if

$$-\frac{3}{4} < \frac{\beta}{R} < -\frac{1}{4} \quad (3.79)$$

#### 4. Estimation from the Sample of the Percent Relative Efficiency

In this section we have obtained the percent relative efficiencies (PRE) of the proposed estimators and their estimates from the sample in three different situations, see Sukhatme et al (1970, pp. 231-232).

##### 4.1 Strategy -I:

Percent relative efficiencies (PREs) of  $t_{1R}$ ,  $t_{1Re}$  and  $t_{1(\hat{a}_1)}$  with respect to  $\bar{y}_{(n-r)}$  are respectively given by

$$\text{PRE}(t_{1R}, \bar{y}_{(n-r)}) = [1 + A(1 - 2k)]^{-1} * 100 \quad (4.1)$$

$$\text{PRE}(t_{1Re}, \bar{y}_{(n-r)}) = [1 + (A/4)(1 - 4k)]^{-1} * 100 \quad (4.2)$$

$$\text{PRE}(t_{1(\hat{a}_1)}, \bar{y}_{(n-r)}) = (1 - \rho^2) * 100, \quad (4.3)$$

where  $A = (R^2 S_x^2 / S_y^2)$  and  $k = (\beta / R)$ .

Further the percent relative efficiencies of  $t_{1(a_1)}$  with respect to  $\bar{y}_{(n-r)}$ ,  $t_{1R}$  and  $t_{1Re}$  are respectively given by

$$\text{PRE}(t_{1(a_1)}, \bar{y}_{(n-r)}) = [1 + a_1(A/4)(a_1 + 4k)]^{-1} * 100 \quad (4.4)$$

$$\text{PRE}(t_{1(a_1)}, t_{1R}) = \frac{[1 + A(1 - 2k)]}{[1 + a_1(A/4)(a_1 + 4k)]} * 100 \quad (4.5)$$

$$\text{PRE}(t_{1(a_1)}, t_{1Re}) = \frac{[1 + (A/4)(1 - 4k)]}{[1 + a_1(A/4)(a_1 + 4k)]} * 100 \quad (4.6)$$

It is observed from (4.1) to (4.6) that the percent relative efficiencies are the functions of unknown population parameters such as  $R$ ,  $S_x^2$ ,  $S_y^2$ ,  $\beta$  and  $\rho$ . In practice, population parameters are not known exactly and hence these percent relative efficiencies formulae can not be used. To overcome this difficulty, it is advisable to estimate the PREs, replacing the population parameters involved in these PREs by their consistent estimates based on the sample. Thus the estimates of these PREs are respectively given by

$$\hat{PRE}(t_{1R}, \bar{y}_{(n-r)}) = [1 + \hat{A}(1 - 2\hat{k})]^{-1} * 100 \quad (4.7)$$

$$\hat{PRE}(t_{1Re}, \bar{y}_{(n-r)}) = [1 + (\hat{A}/4)(1 - 4\hat{k})]^{-1} * 100 \quad (4.8)$$

$$\hat{PRE}(t_{1(\hat{a}_1)}, \bar{y}_{(n-r)}) = (1 - \hat{\rho}_{(n-r)}^2) * 100 \quad (4.9)$$

$$\hat{PRE}(t_{1(a_1)}, \bar{y}_{(n-r)}) = [1 + a_1(\hat{A}/4)(a_1 + 4\hat{k})]^{-1} * 100 \quad (4.10)$$

$$\hat{PRE}(t_{1(a_1)}, t_{1R}) = \frac{[1 + \hat{A}(1 - 2\hat{k})]}{[1 + a_1(\hat{A}/4)(a_1 + 4\hat{k})]} * 100 \quad (4.11)$$

$$\hat{PRE}(t_{1(a_1)}, t_{1Re}) = \frac{[1 + (\hat{A}/4)(1 - 4\hat{k})]}{[1 + a_1(\hat{A}/4)(a_1 + 4\hat{k})]} * 100, \quad (4.12)$$

where  $\hat{A}_{(n-r)} = \hat{R}_{(n-r)}^2 s_{x(n-r)}^2 / s_{y(n-r)}^2$ ,  $\hat{k} = \hat{\beta}_{(n-r)} / \hat{R}_{(n-r)}$ ,  $\hat{R}_{(n-r)} = \bar{y}_{(n-r)} / \bar{x}_{(n-r)}$ ,

$\hat{\beta}_{(n-r)} = s_{xy(n-r)} / s_{x(n-r)}^2$  and  $\hat{\rho}_{(n-r)} = s_{xy(n-r)} / \{s_{x(n-r)} s_{y(n-r)}\}$ .

## 4.2 Strategy-II:

Percent relative efficiencies of  $t_{2R}$ ,  $t_{2Re}$  and  $t_{2(\hat{a}_2)}$  with respect to  $\bar{y}_{(n-r)}$  are respectively given by

$$PRE(t_{2R}, \bar{y}_{(n-r)}) = [1 + FA(1 - 2k)]^{-1} * 100 \quad (4.13)$$

$$PRE(t_{2Re}, \bar{y}_{(n-r)}) = [1 + F(A/4)(1 - 4k)]^{-1} * 100 \quad (4.14)$$

$$PRE(t_{2(\hat{a}_2)}, \bar{y}_{(n-r)}) = (1 - F\rho^2) * 100, \quad (4.15)$$

where  $\theta = (nq + 2p)$  and  $F = \{\theta(N - n)\} / \{n(N - \theta)\}$ .

Further the percent relative efficiencies of  $t_{2(a_2)}$  with respect to  $\bar{y}_{(n-r)}$ ,  $t_{2R}$  and  $t_{2Re}$  are respectively given by

$$PRE(t_{2(a_2)}, \bar{y}_{(n-r)}) = [1 + F(A/4)a_2(a_2 + 4k)]^{-1} * 100 \quad (4.16)$$

$$PRE(t_{2(a_2)}, t_{2R}) = \frac{[1 + FA(1 - 2k)]}{[1 + F(A/4)a_2(a_2 + 4k)]} * 100 \quad (4.17)$$

$$PRE(t_{2(a_2)}, t_{1Re}) = \frac{[1 + F(A/4)(1 - 4k)]}{[1 + F(A/4)a_2(a_2 + 4k)]} * 100 \quad (4.18)$$

The estimates of PREs given at (4.13)-(4.18) based on sample are respectively given by

$$\hat{PRE}(t_{2R}, \bar{y}_{(n-r)}) = [1 + \hat{F}\hat{A}^*(1 - 2\hat{k}^*)]^{-1} * 100 \quad (4.19)$$

$$\hat{PRE}(t_{2Re}, \bar{y}_{(n-r)}) = [1 + \hat{F}(\hat{A}^*/4)(1 - 4\hat{k}^*)]^{-1} * 100 \quad (4.20)$$

$$\hat{PRE}(t_{2(\hat{a}_2)}, \bar{y}_{(n-r)}) = (1 - \hat{F}\hat{\rho}_{(n-r)}^2) * 100 \quad (4.21)$$

$$\hat{PRE}(t_{2(a_2)}, \bar{y}_{(n-r)}) = [1 + \hat{F}(\hat{A}^*/4)a_2(a_2 + 4\hat{k}^*)]^{-1} * 100 \quad (4.22)$$

$$\hat{PRE}(t_{2(a_2)}, t_{2R}) = \frac{[1 + \hat{F}\hat{A}^*(1 - 2\hat{k}^*)]}{[1 + \hat{F}(\hat{A}^*/4)a_2(a_2 + 4\hat{k}^*)]} * 100 \quad (4.23)$$

$$\hat{PRE}(t_{2(a_2)}, t_{2Re}) = \frac{[1 + \hat{F}(\hat{A}^*/4)(1 - 4\hat{k}^*)]}{[1 + \hat{F}(\hat{A}^*/4)a_2(a_2 + 4\hat{k}^*)]} * 100, \quad (4.24)$$

where  $\hat{F} = \{\hat{\theta}(N - n)\} / \{n(N - \hat{\theta})\}$ ,  $\hat{\theta} = (n\hat{q} + 2\hat{p})$ ,  $\hat{A}^* = \hat{R}_{(n-r)}^{*2} s_{x(n-r)}^2 / s_{y(n-r)}^2$ ,

$\hat{k}^* = \hat{\beta}_{(n-r)}^* / \hat{R}_{(n-r)}^*$ ,  $\hat{\beta}_{(n-r)}^* = s_{yx(n-r)} / s_x^2$ ,  $\hat{R}_{(n-r)}^* = \bar{y}_{(n-r)} / \bar{x}$  and  $\hat{\rho}_{(n-r)}^* = s_{yx(n-r)} / \{s_x s_{y(n-r)}\}$ .

However one may also take the estimate of the population correlation coefficient  $\rho$  as  $\hat{\rho}_{(n-r)} = s_{xy(n-r)} / \{s_{x(n-r)} s_{y(n-r)}\}$  in place of  $\hat{\rho}_{(n-r)}^*$ , see Singh et al (2000).

### 4.3 Strategy-III:

Percent relative efficiencies of  $t_{3R}$ ,  $t_{3Re}$  and  $t_{3(\hat{a}_3)}$  with respect to  $\bar{y}_{(n-r)}$  are respectively given by

$$PRE(t_{3R}, \bar{y}_{(n-r)}) = [1 + F^*A(1 - 2k)]^{-1} * 100 \quad (4.25)$$

$$PRE(t_{3Re}, \bar{y}_{(n-r)}) = [1 + F^*(A/4)(1 - 4k)]^{-1} * 100 \quad (4.26)$$

$$PRE(t_{3(\hat{a}_3)}, \bar{y}_{(n-r)}) = (1 - F^*\rho^2) * 100, \quad (4.27)$$

where  $F^* = \{N(n - \theta)\} / \{n(N - \theta)\}$ .

Further the percent relative efficiencies of  $t_{3(a_3)}$  with respect to  $\bar{y}_{(n-r)}$ ,  $t_{3R}$  and  $t_{3Re}$  are respectively given by

$$PRE(t_{3(a_3)}, \bar{y}_{(n-r)}) = [1 + F^*(A/4)a_3(a_3 + 4k)]^{-1} * 100 \quad (4.28)$$

$$PRE(t_{3(a_3)}, t_{3R}) = \frac{[1 + F^*A(1 - 2k)]}{[1 + F^*(A/4)a_2(a_2 + 4k)]} * 100 \quad (4.29)$$

$$PRE(t_{3(a_3)}, t_{1Re}) = \frac{[1 + F^*(A/4)(1 - 4k)]}{[1 + F^*(A/4)a_3(a_3 + 4k)]} * 100. \quad (4.30)$$

The estimates of PREs given at (4.25)-(4.30) based on sample are respectively given by

$$\hat{PRE}(t_{3R}, \bar{y}_{(n-r)}) = [1 + \hat{F}^* \hat{A}^* (1 - 2\hat{k}^*)]^{-1} * 100 \quad (4.31)$$

$$\hat{PRE}(t_{3Re}, \bar{y}_{(n-r)}) = [1 + \hat{F}^* (\hat{A}^* / 4)(1 - 4\hat{k}^*)]^{-1} * 100 \quad (4.32)$$

$$\hat{PRE}(t_{3(\hat{a}_3)}, \bar{y}_{(n-r)}) = (1 - \hat{F}^* \hat{\rho}_{(n-r)}^{*2}) * 100 \quad (4.33)$$

$$\hat{PRE}(t_{3(a_3)}, \bar{y}_{(n-r)}) = [1 + \hat{F}^* (\hat{A}^* / 4)a_3(a_3 + 4\hat{k}^*)]^{-1} * 100 \quad (4.34)$$

$$\hat{PRE}(t_{3(a_3)}, t_{3R}) = \frac{[1 + \hat{F}^* \hat{A}^* (1 - 2\hat{k}^*)]}{[1 + \hat{F}^* (\hat{A}^* / 4)a_3(a_3 + 4\hat{k}^*)]} * 100 \quad (4.35)$$

$$\hat{PRE}(t_{3(a_3)}, t_{3Re}) = \frac{[1 + \hat{F}^* (\hat{A}^* / 4)(1 - 4\hat{k}^*)]}{[1 + \hat{F}^* (\hat{A}^* / 4)a_3(a_3 + 4\hat{k}^*)]} * 100, \quad (4.36)$$

where  $\hat{F}^* = \{N(n - \hat{\theta})\} / \{n(N - \hat{\theta})\}$ .

• **Remark 4.1-** In similar fashion, the PREs of different product-type estimators {i.e.  $PRE(t_{jp}, \bar{y}_{(n-r)})$ ,  $PRE(t_{jPe}, \bar{y}_{(n-r)})$ ,  $PRE(t_{j(\hat{a}_j)}, \bar{y}_{(n-r)})$ ,  $PRE(t_{j(a_j)}, t_{jP})$ ,  $PRE(t_{j(a_j)}, t_{jPe})$ , ( $j = 1, 2, 3$ )} and their estimates {i.e.  $\hat{PRE}(t_{jp}, \bar{y}_{(n-r)})$ ,  $\hat{PRE}(t_{jPe}, \bar{y}_{(n-r)})$ ,  $\hat{PRE}(t_{j(\hat{a}_j)}, \bar{y}_{(n-r)})$ ,  $\hat{PRE}(t_{j(a_j)}, t_{jP})$  and  $\hat{PRE}(t_{j(a_j)}, t_{jPe})$  can be obtained without much difficulty.

## 5. Empirical Study

To illustrate the performance of different estimators of population mean  $\bar{Y}$  we consider the data as given in Singh (2003, p. 990).

The descriptions of the variables are given below:

$y_i$  = Amount (in \$ 000) of real state farm loans in different states during 1997.

$x_i$  = Amount (in \$000) of nonreal state farm loans in different states during 1997.

The estimates of different parameters are given below:

$$\bar{X} = 878.16; N = 50, n = 20, r = 6, \bar{x} = 942.8615, \bar{x}_{(n-r)} = 1068.292, \bar{y}_{(n-r)} = 580.177,$$

$$s_x^2 = 1307911.82, s_{x(n-r)}^2 = 1685838.731, s_{y(n-r)}^2 = 295169.6416,$$

$$s_{xy(n-r)} = 595984.887,$$

$$\hat{\rho}_{(n-r)} = 0.84487, \hat{\rho}_{(n-r)}^* = 0.9392, \hat{p} = 0.34993, \hat{q} = (1 - \hat{p}) = 0.65,$$

$$\hat{\beta}_{(n-r)} = \frac{S_{xy(n-r)}}{S_x^2} = 0.35353,$$

$$\hat{\beta}_{(n-r)}^* = \frac{S_{xy(n-r)}}{S_x^2} = 0.45568, \hat{R}_{(n-r)} = 0.54309, \hat{R}_{(n-r)}^* = 0.61534, s_x = 1143.639725,$$

$$\hat{\theta} = (n\hat{q} + 2\hat{p}) = 13.7, \frac{\hat{\beta}_{(n-r)}}{\hat{R}_{(n-r)}} = 0.6510, \frac{\hat{\beta}_{(n-r)}^*}{\hat{R}_{(n-r)}^*} = 0.7405, s_{x(n-r)} = 1298.398525,$$

$$\{1 - \hat{\rho}_{(n-r)}^2\} = 0.2862, \frac{4\hat{\beta}_{(n-r)}}{\hat{R}_{(n-r)}} = 2.6040, \frac{\hat{R}_{(n-r)}^2 s_{x(n-r)}^2}{s_{y(n-r)}^2} = 1.684565697,$$

$$s_{y(n-r)} = 543.2951699, \hat{A} = \frac{\hat{R}_{(n-r)}^2 s_{x(n-r)}^2}{s_{y(n-r)}^2} = 1.6846, \hat{A}^* = \frac{\hat{R}_{(n-r)}^{*2} s_{x(n-r)}^2}{s_{y(n-r)}^2} = 3.5145,$$

$$\hat{F} = \frac{\hat{\theta}(N - n)}{n(N - \hat{\theta})} = 745.965, \hat{F}^* = \frac{N(n - \hat{\theta})}{n(N - \hat{\theta})} = 571.725, \hat{k} = \frac{\hat{\beta}_{(n-r)}}{\hat{R}_{(n-r)}} = 0.651, \hat{k}^* = \frac{\hat{\beta}_{(n-r)}^*}{\hat{R}_{(n-r)}^*}$$

We have computed the estimates of the PREs of the estimators:

- (a)  $t_{jR}, t_{jRe}$  and  $t_{j(\hat{a}_j)}$ , ( $j = 1, 2, 3$ ) with respect to the usual unbiased estimator  $\bar{y}_{(n-r)}$  using the formulae {(4.7), (4.8), (4.9)}, {(4.13), (4.14), (4.15)} and {(4.19), (4.20), (4.21)} and findings are demonstrated in Table 5.1, 5.6 and 5.11 respectively;
- (b)  $t_{j(a_j)}$  with respect to the usual unbiased estimator  $\bar{y}_{(n-r)}$ ,  $t_{jR}$  and  $t_{jRe}$ , ( $j = 1, 2, 3$ ) using the formulae {(4.10), (4.11), (4.12)}, {(4.16), (4.17), (4.18)} and {(4.22), (4.23), (4.24)} and results are shown in Tables {5.2, 5.3, 5.4}, {5.7, 5.8, 5.9}, and {5.12, 5.13, 5.14} respectively.

we have also computed the estimate of range of scalar  $a_j$  in which the proposed estimator  $t_{j(a_j)}$  is better than  $\bar{y}_{(n-r)}$ ,  $t_{jR}$  and  $t_{jRe}$ , ( $j = 1, 2, 3$ ) and finding are compiled in Tables 5.5, 5.10 and 5.14 respectively.

### Strategy-I:

**Table 5.1:** The estimates of PREs of  $\bar{y}_{(n-r)}, t_{IR}, t_{IRe}$  and  $t_{I(\hat{a}_1)}$

Estimator	$\bar{y}_{(n-r)}$	$t_{IR}$	$t_{IRe}$	$t_{I(\hat{a}_1)}$
PRE $\{., \bar{y}_{(n-r)}\}$	100.00	203.56	308.18	349.55

**Table 5.2: The estimates of PRE of  $t_{l(a_1)}$  with respect to  $\bar{y}_{(n-r)}$  for different values of  $a_1$** 

$a_1$	-2.604	-2.50	-2.25	-2.00	-1.75
$\hat{PRE} \{t_{l(a_1)}, \bar{y}_{(n-r)}\}$	100.00	112.30	150.48	203.56	269.83
$a_1$	-1.50	$a_{l(opt)} = -1.302$	-1.25	-1.00	-0.75
$\hat{PRE} \{t_{l(a_1)}, \bar{y}_{(n-r)}\}$	330.48	349.55	348.16	308.18	241.31
$a_1$	-0.50	-0.25	0.00		
$\hat{PRE} \{t_{l(a_1)}, \bar{y}_{(n-r)}\}$	179.55	132.95	100.00		

**Table 5.3: The estimates of PRE of  $t_{l(a_1)}$  with respect to the ratio estimator  $t_{lR}$  for different values of  $a_1$** 

$a_1$	-2.00	-1.75	-1.50	$a_{l(opt)} = -1.302$	-1.25
$\hat{PRE} \{t_{l(a_1)}, t_{lR}\}$	100.00	132.56	162.35	171.72	171.04
$a_1$	-1.00	-0.75	-0.6020		
$\hat{PRE} \{t_{l(a_1)}, t_{lR}\}$	151.39	118.55	100.00		

**Table 5.4: The estimates of PRE of  $t_{l(a_1)}$  with respect to the modified ratio type estimator  $t_{2Re}$  for different value of  $a_1$** 

$a_1$	-1.604	-1.50	$a_{l(opt)} = -1.3020$	-1.25	-1.00
$\hat{PRE} \{t_{l(a_1)}, t_{2Re}\}$	100.00	107.25	113.43	112.98	100.00

**Table 5.5: The estimates of range of  $a_1$  for  $t_{l(a_1)}$  to be more efficient than different estimators of population mean**

Estimator	$\bar{y}_{(n-r)}$	$t_{lR}$	$t_{lRe}$
Estimate of range of $a_1$	(-2.6040 , 0.0000)	(-2.0000 , -0.6020)	(-1.6040 , -1.0000)



**Strategy – II:**

**Table 5.6:** The estimates of PRE of  $\bar{y}_{(n-r)}$ ,  $t_{2R}$ ,  $t_{2Re}$  and  $t_{2(\hat{a}_2)}$  with respect to  $\bar{y}_{(n-r)}$

Estimator	$\bar{y}_{(n-r)}$	$t_{2R}$	$t_{2Re}$	$t_{2(\hat{a}_2)}$
$\hat{PRE} \{ \cdot, \bar{y}_{(n-r)} \}$	100.00	184.11	187.22	208.69

**Table 5.7:** The estimates of PRE of  $t_{2(a_2)}$  with respect to  $\bar{y}_{(n-r)}$  for different values of  $a_2$

$a_2$	-2.962	-2.50	-2.25	-2.00	-1.75
$\hat{PRE} \{ t_{2(a_2)}, \bar{y}_{(n-r)} \}$	100.00	137.79	161.39	184.11	201.46
$a_2$	-1.50	$\hat{a}_{2(opt)} = -1.4810$	-1.25	-1.00	-0.75
$\hat{PRE} \{ t_{2(a_2)}, \bar{y}_{(n-r)} \}$	208.65	208.69	203.31	187.22	165.00
$a_2$	-0.50	-0.25	0.00		
$\hat{PRE} \{ t_{2(a_2)}, \bar{y}_{(n-r)} \}$	141.30	119.19	100.00		

**Table 5.8:** The estimates of PRE of  $t_{2(a_2)}$  with respect to the ratio estimator  $t_{2R}$  for different values of  $a_2$

$a_2$	-2.00	-1.75	-1.50	$\hat{a}_{2(opt)} = -1.4810$	-1.25
$\hat{PRE} \{ t_{2(a_2)}, t_{2R} \}$	100.00	109.43	113.34	113.36	110.44
$a_2$	-1.00	-0.962			
$\hat{PRE} \{ t_{2(a_2)}, t_{2R} \}$	101.70	100.00			

**Table 5.9:** The estimates of PRE of  $t_{2(a_2)}$  with respect to the modified ratio estimator  $t_{2Re}$

$a_2$	-1.962	-1.50	$\hat{a}_{2(opt)} = -1.4810$	-1.25	-1.00
$\hat{PRE} \{ t_{2(a_2)}, t_{2Re} \}$	100.00	111.44	111.46	108.59	100.00

**Table 5.10: The estimates of range of  $a_2$  for the proposed estimator  $t_{2(a_2)}$  to be more efficient than different estimators of the population mean**

Estimator	$\bar{y}_{(n-r)}$	$t_{2R}$	$t_{2Re}$
The estimates of range of $a_2$	(-2.962 , 0.00)	(-2.000 , -0.962)	(-1.962 , -1.000)

**Strategy – III:**

**Table 5.11: The estimates of PRE of different estimators  $\bar{y}_{(n-r)}, t_{3R}, t_{3Re}$  and  $t_{3(\hat{a}_3)}$  with respect to the usual unbiased estimator  $\bar{y}_{(n-r)}$**

Estimator	$\bar{y}_{(n-r)}$	$t_{3R}$	$t_{3Re}$	$t_{3(\hat{a}_3)}$
$\hat{PRE} \{ (\cdot), \bar{y}_{(n-r)} \}$	100.00	153.88	155.54	166.44

**Table 5.12: The estimates of PRE of  $t_{3(a_3)}$  with respect to the usual unbiased estimator  $\bar{y}_{(n-r)}$  for different values of  $a_3$**

$a_3$	-2.962	-2.50	-2.25	-2.00	-1.75
$\hat{PRE} \{ t_{3(a_3)}, \bar{y}_{(n-r)} \}$	100.00	126.62	141.16	153.89	162.87
$a_3$	-1.50	$\hat{a}_{3(opt)} = -1.4810$	-1.25	-1.00	-0.75
$\hat{PRE} \{ t_{3(a_3)}, \bar{y}_{(n-r)} \}$	166.42	166.44	163.80	155.54	143.25
$a_3$	-0.50	-0.25	0.00		
$\hat{PRE} \{ t_{3(a_3)}, \bar{y}_{(n-r)} \}$	128.87	114.08	100.00		

**Table 5.13: The estimates of PRE of  $t_{3(a_3)}$  with respect to the usual ratio estimator  $t_{3R}$  for different value of  $a_3$**

$a_3$	-2.00	-1.75	-1.50	$\hat{a}_{3(opt)} = -1.4810$	-1.25
$\hat{PRE} \{ t_{3(a_3)}, t_{3R} \}$	100.00	105.84	108.15	108.16	106.45
$a_3$	-1.00	-0.962			
$\hat{PRE} \{ t_{3(a_3)}, t_{3R} \}$	101.08	100.00			

**Table 5.14: The estimates of PRE of  $t_{3(a_3)}$  with respect to modified ratio estimator  $t_{3Re}$  for different values of  $a_3$**

$a_3$	-1.962	-1.50	$\hat{a}_{3(opt)} = -1.4810$	-1.25	-1.00
$\hat{PRE}\{t_{3(a_3)}, t_{3Re}\}$	100.00	107.00	107.01	105.31	100.00

**Table 5.15: The estimates of range of  $a_3$  for the proposed estimator  $t_{3(a_3)}$  to be more efficient than the different estimators of the population mean  $\bar{Y}$**

Estimator	$\bar{y}_{(n-r)}$	$t_{3R}$	$t_{3Re}$
The estimates of range of $a_3$	(-2.962 , 0.000)	(-2.000 , -0.962)	(-1.962 , -1.000)

It is observed from Tables (5.2, 5.4, 5.5) , (5.7, 5.9, 5.10) and (5.12, 5.14, 5.15) that there is enough scope of selecting the scalar  $a_j$ , ( $j = 1,2,3$ ); to obtain estimators better than the usual unbiased estimator  $\bar{y}_{(n-r)}$ , usual ratio estimator  $t_{jR}$ , ( $j = 1,2,3$ ) and modified ratio type estimator  $t_{jRe}$ , ( $j = 1,2,3$ ). Larger estimated gain in efficiency by using the proposed classes of estimators over  $\bar{y}_{(n-r)}$ ,  $t_{jR}$  and  $t_{jRe}$  are observed in the neighborhood of estimated optimum value  $\hat{a}_{j(opt)}$ , ( $j = 1,2,3$ ) of the scalar  $a_j$ , ( $j = 1,2,3$ ). Largest estimated gain in efficiency is seen at estimated optimum value  $\hat{a}_{j(opt)}$  of  $a_j$ , ( $j = 1,2,3$ ). The estimated gain in efficiency by using the estimator  $t_{j(a_j)}$ , ( $j = 1,2,3$ ). over the usual unbiased estimator  $\bar{y}_{(n-r)}$  (which does not utilize the auxiliary information) is larger as well for larger range of the scalar  $a_j$ , ( $j=1,2,3$ ) as compared to the usual ratio estimator  $t_{jR}$  and  $t_{jRe}$ , ( $j = 1,2,3$ ). Tables 5.5, 5.10 and 5.15 exhibit that when  $a_1 \in (-1.604 , -1.000)$ ,  $a_2 \in (-1.962 , -1.000)$  and  $a_3 \in (-1.962 , -1.000)$  the proposed class of estimators  $t_{1(a_1)}$ ,  $t_{2(a_2)}$  and  $t_{3(a_3)}$  are always more efficient than  $\{\bar{y}_{(n-r)}, t_{1R}, t_{1Re}\}$ ,  $\{\bar{y}_{(n-r)}, t_{2R}, t_{2Re}\}$  and  $\{\bar{y}_{(n-r)}, t_{3R}, t_{3Re}\}$  respectively. It is further observed from Tables 5.1, 5.6 and 5.11 that the proposed modified exponential type ratio estimator  $t_{jRe}$ , ( $j=1,2,3$ ) is more efficient than the usual unbiased estimator  $\bar{y}_{(n-r)}$  and the usual ratio estimator  $t_{jR}$ , ( $j = 1,2,3$ ) with substantial estimated gain in efficiency. However the estimators  $t_{1(\hat{a}_1)}$ ,  $t_{2(\hat{a}_2)}$  and  $t_{3(\hat{a}_3)}$  based

on estimated optimum values are more efficient than the estimators  $\{\bar{y}_{(n-r)}, t_{1R}, t_{1Re}\}$ ,  $\{\bar{y}_{(n-r)}, t_{2R}, t_{2Re}\}$  and  $\{\bar{y}_{(n-r)}, t_{3R}, t_{3Re}\}$  respectively with considerable estimated gain in efficiency. Thus the proposed estimators  $t_{jRe}$  and  $t_{j(\hat{a}_j)}$ , ( $j=1,2,3$ ) are recommended for their use in practice. The estimated gain in efficiency due to the estimators proposed in Strategy-I is the largest followed by the estimators in Strategy -II and Strategy -III.

## Acknowledgements

Authors are thankful to the learned referee for his valuable suggestions and pointing out the mistakes on earlier draft of the paper.

## References

1. Ahmed, M.S., Al-Titib, O., Al-Rawib, Z., Abu- Dayyeh, W. (2005): Estimation of finite population variance in presence of random non-response using auxiliary variables. *Information and Management science*, 16(2), 73-82.
2. Chaudhuri, A and Stenger, H. (1992): *Survey Sampling, theory and methods*, Marcel Dekkar, New-York.
3. Chaudhuri, A and Vos, J.W.E. (1998): *Unified theory and strategies of survey sampling*, North-Holland.
4. Cochran, W.G. (1940): The estimation of the yields of cereal experiments by sampling for the ratio gain to total produce. *Jour. Agric. Soc.*, 30, 262-275.
5. Cochran, W.G. (1997): *Sampling techniques*. Third edition, John Willey and Sons, New-York.
6. Heitjan, D.F. and Basu, S. (1966): Distinguishing 'missing at random' and missing completely at Random'. *The American Statistician*, 50, 207-213.
7. Krisnaiah, P.R. and Rao, C.R. (1998): *Hand book of statistics*, Vol.6, North-Holland.
8. Murthy, M.N. (1964): Product method of estimation. *Sankhya*, 26, 69-74.
9. Rubin, D.B. (1976): Inference and missing data. *Biometrika*, 63(3), 581-592.
10. Singh, S. (2003): *Advanced sampling theory with applications*, How Michael 'selected' Any. Vol. I and II Kluwer Academic Publishers.
11. Singh H.P., Chandra, P., Joarder, A. H. and Singh, S. (2007): Family of estimators of mean, ratio and product of a finite population using random non-response. *Test*, 16, 565-597.

12. Singh, H.P. Chandra, P. and Singh, S. (2003): Variance estimation using multiauxiliary information for random non-response in survey sampling. *Statistica*. LXIII, 23-40.
13. Singh, H.P. and Tracy, D.S. (2001): Estimation of population mean in presence of random non-response in sample surveys. *Statistica*, LXI, 231-248.
14. Singh, R. and Mangat, N.S. (1996): Elements of survey sampling. Kluwer Academic Publisher.
15. Singh, S. and Joarder, A.H. (1998): Estimation of finite population variance using random non-response in survey sampling, *Metrika*, 47, 241-249.
16. Singh, S., Joarder, A.H. and Tracy, D.S. (2000): Regression type estimators for random non-response in survey sampling. *Statistica*, LX, (1), 39-43.
17. Sukhatme, P.V. and Sukhatme, B.V. (1970): Sampling theory of surveys with applications. Iowa State University Press, Ames, Iowa.
18. Tracy, D.S, and Osahan, S.S. (1994): Random non-response on study variable versus on study as well as auxiliary variables. *Statistica*, 54, 163-168.
19. Toutenburg, H. and Srivastava, V.K. (1998): Estimation of ratio of population means in survey sampling when some observation are missing. *Metrika*, 48, 177-187.
20. Tuteja, R.K. and Bahl, S. (1991): Ratio and product type exponential estimator. *Information and Optimization Science*, 12, 159-163.