SORT 33 (1) January-June 2009, 71-84

A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information

Housila P. Singh and Sunil Kumar

Abstract

In the present study, we propose a general class of estimators for population mean of the study variable in the presence of non-response using auxiliary information under double sampling. The expression of mean squared error (MSE) of the proposed class of estimators is derived under double (two-stage) sampling. Some estimators are also derived from the proposed class by allocating the suitable values of constants used. Comparisons of the proposed strategy with the usual unbiased estimator and other estimators are carried out. The results obtained are illustrated numerically using an empirical sample considered in the literature.

MSC: 62D05

Keywords: Double sampling, Mean squared error, Non-response, Study variable, Auxiliary variable.

1 Introduction

In practice almost all surveys suffer from non-response. The problem of non-response often happens due to the refusal of the subject, absenteeism and sometimes due to the lack of information. The pioneering work of Hansen and Hurwitz (1946), assumed that a sub-sample of initial non-respondents is recontacted with a more expensive method, suggesting the first attempt by mail questionnaire and the second attempt by a personal interview. In estimating population parameters such as the mean, total or ratio, sample survey experts sometimes use auxiliary information to improve

School of Studies in Statistics, Vikram University, Ujjain-456010, M. P., India
Received: January 2009
Accepted: April 2009

precision of the estimates. Sodipo and Obisesan (2007) have considered the problem of estimating the population mean in the presence of non-response, in sample survey with full response of an auxiliary character x. Other authors such as Cochran (1977), Rao (1986, 1987), Khare and Srivastava (1993, 1995, 1997), Okafor and Lee (2000) and Tabasum and Khan (2004, 2006) and Singh and Kumar (2008a,b) have studied the problem of non-response under double (two-stage) sampling.

From a finite population $U = (U_1, U_2, ..., U_N)$, a large first phase sample of size n' is selected by simple random sampling without replacement (SRSWOR). A smaller second phase sample of size n is selected from n' by SRSWOR. Non-response occurs on the second phase sample of size n in which n_1 units respond and n_2 units do not. From the n_2 non-respondents, by SRSWOR a sample of $r = n_2/k$; $k > 1$ units is selected where k is the inverse sampling rate at the second phase sample of size n. All the r units respond this time round. The auxiliary information can be used at the estimation stage to compensate for units selected for the sample that fail to provide adequate responses and for population units missing from the sampling frame. In a household survey, for example, the household size can be used as an auxiliary variable for the estimation of, say, family expenditure. Information can be obtained completely on the family size during the household listing while there may be non-response on the household expenditure.

An unbiased estimator for the population mean $\bar{Y}$ of the study variable y, proposed by Hansen and Hurwitz (1946), is defined by

$$\bar{y}^* = w_1\bar{y}_1 + w_2\bar{y}_{2r},$$

where $w_1 = n_1/n$ and $w_2 = n_2/n$. The variance of $\bar{y}^*$ is given by

$$Var(\bar{y}^*) = \left(\frac{1-f}{n}\right) S_y^2 + \frac{W_2(k-1)}{n} S_{y(2)}^2,$$

where $f = n/N$, $S_y^2 = \sum_{i=1}^{N} (y_i - \bar{Y})^2 \Big/ (N-1)$, $S_{y(2)}^2 = \sum_{i=1}^{N_2} (y_i - \bar{Y}_2)^2 \Big/ (N_2 - 1)$, $\bar{Y} = \sum_{i=1}^{N} y_i/N$, $\bar{Y}_2 = \sum_{i=1}^{N_2} y_i/N_2$, $W_2 = N_2/N$, $\bar{y}_{2r} = \sum_{i=1}^{r} y_i/r$, N_1 and $N_2\,(= N - N_1)$ are the sizes of the responding and non-responding units from the finite population N.

It is well known that in estimating the population mean, sample survey experts sometimes use auxiliary information to improve the precision of the estimates. Let x denote an auxiliary variable with population mean $\bar{X} = \sum_{i=1}^{N} x_i/N$. Let $\bar{X}_1 = \sum_{i=1}^{N_1} y_i/N_1$ and $\bar{X}_2 = \sum_{i=1}^{N_2} x_i/N_2$ denote the population means of the response and non-response groups (or strata). Let $\bar{x} = \sum_{i=1}^{n} x_i/n$ denote the mean of all the n units. Let $\bar{x}_1 = \sum_{i=1}^{n_1} x_i/n_1$ and $\bar{x}_2 = \sum_{i=1}^{n_2} x_i/n_2$ denote the means of the n_1 responding units and n_2 non-responding units. Further, let $\bar{x}_{2r} = \sum_{i=1}^{r} x_i/r$ denote the mean of the $r\,(= n_2/k), k > 1$ sub-sampled units. With this background we define an unbiased estimator of population mean $\bar{X}$ as

$$\bar{x}^* = w_1\bar{x}_1 + w_2\bar{x}_{2r}.$$

The variance of $\bar{x}^*$ is given by

$$Var\,(\bar{x}^*) = \left(\frac{1-f}{n}\right)S_x^2 + \frac{W_2\,(k-1)}{n}S_{x(2)}^2,$$

where $S_x^2 = \sum_{i=1}^{N}(x_i - \bar{X})^2 \big/ (N-1)$, $S_{x(2)}^2 = \sum_{i=1}^{N_2}(x_i - \bar{X}_2)^2 \big/ (N_2 - 1)$.

In the present study, we propose a general class of estimators for the population mean $\bar{Y}$ of the study variable y in the presence of non-response under double sampling using auxiliary information. The expressions of bias and variance have been obtained to the first degree of approximation of the proposed class of estimators, which will enable us to obtain these expressions for any member of this family. Some estimators are shown to be particular members of this family. Comparison of the proposed strategy with the usual unbiased estimator and other estimators are carried out. An empirical study is presented to expound the performance of the proposed class of estimators.

2 The proposed family of estimators

We define a class of estimators for the population mean $\bar{Y}$ of the study variable y as

$$T_{DS} = \bar{y}^* \left(\frac{a\bar{x}^* + b}{a\bar{x}' + b}\right)^{\alpha} \left(\frac{a\bar{x} + b}{a\bar{x}' + b}\right)^{\beta}, \tag{1}$$

where $\bar{x}'$ denote the sample mean of x based on first phase sample of size n', $a\,(\neq 0)$, b are either real numbers or functions of known parameters such as standard deviation (σ), Coefficient of variation (C_x), Correlation coefficient (ρ) etc. of the auxiliary variable x, and α, β are suitable chosen constants.

To obtain the bias and variance of the class of estimators T_{DS}, we write

$$\bar{y}^* = \bar{Y}(1+\varepsilon_0), \quad \bar{x}^* = \bar{X}(1+\varepsilon_1), \quad \bar{x}' = \bar{X}(1+\varepsilon_1'), \quad \bar{x} = \bar{X}(1+\varepsilon_2)$$

such that

$$E(\varepsilon_0) = E(\varepsilon_1) = E(\varepsilon_1') = E(\varepsilon_2) = 0$$

and

$$E\left(\varepsilon_0^2\right) = \lambda S_y^2 + \lambda^* S_{y(2)}^2, \quad E\left(\varepsilon_1^2\right) = \lambda S_x^2 + \lambda^* S_{x(2)}^2, \quad E\left(\varepsilon_1'^2\right) = \lambda' S_x^2, \quad E\left(\varepsilon_2^2\right) = \lambda S_x^2,$$
$$E(\varepsilon_0\varepsilon_1) = \lambda \rho_{yx} S_y S_x + \lambda^* \rho_{yx(2)} S_{y(2)} S_{x(2)}, \quad E(\varepsilon_2\varepsilon_1') = \lambda' S_x^2, \quad E(\varepsilon_1\varepsilon_1') = \lambda' S_x^2,$$
$$E(\varepsilon_1\varepsilon_2) = \lambda S_x^2, \quad E(\varepsilon_0\varepsilon_2) = \lambda \rho_{yx} S_y S_x, \quad E(\varepsilon_0\varepsilon_1') = \lambda' \rho_{yx} S_y S_x,$$

where ρ_{yx} and $\rho_{yx(2)}$ are respectively the correlation coefficient of response and non-response group between study variable y and auxiliary variable x,

$$\lambda = \left(\frac{1-f}{n}\right), \quad \lambda' = \left(\frac{1-f'}{n'}\right), \quad \lambda^* = \frac{W_2(k-1)}{n} \quad \text{and} \quad f' = n'/N.$$

Now expressing T_{DS} in terms of $\varepsilon's$ we have

$$T_{DS} = \bar{Y}(1+\varepsilon_0)(1+\phi\varepsilon_1)^{\alpha}(1+\phi\varepsilon_1')^{-\alpha}(1+\phi\varepsilon_2)^{\beta}(1+\phi\varepsilon_1')^{-\beta}, \tag{2}$$

where

$$\phi = \left(\frac{a\bar{X}}{a\bar{X}+b}\right).$$

We assume that $|\phi\varepsilon_1| < 1$, $|\phi\varepsilon_1'| < 1$ and $|\phi\varepsilon_2| < 1$ so that the right hand side of (2) is expandable. Now, expanding the right hand side of (2) to the first degree of approximation, we have

$$\begin{aligned}(T_{DS}-\bar{Y}) =\ & \bar{Y}\{\varepsilon_0 + \beta\phi(\varepsilon_2 + \varepsilon_0\varepsilon_2 - \varepsilon_0\varepsilon_1') + \alpha\phi(\varepsilon_1 + \varepsilon_0\varepsilon_1 - \varepsilon_0\varepsilon_1') - (\alpha+\beta)\phi\varepsilon_1' \\ & + \alpha\beta\phi^2\left(\varepsilon_1'^2 - \varepsilon_1'\varepsilon_2 - \varepsilon_1\varepsilon_1' + \varepsilon_1\varepsilon_2\right) - \phi^2\left(\beta^2\varepsilon_2\varepsilon_1' + \alpha^2\varepsilon_1\varepsilon_1'\right) \\ & + \frac{\beta(\beta+1)}{2}\phi^2\left(\varepsilon_1'^2 + \varepsilon_2^2\right) + \frac{\alpha(\alpha+1)}{2}\phi^2\left(\varepsilon_1'^2 + \varepsilon_1^2\right)\Big\}.\end{aligned} \tag{3}$$

Taking expectations of both sides of (3), we get that the bias of T_{DS} to the first degree of approximation is given by

$$B\left(T_{DS}\right) = \bar{Y}\left[\left(\lambda-\lambda'\right)\phi\left\{\alpha\left(K_{yx}+\frac{\alpha-1}{2}\phi\right)+\beta\left(K_{yx}+\alpha\phi+\frac{\beta-1}{2}\phi\right)\right\}C_x^2 + \lambda^*\alpha\phi\left(K_{yx(2)}+\frac{\alpha-1}{2}\phi\right)C_{x(2)}^2\right], \tag{4}$$

where

$$C_x^2=\frac{S_x^2}{\bar{X}^2},\quad C_{x(2)}^2=\frac{S_{x(2)}^2}{\bar{X}^2},\quad K_{yx}=\frac{\beta_{yx}}{R},\quad K_{yx(2)}=\frac{\beta_{yx(2)}}{R},\quad R=\frac{\bar{Y}}{\bar{X}},$$

$$\beta_{yx}=\frac{S_{yx}}{S_x^2},\quad \beta_{yx(2)}=\frac{S_{yx(2)}}{S_{x(2)}^2},\quad S_{yx}=\sum_{i=1}^{N}\left(x_i-\bar{X}\right)\left(y_i-\bar{Y}\right)\big/\left(N-1\right),$$

$$S_{yx(2)}=\sum_{i=1}^{N_2}\left(x_i-\bar{X}_2\right)\left(y_i-\bar{Y}_2\right)\big/\left(N_2-1\right).$$

Squaring both sides of (3) and neglecting terms of $\varepsilon's$ involving power greater than two, we have

$$\begin{aligned}\left(T_{DS}-\bar{Y}\right)^2 &= \bar{Y}^2\left\{\varepsilon_0+\alpha\phi\varepsilon_1+\beta\phi\varepsilon_2-(\alpha+\beta)\phi\varepsilon_1'\right\}^2\\ &= \bar{Y}^2\Big\{\varepsilon_0^2+\left(\alpha^2\varepsilon_1^2+\beta^2\varepsilon_2^2+2\alpha\beta\varepsilon_1\varepsilon_2\right)\phi^2+(\alpha+\beta)^2\phi^2\varepsilon_1'^2\\ &+ 2\phi\left(\alpha\varepsilon_0\varepsilon_1+\beta\varepsilon_0\varepsilon_2\right)-2(\alpha+\beta)\phi\varepsilon_1'\varepsilon_0-2(\alpha+\beta)\phi^2\left(\alpha\varepsilon_1\varepsilon_1'+\beta\varepsilon_2\varepsilon_1'\right)\Big\}.\end{aligned} \tag{5}$$

Taking expectations of both sides of (5), we get the variance of T_{DS} to the first degree of approximation, we get

$$Var\left(T_{DS}\right) = \bar{Y}^2\left[\left(\lambda-\lambda'\right)\left\{C_y^2+(\alpha+\beta)\phi\left((\alpha+\beta)\phi+2K_{yx}\right)C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\alpha\phi\left(\alpha\phi+2K_{yx(2)}\right)C_{x(2)}^2\right\}+\lambda'C_y^2\right], \tag{6}$$

The variance of T_{DS} is minimized for

$$\begin{aligned}\alpha &= -\frac{K_{yx(2)}}{\phi}\\ \beta &= \left(\frac{1}{\phi}\right)\left(K_{yx(2)}-K_{yx}\right) = -\left(\frac{1}{\phi}\right)\left(K_{yx}-K_{yx(2)}\right),\end{aligned} \tag{7}$$

Substituting (7) in (1), we get the asymptotically optimum estimator (AOE) as

$$T_{DS(opt)}=\bar{y}^*\left\{\frac{\left(a\bar{x}'+b\right)^2}{\left(a\bar{x}^*+b\right)\left(a\bar{x}+b\right)}\right\}^{K_{yx(2)}/\phi}\left(\frac{a\bar{x}'+b}{a\bar{x}+b}\right)^{K_{yx}/\phi}. \tag{8}$$

The minimum variance of T_{DS} is given by

$$\begin{aligned} \min Var(T_{DS}) &= \bar{Y}^2\left[(\lambda-\lambda')\left(1-\rho_{yx}^2\right)C_y^2+\lambda^*\left(1-\rho_{yx(2)}^2\right)C_{y(2)}^2+\lambda' C_y^2\right] \\ &= Var\left(T_{DS(opt)}\right). \end{aligned} \tag{9}$$

2.1 Some members of the proposed class of estimators T_{DS}

The following are the estimators of the population mean which can be obtained by suitable choices of constants α, β, a and b.

Estimator	α	β	a	b
$T_{DS}^{(0)}=\bar{y}^*$ Usual unbiased estimator	0	0	a	b
$T_{DS}^{(1)}=\bar{y}^*\left(\frac{\bar{x}'}{\bar{x}^*}\right)$ Khare and Srivastava (1993), Tabasum and Khan's (2004) ratio estimator	−1	0	1	0
$T_{DS}^{(2)}=\bar{y}^*\left(\frac{\bar{x}'}{\bar{x}}\right)$ Khare and Srivastava (1993), Tabasum and Khan's (2006) ratio estimator	0	−1	1	0
$T_{DS}^{(3)}=\bar{y}^*\left(\frac{\bar{x}'}{\bar{x}^*}\right)\left(\frac{\bar{x}'}{\bar{x}}\right)$ Singh and Kumar's (2008a) ratio estimator	−1	−1	1	0
$T_{DS}^{(4)}=\bar{y}^*\left(\frac{\bar{x}^*}{\bar{x}'}\right)\left(\frac{\bar{x}}{\bar{x}'}\right)$ Singh and Kumar's (2008a) product estimator	1	1	1	0
$T_{DS}^{(5)}=\bar{y}^*\left(\frac{\bar{x}'+C_x}{\bar{x}^*+C_x}\right)$	−1	0	1	C_x
$T_{DS}^{(6)}=\bar{y}^*\left(\frac{\bar{x}'+C_x}{\bar{x}+C_x}\right)$	0	−1	1	C_x
$T_{DS}^{(7)}=\bar{y}^*\left(\frac{\bar{x}'+C_x}{\bar{x}^*+C_x}\right)\left(\frac{\bar{x}'+C_x}{\bar{x}+C_x}\right)$	−1	−1	1	C_x
$T_{DS}^{(8)}=\bar{y}^*\left(\frac{\bar{x}'+\rho}{\bar{x}^*+\rho}\right)$	−1	0	1	ρ
$T_{DS}^{(9)}=\bar{y}^*\left(\frac{\bar{x}'+\rho}{\bar{x}+\rho}\right)$	0	−1	1	ρ
$T_{DS}^{(10)}=\bar{y}^*\left(\frac{\bar{x}'+\rho}{\bar{x}^*+\rho}\right)\left(\frac{\bar{x}'+\rho}{\bar{x}+\rho}\right)$	−1	−1	1	ρ

where C_x is the coefficient of variation of the auxiliary variable x and ρ the correlation coefficient between the study variable y and the auxiliary variable x.

Many more estimators can also be generated from the proposed estimator in (1) just by putting different values of α, β, a and b. The expressions of bias and variance of the above estimators can be obtained by mere substituting the values of α, β, a and b in (4) and (6), respectively. Up to the first degree of approximation, the bias and variance expressions of various estimators are

$$B\left(T_{DS}^{(0)}\right) = \bar{Y}, \tag{10}$$

$$B\left(T_{DS}^{(1)}\right) = \bar{Y}\left\{(\lambda-\lambda')(1-K_{yx})C_x^2 + \lambda^*\left(1-K_{yx(2)}\right)C_{x(2)}^2\right\}, \tag{11}$$

$$B\left(T_{DS}^{(2)}\right) = \bar{Y}(\lambda-\lambda')(1-K_{yx})C_x^2, \tag{12}$$

$$B\left(T_{DS}^{(3)}\right) = \bar{Y}\left\{(\lambda-\lambda')(3-2K_{yx})C_x^2 + \lambda^*\left(1-K_{yx(2)}\right)C_{x(2)}^2\right\}, \tag{13}$$

$$B\left(T_{DS}^{(4)}\right) = \bar{Y}\left\{(\lambda-\lambda')(1+2K_{yx})C_x^2 + \lambda^* K_{yx(2)}C_{x(2)}^2\right\}, \tag{14}$$

$$B\left(T_{DS}^{(5)}\right) = \bar{Y}\left\{(\lambda-\lambda')\phi'(\phi'-K_{yx})C_x^2 + \lambda^*\phi'\left(\phi'-K_{yx(2)}\right)C_{x(2)}^2\right\}, \tag{15}$$

$$B\left(T_{DS}^{(6)}\right) = \bar{Y}(\lambda-\lambda')\phi'(\phi'-K_{yx})C_x^2, \tag{16}$$

$$B\left(T_{DS}^{(7)}\right) = \bar{Y}\left\{(\lambda-\lambda')\phi'(3\phi'-2K_{yx})C_x^2 + \lambda^*\phi'\left(\phi'-K_{yx(2)}\right)C_{x(2)}^2\right\}, \tag{17}$$

$$B\left(T_{DS}^{(8)}\right) = \bar{Y}\left\{(\lambda-\lambda')\phi^*(\phi^*-K_{yx})C_x^2 + \lambda^*\phi^*\left(\phi^*-K_{yx(2)}\right)C_{x(2)}^2\right\}, \tag{18}$$

$$B\left(T_{DS}^{(9)}\right) = \bar{Y}(\lambda-\lambda')\phi^*(\phi^*-K_{yx})C_x^2, \tag{19}$$

$$B\left(T_{DS}^{(10)}\right) = \bar{Y}\left\{(\lambda-\lambda')\phi^*(3\phi^*-2K_{yx})C_x^2 + \lambda^*\phi^*\left(\phi^*-K_{yx(2)}\right)C_{x(2)}^2\right\}, \tag{20}$$

$$Var\left(T_{DS}^{(0)}\right) = \bar{Y}^2\left\{\lambda C_y^2 + \lambda^* C_{y(2)}^2\right\}, \tag{21}$$

$$Var\left(T_{DS}^{(1)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+(1-2K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\left(1-2K_{yx(2)}\right)C_{x(2)}^2\right\} + \lambda' C_y^2\right], \tag{22}$$

$$Var\left(T_{DS}^{(2)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+(1-2K_{yx})C_x^2\right\} + \lambda^* C_{y(2)}^2 + \lambda' C_y^2\right], \tag{23}$$

$$Var\left(T_{DS}^{(3)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+4(1-K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\left(1-2K_{yx(2)}\right)C_{x(2)}^2\right\} + \lambda' C_y^2\right], \tag{24}$$

$$Var\left(T_{DS}^{(4)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+4(1+K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\left(1+2K_{yx(2)}\right)C_{x(2)}^2\right\} + \lambda' C_y^2\right], \tag{25}$$

$$Var\left(T_{DS}^{(5)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+\phi'(\phi'-2K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\phi'\left(\phi'-2K_{yx(2)}\right)C_{x(2)}^2\right\} + \lambda' C_y^2\right], \tag{26}$$

$$Var\left(T_{Ds}^{(6)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+\phi'(\phi'-2K_{yx})C_x^2\right\} + \lambda^* C_{y(2)}^2 + \lambda' C_y^2\right], \tag{27}$$

$$Var\left(T_{DS}^{(7)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+4\phi'(\phi'-K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\phi'\left(\phi'-2K_{yx(2)}\right)C_{x(2)}^2\right\} + \lambda' C_y^2\right], \tag{28}$$

$$Var\left(T_{DS}^{(8)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+\phi^*(\phi^*-2K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2+\phi^*\left(\phi^*-2K_{yx(2)}\right)C_{x(2)}^2\right\} + \lambda' C_y^2\right], \tag{29}$$

$$Var\left(T_{DS}^{(9)}\right) = \bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+\phi^*(\phi^*-2K_{yx})C_x^2\right\} + \lambda^* C_{y(2)}^2 + \lambda' C_y^2\right], \tag{30}$$

$$Var\left(T_{DS}^{(10)}\right)=\bar{Y}^2\left[(\lambda-\lambda')\left\{C_y^2+4\phi^*\left(\phi^*-K_{yx}\right)C_x^2\right\}+\lambda^*\left\{C_{y(2)}^2+\phi^*\left(\phi^*-2K_{yx(2)}\right)C_{x(2)}^2\right\}+\right.$$
$$\left.+\lambda' C_y^2\right], \tag{31}$$

where $\phi'=\left(\frac{\bar{X}}{\bar{X}+C_x}\right)$ and $\phi^*=\left(\frac{\bar{X}}{\bar{X}+\rho}\right)$.

2.2 Efficiency comparison

The proposed class of estimators T_{DS} is more efficient than

(i) usual unbiased estimator $T_{DS}^{(0)}=\bar{y}^*$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left\{-\left(\frac{2K_{yx}}{\phi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\phi}\right)\right\} \\ or \max.\left\{-\left(\frac{2K_{yx}}{\phi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\phi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{32}$$

(ii) usual ratio estimator $T_{DS}^{(1)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left\{-\left(\frac{2K_{yx}-1}{\phi}+\beta\right);-\left(\frac{2K_{yx(2)}-1}{\phi}\right)\right\} \\ or \max.\left\{-\left(\frac{2K_{yx}-1}{\phi}+\beta\right);-\left(\frac{2K_{yx(2)}-1}{\phi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{33}$$

(iii) the ratio estimator $T_{DS}^{(2)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left\{-\left(\frac{2\left(K_{yx}-1\right)}{\phi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\phi}\right)\right\} \\ or \max.\left\{-\left(\frac{2\left(K_{yx}-1\right)}{\phi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\phi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{34}$$

(iv) the ratio estimator $T_{DS}^{(3)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left[-\left\{\frac{2\left(K_{yx}-1\right)}{\phi}+\beta\right\};-\left(\frac{2K_{yx(2)}-1}{\phi}\right)\right] \\ or \max.\left[-\left\{\frac{2\left(K_{yx}-1\right)}{\phi}+\beta\right\};-\left(\frac{2K_{yx(2)}-1}{\phi}\right)\right]<\alpha<0 \end{array}\right\}, \tag{35}$$

(v) the product estimator $T_{DS}^{(4)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left[-\left\{\dfrac{2\left(K_{yx}+1\right)}{\phi}+\beta\right\};-\left(\dfrac{2K_{yx(2)}+1}{\phi}\right)\right] \\ or\ \ \max.\left[-\left\{\dfrac{2\left(K_{yx}+1\right)}{\phi}+\beta\right\};-\left(\dfrac{2K_{yx(2)}+1}{\phi}\right)\right]<\alpha<0 \end{array}\right\}, \tag{36}$$

(vi) the estimator $T_{DS}^{(5)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left[-\left\{\left(\dfrac{2K_{yx}-\phi'}{\phi}\right)+\beta\right\};-\left(\dfrac{2K_{yx(2)}-\phi'}{\phi}\right)\right] \\ or\ \ \max.\left[-\left\{\left(\dfrac{2K_{yx}-\phi'}{\phi}\right)+\beta\right\};-\left(\dfrac{2K_{yx(2)}-\phi'}{\phi}\right)\right]<\alpha<0 \end{array}\right\}, \tag{37}$$

(vii) the estimator $T_{DS}^{(6)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left\{-\left(\dfrac{2K_{yx}-\phi'}{\phi}+\beta\right);-\left(\dfrac{2K_{yx(2)}}{\phi}\right)\right\} \\ or\ \ \max.\left\{-\left(\dfrac{2K_{yx}-\phi'}{\phi}+\beta\right);-\left(\dfrac{2K_{yx(2)}}{\phi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{38}$$

(viii) the estimator $T_{DS}^{(7)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left[-\left\{\dfrac{2\left(K_{yx}-\phi'\right)}{\phi}+\beta\right\};\left(\dfrac{2K_{yx(2)}-\phi'}{\phi}\right)\right] \\ or\ \ \max.\left[-\left\{\dfrac{2\left(K_{yx}-\phi'\right)}{\phi}+\beta\right\};\left(\dfrac{2K_{yx(2)}-\phi'}{\phi}\right)\right]<\alpha<0 \end{array}\right\}, \tag{39}$$

(ix) the estimator $T_{DS}^{(8)}$ if

$$\left.\begin{array}{r} 0<\alpha<\min.\left\{-\left(\dfrac{2K_{yx}-\phi^{*}}{\phi}+\beta\right);-\left(\dfrac{2K_{yx(2)}-\phi^{*}}{\phi}\right)\right\} \\ or\ \ \max.\left\{-\left(\dfrac{2K_{yx}-\phi^{*}}{\phi}+\beta\right);-\left(\dfrac{2K_{yx(2)}-\phi^{*}}{\phi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{40}$$

(x) the estimator $T_{DS}^{(9)}$ if

$$\left.\begin{array}{l} 0<\alpha<\min.\left\{-\left(\dfrac{2K_{yx}-\phi^{*}}{\phi}+\beta\right);\ -\left(\dfrac{2K_{yx(2)}}{\phi}\right)\right\} \\ or\ \ \min.\left\{-\left(\dfrac{2K_{yx}-\phi^{*}}{\phi}+\beta\right);\ -\left(\dfrac{2K_{yx(2)}}{\phi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{41}$$

(xi) the estimator $T_{DS}^{(10)}$ if

$$\left.\begin{array}{l} 0<\alpha<\min.\left[-\left\{\dfrac{2\left(K_{yx}-\phi^{*}\right)}{\phi}+\beta\right\};\ -\left(\dfrac{2K_{yx(2)}-\phi^{*}}{\phi}\right)\right] \\ or\ \ \max.\left[-\left\{\dfrac{2\left(K_{yx}-\phi^{*}\right)}{\phi}+\beta\right\};\ -\left(\dfrac{2K_{yx(2)}-\phi^{*}}{\phi}\right)\right]<\alpha<0 \end{array}\right\}. \tag{42}$$

3 Empirical study

To illustrate the properties of the proposed estimators of the population mean $\bar{Y}$, we consider a real data set considered before by Srivastava (1993). The description of the sample is given below.

The sample of 100 consecutive trips (after omitting 20 outlier values) measured by two fuel meters for a small family car in normal usage given by Lewis et al (1991) has been taken into consideration. The measurement of turbine meter (in ml) is considered as main character y and the measurement of displacement meter (in cm^3) is considered as auxiliary character x. We treat the last 25% values as non-response values. The values of the parameters are as follows:

$$\bar{Y}=3500.12,\quad \bar{X}=260.84,\quad S_y=2079.30,\quad S_x=156.40,\quad \bar{Y}_2=3401.08,$$
$$\bar{X}_2=259.96,\quad S_{y(2)}=1726.02,\quad S_{x(2)}=134.36,\quad \rho=0.985,\quad \rho_2=0.995.$$

Here, we have computed (i) the absolute relative bias of different suggested estimators of $\bar{Y}$ using the formula:

$$ARB\,(\,.\,)=\left|\frac{Bias\,(\,.\,)}{\bar{Y}}\right|;$$

(ii) the percent relative efficiencies (PRE) of different suggested estimators with respect to the usual unbiased estimator $\bar{y}^{*}$, for different values of k.

Table 1: *Absolute relative bias (ARB) of different proposed estimators.*

Estimators	$N = 100,\quad n' = 50,\quad n = 30$			
	$(1/k)$			
	$(1/5)$	$(1/4)$	$(1/3)$	$(1/2)$
$\bar{y}^* = T_{DS}^{(0)}$	0.0000	0.0000	0.0000	0.0000
$T_{DS}^{(1)}$	0.00053	0.00043	0.00032	0.00022
$T_{DS}^{(2)}$	0.00012	0.00012	0.00012	0.00012
$T_{DS}^{(3)}$	0.00543	0.00533	0.00522	0.00512
$T_{DS}^{(4)}$	0.02251	0.02041	0.01831	0.01621
$T_{DS}^{(5)}$	0.00050	0.00040	0.00030	0.00020
$T_{DS}^{(6)}$	0.00010	0.00010	0.00010	0.00010
$T_{DS}^{(7)}$	0.00536	0.00527	0.00517	0.00507
$T_{DS}^{(8)}$	0.00048	0.00038	0.00029	0.00019
$T_{DS}^{(9)}$	0.00010	0.00010	0.00010	0.00010
$T_{DS}^{(10)}$	0.00532	0.00523	0.00513	0.00503

Table 2: *Percent relative efficiency (PRE) of the estimators with respect to $\bar{y}^*$.*

Estimators	$N = 100,\quad n' = 50,\quad n = 30$			
	$(1/k)$			
	$(1/5)$	$(1/4)$	$(1/3)$	$(1/2)$
$\bar{y}^* = T_{DS}^{(0)}$	100.00	100.00	100.00	100.00
$T_{DS}^{(1)}$	433.12	381.95	330.09	277.54
$T_{DS}^{(2)}$	138.74	146.79	159.06	180.07
$T_{DS}^{(3)}$	185.74	163.18	140.48	117.65
$T_{DS}^{(4)}$	35.15	30.83	26.49	22.17
$T_{DS}^{(5)}$	309.02	284.54	257.42	227.21
$T_{DS}^{(6)}$	132.15	138.46	147.83	163.25
$T_{DS}^{(7)}$	322.99	298.09	270.38	239.33
$T_{DS}^{(8)}$	433.56	382.27	330.30	277.65
$T_{DS}^{(9)}$	138.75	146.80	159.08	180.09
$T_{DS}^{(10)}$	187.38	164.60	141.70	118.66
$T_{DS(opt)}$	435.74	383.77	331.23	278.12

It is observed from Table 1 that the absolute relative bias (ARB) of the estimators $T_{DS}^{(j)};\ j = 1, 3, 4, 5, 7, 8, 10$ decreases with the increase of $(1/k)$ while it remains constant for $T_{DS}^{(2)}$, $T_{DS}^{(6)}$ and $T_{DS}^{(9)}$. The ARB of the estimator $T_{DS}^{(4)}$ is larger than all other estimators. It may be due to positive correlation. The estimator $T_{DS}^{(4)}$ is a product type estimator which is appropriate in the situations where the correlation between y and x is negative. It is further observed from Table 1 that

$$ARB\left(T_{DS}^{(0)}\right) < ARB\left(T_{DS}^{(9)}\right) = ARB\left(T_{DS}^{(6)}\right) < ARB\left(T_{DS}^{(2)}\right) < ARB\left(T_{DS}^{(8)}\right) < ARB \left(T_{DS}^{(5)}\right) < ARB\left(T_{DS}^{(1)}\right) < ARB\left(T_{DS}^{(10)}\right) < ARB\left(T_{DS}^{(7)}\right) < ARB\left(T_{DS}^{(3)}\right) < ARB\left(T_{DS}^{(4)}\right)$$

which clearly indicates that the estimator $T_{DS}^{(8)}$ (based on the knowledge of correlation coefficient) has least magnitude of relative bias followed by $T_{DS}^{(1)}$.

From Table 2, we see that the proposed general class of estimators is more desirable over all the considered estimators under optimum condition. It is observed from Table 2 that the percent relative efficiency of the estimators $T_{DS}^{(1)}$, $T_{DS}^{(3)}$, $T_{DS}^{(4)}$, $T_{DS}^{(5)}$, $T_{DS}^{(7)}$, $T_{DS}^{(8)}$ and $T_{DS(opt)}$ decreases as $\left(1/k\right)$ increases, but for the estimators $T_{DS}^{(2)}$, $T_{DS}^{(6)}$ and $T_{DS}^{(9)}$, it increases with the increase in the value of $\left(1/k\right)$.

From Tables 1 and 2, it is further observed that the estimator $T_{DS}^{(8)}$ (based on known correlation coefficient) seems to be more appropriate estimator in comparison to others as it has appreciable efficiency (close to the efficiency of the optimum estimator $T_{DS(opt)}$) as well as negligible magnitude of relative bias. However, the estimators $T_{DS}^{(5)}$ and $T_{DS}^{(7)}$ (based on known coefficient of variation) are also appropriate choices among the estimators as they have considerable gain in efficiency as well as lower relative bias.

Finally, we conclude that the proposed estimator $T_{DS}^{(5)}$, $T_{DS}^{(7)}$ and $T_{DS}^{(8)}$ are better alternatives of the optimum estimator $T_{DS(opt)}$.

Acknowledgement

Authors are thankful to the referee for his valuable constructive suggestions regarding improvement of the paper.

References

Cochran, W. G. (1977). *Sampling Techniques*, 3rd *ed.*, John Wiley and Sons, New York.

Hansen, M. H. and Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *J. Amer. Statist. Assoc.,* 41, 517-529.

Khare, B. B. and Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response. *Nat. Acad. Sc. Letters, India*, 16, 111-114.

Khare, B. B. and Srivastava, S. (1995). Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response. *Proc. Nat. Acad. Sci. India,* 65(A), 195-203.

Khare, B. B. and Srivastava, S. (1997). Transformed ratio type estimators for the population mean in the presence of non-response. *Comm. Statist.-Theory Methods*, 26, 1779-1791.

Lewis, P. A., Jones, P. W., Polak, J. W. and Tillotson, H. T. (1991). The problem of conversion in method comparison studies. *Applied Statistics*, 40, 105-112.

Okafor, F. C. and Lee, H. (2000). Double sampling for ratio and regression estimation with sub sampling the non-respondent. *Survey Methodology*, 26, 183-188.

Rao, P. S. R. S. (1986). Ratio Estimation with sub sampling the non-respondents. *Survey Methodology*, 12, 217-230.

Rao, P. S. R. S. (1987). Ratio and regression estimates with sub-sampling the non-respondents. *Paper presented at a special contributed session of the International Statistical Association Meetings*, September, 2-16, Tokyo, Japan.

Singh, H. P. and Kumar, S. (2008a). Estimation of mean in presence of non-response using two phase sampling scheme. *Statistical Papers*, DOI10.1007/s00362-008-0140-5.

Singh, H. P. and Kumar, S. (2008b). A regression approach to the estimation of finite population mean in presence of non-response. *Aust. N. Z. J. Stat.*, 50, 395-408.

Sodipo, A. A. and Obisesan, K. O. (2007). Estimation of the population mean using difference cum ratio estimator with full response on the auxiliary character. *Res. J. Applied Sci*., 2, 769-772.

Srivastava, S. (1993). Some problems on the estimation of population mean using auxiliary character in presence of non-response in sample surveys. *Thesis submitted to Banaras Hindu University, Varanasi, India*.

Tabasum, R. and Khan, I. A. (2004). Double sampling for ratio estimation with non-response. *J. Ind. Soc. Agril. Statist.*, 58, 300-306.

Tabasum, R. and Khan, I. A. (2006). Double sampling ratio estimator for the population mean in presence of non-response. *Assam Statist. Review*, 20, 73-83.