

## On Three Methods of Estimation In Two-Stage Sampling Using Auxiliary Information

L.N. Sahoo and Nirupama Mahapatra<sup>1</sup>

Received: March, 2007; Revised: August, 2007

### ABSTRACT

This paper discusses three different estimation methodologies on gathering auxiliary information available at different stages of a survey under a two-stage sampling set-up. In order to make a comparative study of the discussed methods, we compare performances of three different classes of estimators based on these methods available in the literature.

**Key words :** Asymptotic variance, auxiliary variable, two-stage sampling.

**AMS Subject Classification :** 62 D05

### I. INTRODUCTION

Consider a finite population  $U = \{1, \dots, i, \dots, N\}$  of  $N$  first stage units (*fsu*) such that the  $i$ th *fsu*  $U_i = \{1, \dots, j, \dots, M_i\}$  contains  $M_i$  second stage units (*ssu*) and  $M = \sum_{i=1}^N M_i$ . Let  $y_{ij}$  and  $x_{ij}$  denote the values of the study variable  $y$  and an auxiliary variable  $x$  for the  $j$ th *ssu* of  $U_i$  ( $j = 1, 2, \dots, M_i; i = 1, 2, \dots, N$ ). Let  $\bar{Y}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} y_{ij}$ ,  $\bar{X}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} x_{ij}$ ,  $\bar{Y} = \frac{1}{N} \sum_{i=1}^N \alpha_i \bar{Y}_i$  and  $\bar{X} = \frac{1}{N} \sum_{i=1}^N \alpha_i \bar{X}_i$  where  $\alpha_i = NM_i/M$ . To estimate the population mean  $\bar{Y}$ , let us assume that a sample  $s$  ( $s \subset U$ ) of  $n$  *fsus* is drawn from  $U$  and then a sample  $s_i$  ( $s_i \subset U_i$ ) of  $m_i$  *ssus* from the selected  $U_i$ ,  $i \in s$ , according to the design simple random sampling without replacement (SRSWOR). Define  $\bar{y}_i = \frac{1}{m_i} \sum_{j \in s_i} y_{ij}$ ,  $\bar{x}_i = \frac{1}{m_i} \sum_{j \in s_i} x_{ij}$ ,  $\bar{y} = \frac{1}{n} \sum_{i \in s} \alpha_i \bar{y}_i$ ,  $\bar{x} = \frac{1}{n} \sum_{i \in s} \alpha_i \bar{x}_i$  and  $t_x = \frac{1}{n} \sum_{i \in s} \alpha_i \bar{X}_i$ .

In two-stage sampling, precision of an estimator not only depends on the strength of the relationship between  $y$  and  $x$  but also on how well the kind and extent of the available auxiliary information can be utilized at different stages. Many different cases can be considered according to the nature of the available auxiliary information [see for example, Sarndal, Swensson and Wretman (1992, p.304)]. But, here we are confined to two cases only. In case A, emphasis is given on the availability of overall population mean or total of  $x$ . In case B, it is assumed that the mean or total of  $x$  for the selected  $U_i, i \in s$ , is either known or can be known easily or inexpensively. In survey sampling literature, a great variety of

<sup>1</sup> First author is connected with Department of Statistics, Utkal University, Bhubaneswar 751004, India with e-mail Address: [Insahoostatuu@rediffmail.com](mailto:Insahoostatuu@rediffmail.com) and second is with the Department of Statistics, Salipur College, Cuttack 754202, India

estimation techniques consider only case A. Therefore, these techniques may not fully exhaust the information contents of the auxiliary variable. On the other hand, case B has also some practical merits. Because, when  $fsus$  are selected, information on means or totals of  $x$  for the selected  $fsus$  can be easily obtained from the available records on administration, geographical area, or from demographic sources, a census or a current population survey. As an example, we may refer to a crop survey where  $y$  and  $x$  are respectively the yield of a crop and area under the crop. Then information on the total crop area for each selected block (cluster of villages considered as  $fsus$ ) can be obtained easily from the block records.

The subject matter of this paper is to discuss and compare three different estimation methods *viz.*, classical method, chain method and predictive method. On the use of auxiliary information, the classical method considers case A while the other two methods are based on the consideration of an amalgam of cases A and B. To make a comparison between the said methods, we compare the performance of three different classes of estimators of  $\bar{Y}$  based on each of these methods which are already available in the literature.

## II. CLASSES OF ESTIMATORS UNDER STUDY

### 2.1 Class of Classical Estimators [Srivastava (1980)]

The class of estimators of Srivastava (1980) is defined by

$$t_c = h(\bar{y}, \bar{x}),$$

where  $(\bar{y}, \bar{x}) \in R_2$ , a subspace of the 2-dimensional real space containing the point  $(\bar{Y}, \bar{X})$ ,  $h(\bar{y}, \bar{x})$  is a known function of  $\bar{y}$  and  $\bar{x}$  such that  $h(\bar{Y}, \bar{X}) = \bar{Y}$ , and the function  $h$  satisfies the following regularity conditions:

- It is continuous in  $R_2$
- The first and second order partial derivatives of the function with respect to  $\bar{y}$  and  $\bar{x}$  exist and are also continuous in  $R_2$ .

An expression for the asymptotic variance of  $t_c$ , derived through the Taylor linearization method, is given by:

$$V(t_c) = \lambda(S_y^2 + h_1^2 S_x^2 + 2h_1 S_{yx}) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + h_1^2 S_{ix}^2 + 2h_1 S_{iyx}), \quad (2.1)$$

where  $\lambda_i = \left(\frac{1}{m_i} - \frac{1}{M_i}\right)$ ,  $\lambda = \left(\frac{1}{n} - \frac{1}{N}\right)$ ,  $S_{iy}^2 = \frac{1}{M_i-1} \sum_{j=1}^{M_i} (y_{ij} - \bar{Y}_i)^2$ ,  $S_{ix}^2 = \frac{1}{M_i-1} \sum_{j=1}^{M_i} (x_{ij} - \bar{X}_i)^2$ ,  
 $S_{iyx} = \frac{1}{M_i-1} \sum_{j=1}^{M_i} (y_{ij} - \bar{Y}_i)(x_{ij} - \bar{X}_i)$ ,  $S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (\alpha_i \bar{Y}_i - \bar{Y})^2$ ,  $S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (\alpha_i \bar{X}_i - \bar{X})^2$ ,  
 $S_{yx} = \frac{1}{N-1} \sum_{i=1}^N (\alpha_i \bar{Y}_i - \bar{Y})(\alpha_i \bar{X}_i - \bar{X})$  and  $h_1 = \left. \frac{\partial t_c}{\partial \bar{x}} \right|_{(\bar{y}, \bar{x}) = (\bar{Y}, \bar{X})}$ .

The asymptotic minimum variance bound (MVB) and the resulting minimum variance bound estimator (MVBE) of  $t_c$  are given by

$$\min V(t_c) = \lambda S_y^2 (1 - \rho_c^2) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{iy}^2 (1 - \rho_c^2) \tag{2.2}$$

and  $t_{RG}^{(c)} = \bar{y} - \beta_c (\bar{x} - \bar{X})$ ,

where  $\rho_c$  is the correlation coefficient between  $\bar{y}$  and  $\bar{x}$ , and  $\beta_c$  is the regression coefficient of  $\bar{y}$  on  $\bar{x}$ .

**2.2 Class of Chain Estimators [Sahoo and Panda (1997)]**

For any arbitrary second stage sample  $s_i, i \in s$ , let  $(\bar{y}_i, \bar{x}_i)$  assume values in  $R_2$  containing the point  $(\bar{Y}_i, \bar{X}_i)$  and  $f_i(\bar{y}_i, \bar{x}_i)$  be a function of  $(\bar{y}_i, \bar{x}_i)$  such that  $f_i(\bar{Y}_i, \bar{X}_i) = \bar{Y}_i$ . Further, for a given first stage sample  $s$ , let  $(t_y, t_x)$ , where  $t_y = \frac{1}{n} \sum_{i \in s} \alpha_i f_i(\bar{y}_i, \bar{x}_i)$ , assume values in  $R_2$  containing the point  $(\bar{Y}, \bar{X})$  and  $f(t_y, t_x)$  be a function of  $(t_y, t_x)$  such that  $f(\bar{Y}, \bar{X}) = \bar{Y}$ . Assuming that the functions  $f_i(\bar{y}_i, \bar{x}_i), i \in s$ , and  $f(t_y, t_x)$  admit Srivastava's (1980) regularity conditions, the class of estimators of Sahoo and Panda (1997) is defined by:

$$t_s = f(t_y, t_x).$$

The formula for the asymptotic variance of  $t_s$  is given by:

$$V(t_s) = \lambda (S_y^2 + f_1^2 S_x^2 + 2f_1 S_{yx}) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + f_{i1}^2 S_{ix}^2 + 2f_{i1} S_{iyx}), \tag{2.3}$$

where  $f_{i1} = \left. \frac{\partial f_i}{\partial \bar{x}_i} \right|_{(\bar{y}_i, \bar{x}_i) = (\bar{Y}_i, \bar{X}_i)}$  and  $f_1 = \left. \frac{\partial f}{\partial t_x} \right|_{(t_y, t_x) = (\bar{Y}, \bar{X})}$ . Hence, the MVB and the corresponding

MVBE of the class can be obtained as

$$\min V(t_s) = \lambda S_y^2 (1 - \rho^2) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{iy}^2 (1 - \rho_i^2) \tag{2.4}$$

and  $t_{RG}^{(s)} = \frac{1}{n} \sum_{i \in s} \alpha_i [\bar{y}_i + \beta_{iyx} (\bar{x}_i - \bar{X}_i)] - \beta_{yx} (t_x - \bar{X})$ ,

where  $\rho_i = S_{iyx} / S_{iy} S_{ix}$ ,  $\rho = S_{yx} / S_y S_x$ ,  $\beta_{iyx} = S_{iyx} / S_{ix}^2$  and  $\beta_{yx} = S_{yx} / S_x^2$ . This regression estimator was also considered by Sahoo (1987).

### 2.3 Class of Predictive Estimators

A class of estimators, that can be easily developed by utilizing prediction criterion given in Basu (1971, p.212, example 3) and subsequently studied by Sampford (1978). As a starting point, let us express the quantity  $\bar{Y}$  in the following form:

$$\bar{Y} = \frac{1}{M} \left[ \sum_{i \in s} \{m_i \bar{y}_i + (M_i - m_i) \bar{Y}_{ir}\} \right] + \frac{N-n}{N} \bar{Y}_r, \quad (2.5)$$

where  $(N-n)\bar{Y}_r = \sum_{i \in \bar{s}} \alpha_i \bar{Y}_i$  and  $(M_i - m_i)\bar{Y}_{ir} = \sum_{j \in \bar{s}_i} y_{ij}$ . According to this expression, we define a predictor  $\hat{\bar{Y}}$  for  $\bar{Y}$  as:

$$\hat{\bar{Y}} = \frac{1}{M} \left[ \sum_{i \in s} \{m_i \bar{y}_i + (M_i - m_i) T_i\} \right] + \frac{N-n}{N} T, \quad (2.6)$$

where  $T_i$  ( $i \in s$ ) and  $T$  are the respective predictors of  $\bar{Y}_{ir}$  and  $\bar{Y}_r$ .

For the realized samples  $s_i$  and  $s$ , let  $e_i = (\bar{y}_i, \bar{x}_i, \bar{X}_{ir})$  and  $e = (\bar{y}, t_x, \bar{X}_r)$  assume values in  $R_3$  containing the points  $E_i = (\bar{Y}_i, \bar{X}_i, \bar{X}_{ir})$  and  $E = (\bar{Y}, \bar{X}, \bar{X}_r)$ , where  $(N-n)\bar{X}_r = \sum_{i \in \bar{s}} \alpha_i \bar{X}_i$ ,  $(M_i - m_i) = \sum_{j \in \bar{s}_i} x_{ij}$ . Further, let  $g_i(e_i)$  and  $g(e)$  be some functions of  $e_i$  and  $e$  respectively such that  $g_i(E_i) = \bar{Y}_i, i \in s$ , and  $g(E) = \bar{Y}$ , and these functions satisfy Srivastava's regularity conditions. Thus, based on information available through  $s_i$  and  $s$ ,  $g_i(e_i)$  and  $g(e)$  clearly define classes of estimators for  $\bar{Y}_i, i \in s$ , and  $\bar{Y}$  respectively. Hence, on using  $g_i(e_i)$  and  $g(e)$  as predictors in places of  $T_i$  and  $T$  in our predictive equation (2.6), we may define a class of predictive estimators for  $\bar{Y}$  by:

$$t_p = \frac{1}{M} \left[ \sum_{i \in s} \{m_i \bar{y}_i + (M_i - m_i) g_i(e_i)\} \right] + \frac{N-n}{N} g(e),$$

with an asymptotic variance expression:

$$V(t_p) = \lambda(S_y^2 + 2g_1 S_{yx} + g_1^2 S_x^2) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + 2\theta g_{i1} S_{iyx} + \theta^2 g_{i1}^2 S_{ix}^2), \quad (2.7)$$

where  $\theta = \frac{n}{N}$ ,  $g_{i1} = \left. \frac{\partial g_i}{\partial \bar{x}_i} \right|_{e_i=E_i} = - \left. \frac{\partial g_i}{\partial \bar{X}_{ir}} \right|_{e_i=E_i}$  and  $g_1 = \left. \frac{\partial g}{\partial t_x} \right|_{e=E} = - \left. \frac{\partial g}{\partial \bar{X}_r} \right|_{e=E}$ . Thus, we note that

the MVB and the corresponding MVBE of  $t_p$  are:

$$\min V(t_p) = \min V(t_s)$$

and  $t_{RG}^{(p)} = t_{RG}^{(s)}$ .

### III. COMPARISON OF $t_c$ , $t_s$ AND $t_p$

Now, it seems necessary to make a good choice among the three classes of estimators defined by  $t_c$ ,  $t_s$  and  $t_p$  on the basis of certain performance criteria. But, as the literature to date offers little guidance in this choice, here for simplicity, we accept efficiency as the performance benchmark. With this performance tool, while comparing the analogous expressions (2.1), (2.3) and (2.7), we obtained sufficient conditions (suppressed to save space) under which one class could be claimed to be more efficient than others. But, these conditions depend heavily on the choices of various functions involved in composing of the classes and therefore cannot lead to any meaningful conclusions unless natures of these functions are clearly specified. Hence, to avoid this difficulty in direct comparison of efficiencies, we compare the classes by judging the (i) efficiencies of MVBs, and (ii) efficiencies of similar/equivalent estimators.

#### 3.1 Efficiencies of the Minimum Variance Bound Estimators

By accepting efficiency of MVBE as an intrinsic measure of efficiency of a class, we concentrate on the comparison of MVB expressions given in (2.2), (2.4) and (2.8). From these expressions, we now get:

$$\min V(t_p) = \min V(t_s) \leq \min V(t_c)$$

i.e.,  $t_{RG}^{(p)}$  and  $t_{RG}^{(s)}$  are more efficient than  $t_{RG}^{(c)}$  if

$$\rho_c \leq \rho \text{ and } \rho_i \quad \forall i. \tag{3.1}$$

Hence, we may conclude that in respect of MVB criterion and under the conditions (3.1), both chain and predictive methods are superior to the classical method. However, on this ground no conclusion can be drawn regarding the efficiency of chain method over predictive method or vice versa.

#### 3.2 Efficiencies of Equivalent Estimators

The primary purpose here is to select similar or equivalent type estimators belonging to different classes and then to make a comparative study of these estimators in respect of efficiency. But, for this selection, among the various available alternative estimators, we consider only those estimators having either standard ratio or standard regression features.

The regression estimator considered in Sukhatme *et al.* (1984, p.233) is defined by:

$$t_{RG} = \bar{y} - \beta_{yx}(\bar{x} - \bar{X}).$$

Expressing  $t_{RG}$  in a different form as:

$$t_{RG} = \frac{1}{n} \sum_{i \in S} \alpha_i [\bar{y}_i - \beta_{yx}(\bar{x}_i - \bar{X}_i)] - \beta_{yx}(t_x - \bar{X}),$$

we note that the estimator comes out as a common member of both  $t_c$  and  $t_s$ .

Sahoo and Panda (1999) developed a predictive regression-type estimator:

$$t_{PRG} = \frac{1}{n} \sum_{i \in s} \alpha_i [\bar{y}_i - \delta_i (\bar{x}_i - \bar{X}_i)] - \beta_{yx} (t_x - \bar{X}),$$

that can be easily viewed as a special case of  $t_p$ , where  $\delta_i = \beta_{yx} - \theta(\beta_{yx} - \beta_{iyx})$ .

Comparing asymptotic variance expressions of  $t_{RG}$  and  $t_{PRG}$  given by

$$V(t_{RG}) = \lambda S_y^2 (1 - \rho^2) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + \beta_{yx}^2 S_{ix}^2 - 2\beta_{yx} S_{iyx}) \quad (3.2)$$

$$V(t_{PRG}) = \lambda S_y^2 (1 - \rho^2) + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + \delta_i^2 S_{ix}^2 - 2\delta_i S_{iyx}), \quad (3.3)$$

it can be easily checked that  $t_{PRG}$  is more efficient than  $t_{RG}$ .

Turning our attention to ratio-type estimators, we consider the following estimators as particular cases of  $t_c$ ,  $t_s$  and  $t_p$  respectively:

$$\text{Classical ratio estimator : } t_R = \frac{\bar{y}}{\bar{x}} \bar{X}$$

$$\text{Chain ratio estimator : } t_{CR} = \frac{\frac{1}{n} \sum_{i \in s} \alpha_i \frac{\bar{y}_i}{\bar{x}_i} \bar{X}_i}{t_x} \bar{X} \quad [\text{Murthy (1977, p.390)}]$$

$$\text{Predictive ratio estimator : } t_{PR} = t_R + \frac{\theta}{n} \sum_{i \in s} \alpha_i \left( \frac{\bar{y}_i}{\bar{x}_i} - \frac{\bar{y}}{\bar{x}} \right) \bar{X}_i \quad [\text{Sahoo and Panda (1998)}]$$

Asymptotic variance expressions of these estimators are as follows:

$$V(t_R) = V_R + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + R^2 S_{ix}^2 - 2RS_{iyx}) \quad (3.4)$$

$$V(t_{CR}) = V_R + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + R_i^2 S_{ix}^2 - 2R_i S_{iyx}) \quad (3.5)$$

$$V(t_{PR}) = V_R + \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i (S_{iy}^2 + \phi_i^2 S_{ix}^2 - 2\phi_i S_{iyx}) \quad (3.6)$$

where  $V_R = \lambda(S_y^2 + R^2 S_x^2 - 2RS_{yx})$ ,  $R = \bar{Y}/\bar{X}$ ,  $R_i = \bar{Y}_i/\bar{X}_i$  and  $\phi_i = R - \theta(R - R_i)$ . From these expressions we obtain:

$$V(t_R) - V(t_{CR}) = \frac{1}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{ix}^2 (R - R_i) \{ (R + R_i) - 2\beta_{iyx} \} \quad (3.7)$$

$$V(t_R) - V(t_{PR}) = \frac{\theta}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{ix}^2 (R - R_i) \{ (R + \phi_i) - 2\beta_{iyx} \} \quad (3.8)$$

$$V(t_{PR}) - V(t_{CR}) = \frac{1-\theta}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{ix}^2 (R - R_i) \{ (R_i + \phi_i) - 2\beta_{iyx} \}. \quad (3.9)$$

Hence,  $V(t_{CR}) \leq V(t_R), V(t_{PR}) \leq V(t_R)$  and  $V(t_{CR}) \leq V(t_{PR})$  if  $\beta_{iyx} \leq \frac{1}{2}(R + R_i), \frac{1}{2}(R + \phi_i)$  and  $\frac{1}{2}(R_i + \phi_i)$  respectively provided  $R > R_i, \forall i$ . These results may be more compactly stated as follows:

$$\text{When } R > R_i, \forall i, V(t_{CR}) \leq V(t_{PR}) \leq V(t_R) \text{ if } \beta_{iyx} \leq \frac{1}{2}(R_i + \phi_i). \tag{3.10}$$

The requirement  $R > R_i$  need not be fulfilled for all *fsus* in *U* but for at least a majority of *fsus*.

We also examine the relative efficiencies of  $t_R, t_{CR}$  and  $t_{PR}$  under a super population model considered by Pfeffermann and Nathan (1981). The model assumes that:

$y_{ij} = \beta_i x_{ij} + e_{ij}, j = 1, 2, \dots, M_i; i = 1, 2, \dots, N,$   
 with  $E(e_{ij}/x_{ij}) = 0, E(e_{ij}^2/x_{ij}) = \sigma_i^2$  and  $E(e_{ij}e_{kl}/x_{ij}, x_{kl}) = 0$ , if either  $i \neq j$  or  $k \neq l$  or both. Further, following Scott and Smith (1969), let us assume the following random effects model for the  $\beta_i$ 's :

$$\beta_i = \beta + v_i, i = 1, 2, \dots, N,$$

such that  $E(v_i) = 0, E(v_i^2) = \delta^2$  and  $E(v_i v_j) = 0, i \neq j$ .

Under the model, after a considerable simplification, we derive the following expressions:

$$V(t_R) - V(t_{CR}) = \frac{\delta^2}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{ix}^2 \tag{3.11}$$

$$V(t_R) - V(t_{PR}) = \frac{\theta(2 - \theta)\delta^2}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{ix}^2 \tag{3.12}$$

$$V(t_{PR}) - V(t_{CR}) = \frac{(1 - \theta)\delta^2}{nN} \sum_{i=1}^N \alpha_i^2 \lambda_i S_{ix}^2. \tag{3.13}$$

These expressions are always positive. Thus, it seems that under the discussed model:

$$V(t_{CR}) \leq V(t_{PR}) \leq V(t_R).$$

In concluding this section, note that predictive regression estimator appears to be more efficient than other regression estimators whereas chain ratio estimator appears to be more efficient than other ratio estimators. This leads to an important conclusion that the chain method is preferred to others if the population regression line of *y* on *x* is linear and passing through the origin. Otherwise, the predictive method may have an increasing advantage over others.

#### IV. A SIMULATION STUDY

In this section, we report the results of a simulation study that was carried out to compare efficiencies of different estimators discussed in the preceding sections. We considered three sets of benchmark data of the MU 284 population consisting of  $M = 284$  municipalities (*ssu*) divided into  $N = 50$  clusters (*fsu*) available in Sarndal, Swensson and Wretman (1992, p.652). The two variables for these data were as follows:

*Data Set I* :  $y$  = Revenues from the 1985 municipal taxation,  $x$  = Number of Social-Democratic seats in municipal council.

*Data Set II* :  $y$  = 1985 population,  $x$  = 1975 population.

*Data Set III* :  $y$  = Number of municipal employees in 1984,  $x$  = Total number of seats in municipal council.

Our simulation consisted of 1000 independent first-stage samples each of size  $n = 16$ . From every selected  $U_i$  ( $i = 1, 2, \dots, 16$ ) in a first-stage sample, a second-stage sample of  $m_i = 2$  *ssus* was again selected. Thus, we had 1000 independent samples each of size 32 *ssus*. For each sample from 1 to 1000, values of the comparable ratio and regression estimators were computed and then their simulated mean square errors were calculated. Relative efficiencies of different estimators compared to the direct estimator  $\bar{y}$  are displayed in table 4.1.

Thus, as expected,  $t_{RG}^{(s)}$  leads to substantial increase in efficiency over others. On the other hand,  $t_{PRG}$  turns out to be more efficient than other regression estimators. Among the ratio estimators, the chain ratio estimator  $t_{CR}$  has a better performance than others. Although the findings of this empirical study agree with our theoretical findings, the scope is limited and may not fit other situations.

**Table 1 : Relative Efficiencies of the Estimators wrt  $\bar{y}$**

Data Set	Ratio Estimators			Regression Estimators			
	$t_R$	$t_{CR}$	$t_{PR}$	$t_{RG}^{(c)}$	$t_{RG}^{(s)}$	$t_{RG}$	$t_{PRG}$
I	137	149	143	125	163	107	126
II	159	204	181	165	192	149	179
III	118	135	121	110	144	121	139

#### Acknowledgement

The authors are grateful to the editor and referee for their useful comments that have helped to improve the paper.