Available online at www.isas.org.in

JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 63(2) 2009 175-180

A Class of Predictive Estimators in Two-Stage Sampling

L.N. Sahoo[1*], B.C. Das[1] and J. Sahoo[2]

[1]*Utkal University, Bhubaneswar 751004*

[2]*Department of Statistics, S.K.C.G. College, Parlakhemundi 761200*

(Received: May 2005, Accepted: August 2009)

SUMMARY

Under the well known prediction approach of Basu (1971), we introduce a new class of estimators for the finite population mean availing information on two auxiliary variables in a two-stage sampling.

Keywords: Asymptotic variance, Auxiliary variable, Prediction approach, Two-stage sampling.

1. INTRODUCTION

Consider a finite population U, partitioned into N first stage units (fsu) denoted by $U_1, U_2, ..., U_N$ such that the number of second stage units (ssu) in U_i is M_i and $M = \sum_{i=1}^{N} M_i$. Let y_{ij} and x_{ij} denote values of the study variable y and an auxiliary variable x respectively, for the j^{th} ssu of U_i ($j = 1, 2, ..., M_i$; $i = 1, 2, ..., N$). Define $\bar{Y}_i = \frac{1}{M_i}\sum_{j=1}^{M_i} y_{ij}$, $\bar{X}_i = \frac{1}{M_i}\sum_{j=1}^{M_i} x_{ij}$ as the means of U_i and $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} u_i \bar{Y}_i$, $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} u_i \bar{X}_i$ as the overall population means, where $u_i = NM_i/M$. To estimate $\bar{Y}$, assume that a sample s of n fsus is drawn from U and then a sample s_i of m_i ssus from the selected U_i is drawn according to the design simple random sampling without replacement. Let $\bar{y}_i = \frac{1}{m_i}\sum_{j \in s_i} y_{ij}$, $\bar{x}_i = \frac{1}{m_i}\sum_{j \in s_i} x_{ij}$, $\bar{y} = \frac{1}{n}\sum_{i \in s} u_i \bar{y}_i$, $\bar{x} = \frac{1}{n}\sum_{i \in s} u_i \bar{x}_i$ and $\bar{x}' = \frac{1}{n}\sum_{i \in s} u_i \bar{X}_i$.

When $\bar{X}$ is known accurately, Srivastava's (1980) class of estimators is defined by $t_s = \gamma(\bar{y}, \bar{x})$, where $\gamma(\bar{y}, \bar{x})$ is a function of $\bar{y}$ and $\bar{x}$, such that $\gamma(\bar{Y}, \bar{X}) = \bar{Y}$ and satisfies certain regularity conditions in R_2, a 2-dimensional real space containing the point $(\bar{Y}, \bar{X})$. But, in a two-stage sampling plan it is usually felt that efficiency of an estimator depends on how well the auxiliary information can be utilized at different stages. With this spirit, using known values of $\bar{X}_i$'s for the selected fsus, Sahoo and Panda (1997) considered a class of estimators defined by

$$t_{sp} = \mu\left(\frac{1}{n}\sum_{i \in s} u_i \mu_i(\bar{y}_i, \bar{x}_i), \bar{x}'\right)$$

[*] *Corresponding author* : L.N. Sahoo

E-mail address : lnsahoostatuu@rediffmail.com

such that $\mu(\bar{Y}, \bar{X}) = \bar{Y}$, $\mu_i(\bar{Y}_i, \bar{X}_i) = \bar{Y}_i$, $i \in s$ and the functions $\mu(.,.)$ and $\mu_i(.,.)$ admit Srivastava's (1980) regularity conditions in R_2.

In certain practical situations we get information on another strong auxiliary variable z, taking value z_{ij} for the j^{th} ssu of U_i in such a way that the overall population mean $\bar{Z}$ is unknown, but the population means for the selected fsus *i.e.*, $\bar{Z}_i$, $i \in s$, are known. In this context, using both auxiliary variables, Sahoo and Sahoo (2005) composed a class of estimators defined by

$$t_{ss} = \alpha\left(\frac{1}{n}\sum_{i\in s} u_i \alpha_i(\bar{y}_i, \bar{z}_i), \bar{x}'\right)$$

where $\bar{z}_i = \frac{1}{m_i}\sum_{j\in s_i} z_{ij}$ and the functions α and α_i satisfy regularity conditions in R_2. The basic assumption behind the construction of t_{ss} is that $\bar{X}_i, \bar{Z}_i, i \in s$ and $\bar{X}$ are known but $\bar{Z}$ is unknown. However, guided by these assumptions, here we develop a general class of estimators for $\bar{Y}$ motivated by the predictive approach of Basu (1971, p. 212, example 3).

As an example of this type of situation, we may refer to a crop survey conducted in a district with block (cluster of villages) as the fsu and village as the ssu. If y, x and z represent respectively yield, cultivated area and area under wheat, then information on the average area under cultivation per village in the i^{th} block, *i.e.*, $\bar{X}_i$ for $i \in s$, can be obtained at a low cost from the block records and average area under cultivation for the district *i.e.*, $\bar{X}$ can be known from the district records. Information on $\bar{Z}_i$ *i.e.*, average area under wheat for the i^{th} selected block can also be easily available from the block level records.

2. PREDICTION CRITERION IN TWO-STAGE SAMPLING

Let $\bar{s}$ denote the set of $(N - n)$ fsus of U which are not included in s and $\bar{s}_i$, the set of $(M_i - m_i)$ ssus of U_i which are not included in s_i, $i \in s$. Under the usual predictive set-up, it is possible to express

$$\bar{Y} = \frac{1}{M}\left[\sum_{i\in s}\left\{\sum_{j\in s_i} y_{ij} + \sum_{j\in \bar{s}_i} y_{ij}\right\} + \sum_{i\in\bar{s}} M_i \bar{Y}_i\right] \quad (1)$$

Writing $(N - n)\bar{Y}_r = \sum_{i\in\bar{s}} u_i \bar{Y}_i$

and $(M_i - m_i)\bar{Y}_{ir} = \sum_{j\in\bar{s}_i} y_{ij}$, we have

$$\bar{Y} = \frac{1}{M}\left[\sum_{i\in s}\left\{m_i \bar{y}_i + (M_i - m_i)\bar{Y}_{ir}\right\}\right] + \frac{N-n}{N}\bar{Y}_r \quad (2)$$

To estimate $\bar{Y}$, we, therefore have to predict the quantities $\bar{Y}_{ir}$ and $\bar{Y}_r$ from the sample data because the first component of the right hand side of (2) is already known. Using T_i and T as their predictors, a predictor $\hat{\bar{Y}}$ of $\bar{Y}$ of may be defined by the equation

$$\hat{\bar{Y}} = \frac{1}{M}\left[\sum_{i\in s}\left\{m_i \bar{y}_i + (M_i - m_i)T_i\right\}\right] + \frac{N-n}{N}T \quad (3)$$

Note that if $m_i = M_i$ and $n = N$; $\hat{\bar{Y}} = \bar{Y}$ the target of our prediction.

Corresponding to various suitable choices of the predictors $T_i\,(i \in s)$ and T, equation (3) generates a class of estimators. But, we achieve this objective by defining these predictors in terms of two auxiliary variables *i.e.*, x and z.

3. THE CLASS OF PREDICTIVE ESTIMATORS

For given s_i and s, let

$$e_i = (\bar{y}_i, \bar{x}_i, \bar{z}_i, \bar{X}_{ir}, \bar{Z}_{ir}) \in R_5$$

and $$e = (\bar{y}, \bar{x}', \bar{X}_r) \in R_3,$$

where $(M_i - m_i)\bar{X}_{ir} = \sum_{j\in\bar{s}_i} x_{ij}$, $(M_i - m_i)\bar{Z}_{ir} = \sum_{j\in\bar{s}_i} z_{ij}$, $(N-n)\bar{X}_r = \sum_{i\in\bar{s}} u_i \bar{X}_i$, R_5 and R_3 are 5- and 3-dimensional real spaces containing the points $E_i = (\bar{Y}_i, \bar{X}_i, \bar{Z}_i, \bar{X}_i, \bar{Z}_i)$, $i \in s$ and $E = (\bar{Y}, \bar{X}, \bar{X})$ respectively. Further, let $h_i(e_i)$ and $h(e)$ be some known functions of e_i and e respectively such that $h_i(E_i) = \bar{Y}_i$, $i \in s$, and $h(E) = \bar{Y}$. Let us assume that

(a) the functions h_i and h are continuous in R_5 and R_3 respectively, and

(b) the first and second order partial derivatives of these functions with respect to their arguments exist and are also continuous in their respective range spaces.

Thus, based on information available on s_i and s, $h_i(e_i)$ and $h(e)$ clearly define classes of estimators for $\bar{Y}_i, i \in s$, and $\bar{Y}$ respectively. Using $h_i(e_i)$ and $h(e)$ as predictors in places of T_i and T in our predictive equation (3), we now define a class of predictive estimators for $\bar{Y}$ by

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

Many estimators may turn out as special cases of t_h corresponding to various selections of h_i and h. Let us consider the following simple cases:

(i) If the information on x is completely ignored, *i.e.*, if $h_i = \bar{y}_i$ and $h = \bar{y}$ then t_h becomes $\bar{y}$, the simple expansion estimator of $\bar{Y}$.

(ii) When $h_i = \dfrac{\bar{y}_i\bar{X}_{ir}\bar{Z}_{ir}}{\bar{x}_i\bar{z}_i}$ and $h = \dfrac{\bar{y}\bar{X}_r}{\bar{x}'}$, then

$$t_h \to t_R^{(h)} = \bar{y}\frac{\bar{X}}{\bar{x}'} - \frac{f}{n}\sum_{i\in s}u_i\bar{y}_i \times \left[(1-f_i) - \frac{1}{1-f_i}\left(\frac{\bar{X}_i}{\bar{x}_i} - f_i\right)\left(\frac{\bar{Z}_i}{\bar{z}_i} - f_i\right)\right]$$

a ratio-type estimator, where $f = \dfrac{n}{N}$ and $f_i = \dfrac{m_i}{M_i}$.

(iii) When $h_i = \dfrac{\bar{y}_i\bar{x}_i\bar{z}_i}{\bar{X}_{ir}\bar{Z}_{ir}}$ and $h = \dfrac{\bar{y}\,\bar{x}'}{\bar{X}_r}$, then

$$t_h \to t_P^{(h)} = \frac{f}{n}\sum_{i\in s}u_i \times \left[f_i\bar{y}_i + (1-f_i)^3\frac{\bar{y}_i\bar{x}_i\bar{z}_i}{(\bar{X}_i - f_i\bar{x}_i)(\bar{Z}_i - f_i\bar{z}_i)}\right] + (1-f)^2\frac{\bar{y}\,\bar{x}'}{\bar{X} - f\,\bar{x}'}$$

a product-type estimator.

(iv) When $h_i = \bar{y}_i - \beta_{iyx}(\bar{x}_i - \bar{X}_{ir}) - \beta_{iyz}(\bar{z}_i - \bar{Z}_{ir})$ and

$$h = \bar{y} - \beta_{byx}(\bar{x}' - \bar{X}_r)$$

$$t_h \to t_{RG}^{(h)} = \bar{y} - \frac{f}{n}\sum_{i\in s}u_i\left\{\beta_{iyx}(\bar{x}_i - \bar{X}_i) + \beta_{iyz}(\bar{z}_i - \bar{Z}_i)\right\} - \beta_{byx}(\bar{x}' - \bar{X})$$

a regression-type estimator, where

$\beta_{\text{iyz}} = S_{iyx} / S_{ix}^2$,

$\beta_{\text{iyz}} = S_{iyz} / S_{iz}^2$, $\beta_{\text{byz}} = S_{byx} / S_{bx}^2$ such that

$$S_{iyx} = \frac{1}{M_i - 1}\sum_{j=1}^{M_i}(y_{ij} - \bar{Y}_i)(x_{ij} - \bar{X}_i)$$

$$S_{byx} = \frac{1}{N-1}\sum_{i=1}^{N}(u_i\bar{Y}_i - \bar{Y})(u_i\bar{X}_i - \bar{X})$$

$$S_{iz}^2 = \frac{1}{M_i - 1}\sum_{j=1}^{M_i}(z_{ij} - \bar{Z}_i)^2$$

$$S_{by}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(u_i\bar{Y}_i - \bar{Y})^2, \text{ etc.}$$

(v) If the estimation procedure is carried out with the involvement of x only, then $h_i = d_i(\bar{y}_i, \bar{x}_i, \bar{X}_{ir})$ so that $t_h \to t_h^{(d)}$, a class of predictive estimators defined by

$$t_h^{(d)} = \frac{1}{M}\left[\sum_{i\in s}\{m_i\bar{y}_i + (M_i - m_i)d_i(\bar{y}_i, \bar{x}_i, \bar{X}_{ir})\}\right] + \frac{N-n}{N}h(e)$$

(vi) As a specific case of t_h, we may also consider another subclass of predictive estimators defined by

$$t_h^{(k)} = \frac{1}{M}\left[\sum_{i\in s}\{m_i\bar{y}_i + (M_i - m_i)k_i(\bar{y}_i, \bar{z}_i, \bar{Z}_{ir})\}\right] + \frac{N-n}{N}h(e)$$

on considering $h_i = k_i\ (\bar{y}_i, \bar{z}_i, \bar{Z}_{ir})$.

4. ASYMPTOTIC VARIANCE OF t_h

Expanding $h_i(e_i)$ and around the points E_i and E respectively in a first order Taylor's series and then neglecting the remainder term, we get

$$h_i(e_i) = h_i(E_i) + h_{i0}(\bar{y}_i - \bar{Y}_i) + h_{i1}(\bar{x}_i - \bar{X}_i) + h_{i2}(\bar{z}_i - \bar{Z}_i) + h_{i3}(\bar{X}_{ir} - \bar{X}_i) + h_{i4}(\bar{Z}_{ir} - \bar{Z}_i) \quad (4)$$

and

$$h(e) = h(E) + h_0(\bar{y} - \bar{Y}) + h_1(\bar{x}' - \bar{X}) + h_2(\bar{X}_r - \bar{X}) \quad (5)$$

where h_{i0}, h_{i1}, h_{i2}, h_{i3}, and h_{i4} are respectively the values of first order partial derivatives of $h_i(e_i)$ with respect to $\bar{y}_i, \bar{x}_i, \bar{z}_i, \bar{X}_{ir}$ and $\bar{Z}_{ir}$ at E_i and h_0, h_1 and h_2 are respectively the values of first order partial derivatives of $h(e)$ with respect to $\bar{y}, \bar{x}'$ and $\bar{X}_r$ at E.

Noting that $h_{i0} = 1$, $h_{i1} = -h_{i3}$, $h_{i2} = -h_{i4}$, $\bar{X}_{ir} = \dfrac{M_i\bar{X}_i - m_i\bar{x}_i}{M_i - m_i}$, $\bar{Z}_{ir} = \dfrac{M_i\bar{Z}_i - m_i\bar{z}_i}{M_i - m_i}$ we have after a considerable simplification

$$t_h = \bar{y} + \frac{f}{n}\sum_{i\in s} u_i[h_{i1}(\bar{x}_i - \bar{X}_i) + h_{i2}(\bar{z}_i - \bar{Z}_i)] + h_1(\bar{x}' - \bar{X}) \quad (6)$$

Hence, after a few tedious algebraic steps (suppressed to save space), the asymptotic variance of t_h is obtained as

$$V(t_h) = \frac{1-f}{n}\left(S_{by}^2 + h_1^2 S_{bx}^2 + 2h_1 S_{byx}\right) + \frac{1}{nN}\sum_{i=1}^{N} u_i^2 \frac{1-f_i}{m_i} V \quad (7)$$

where $V_i = S_{iy}^2 + f^2 h_{i1}^2 S_{ix}^2 + f^2 h_{i2}^2 S_{iz}^2 + 2 f h_{i1} S_{iyx} + 2 f h_{i2} S_{iyz} + 2 f^2 h_{i1} h_{i2} S_{ixz}$

Minimizing $V(t_h)$ over h_{i1}, h_{i2} and h_1 we get

$$h_{i1} = -\frac{1}{f}\frac{\beta_{iyx} - \beta_{iyz}\beta_{izx}}{1 - \beta_{izx}\beta_{ixz}} = h_{i1}^* \quad \text{(say)}$$

$$h_{i2} = -\frac{1}{f}\frac{\beta_{iyz} - \beta_{iyx}\beta_{ixz}}{1 - \beta_{izx}\beta_{ixz}} = h_{i2}^* \quad \text{(say)}$$

and $\quad h_1 = -\beta_{byx}$

where $\beta_{izx} = S_{izx} / S_{ix}^2$, $\beta_{ixz} = S_{ixz} / S_{iz}^2$. Use of these optimum values in (7) yields the minimum asymptotic variance of the class (may be called as the asymptotic minimum variance bound (MVB) of the class) is given by

$$\min V(t_h) = \frac{1-f}{n} S_{by}^2 (1 - \rho_{byx}^2) + \frac{1}{nN}\sum_{i=1}^{N} u_i^2 \frac{1-f_i}{m_i} S_{iy}^2 (1 - \rho_i^2) \quad (8)$$

where $\quad \rho_{byx} = S_{byx} / S_{by}S_{bx}$ and

$$\rho_i = \sqrt{\frac{\rho_{iyx}^2 + \rho_{iyz}^2 - 2\rho_{iyx}\rho_{iyz}\rho_{ixz}}{1 - \rho_{ixz}^2}}$$

the multiple correlation coefficient of y on x and z in U_i such that $\rho_{iyx} = S_{iyx} / S_{iy} S_{ix}$ etc. An estimator attaining this bound is called as an MVB estimator. In the present context our MVB estimator is a regression-type estimator of the form

$$t_{RG}^0 = \bar{y} - \frac{1}{n}\sum_{i\in s} u_i\left[h_{i1}^*(\bar{x}_i - \bar{X}_i) + h_{i2}^*(\bar{z}_i - \bar{Z}_i)\right] - \beta_{byx}(\bar{x}' - \bar{X})$$

The parametric functions h_{i1}^*, h_{i2}^* and β_{byx} can be replaced by their consistent estimates computed from the sample itself. But, the asymptotic variance of the resulting estimator remains unchanged and is given by (8).

5. PRECISION OF t_h

In an effort to study the efficiency aspect of the predictive method of estimation developed in this work in relation to the classical method, our first attempt is to compare the efficiency of t_h with that of t_s. The asymptotic variance of t_s obtained through Taylor linearization is given by

$$V(t_s) = \frac{1-f}{n}(S_{by}^2 + \gamma_1^2 S_{bx}^2 + 2\gamma_1 S_{byx}) + \frac{1}{nN}\sum_{i=1}^{N} u_i^2 \frac{1-f_i}{m_i}(S_{iy}^2 + \gamma_1^2 S_{ix}^2 + 2\gamma_1 S_{iyx}) \quad (9)$$

where γ_1 is the first order partial derivative of $\gamma(\bar{y}, \bar{x})$ with respect to $\bar{x}$ when evaluated at $(\bar{Y}, \bar{X})$.

From (7) and (9), it follows that $V(t_h) \le V(t_s)$ *i.e.*, an estimator of t_h is more precise than an estimator of t_s if

$$\left|\gamma_1 + \beta_{byx}\right| \ge \left|h_1 + \beta_{byx}\right|$$

and $S_{ix}^2\left[(\gamma_1 + \beta_{iyx})^2 - (fh_{i1} + \beta_{iyx})^2\right]$

$$\ge fh_{i1} S_{iz}^2 (fh_{i2} + 2\beta_{iyz} + 2fh_{i1}\beta_{ixz}) \quad \forall \; i \quad (10)$$

These sufficient conditions basically depend on the choices of different functions for composing t_h and t_s. However, they give some indication that there is enough scope for improving upon the estimators through our predictive method over classical method. But, these conditions can not lead to any straight forward conclusions if the characteristics of the functions are unknown. However, for simplicity, if we accept MVB as an intrinsic measure of precision of a class, the problem of precision comparison seems to be easier and our attention will be concentrated on the MVB estimators only.

The minimum asymptotic variance of t_s is

$$\min V(t_s) = \frac{1-f}{n} S_{by}^2 (1-\rho^2) + \frac{1}{nN}\sum_{i=1}^{N} u_i^2 \frac{1-f_i}{m_i} S_{iy}^2 (1-\rho^2) \quad (11)$$

and the corresponding MVB estimator is

$$t_{RG}^{(s)} = \bar{y} - \beta(\bar{x} - \bar{X})$$

where ρ is the correlation coefficient between $\bar{y}$ and $\bar{x}$ and β is the regression coefficient of $\bar{y}$ on $\bar{x}$. Hence, we see that

$$\min V(t_h) \le \min V(t_s)$$

i.e., t_{RG}^0 is more efficient than $t_{RG}^{(s)}$ if

$$\rho^2 \le \rho_{byx}^2 \text{ and } \rho_i^2 \quad \forall \; i \quad (12)$$

Turning our attention to study the precision of t_h compared to other classes of classical and predictive estimators *viz*., t_{sp}, t_{ss}, $t_h^{(d)}$ and $t_h^{(k)}$ on the ground of MVB criterion, we see that

$$\min V(t_{sp}) = \frac{1-f}{n} S_{by}^2 (1-\rho_{byx}^2) + \frac{1}{nN}\sum_{i=1}^{N} u_i^2 \frac{1-f_i}{m_i} S_{iy}^2 (1-\rho_{iyx}^2) \quad (13)$$

$$\min V(t_{ss}) = \frac{1-f}{n} S_{by}^2 (1-\rho_{byx}^2) + \frac{1}{nN}\sum_{i=1}^{N} u_i^2 \frac{1-f_i}{m_i} S_{iy}^2 (1-\rho_{iyz}^2) \quad (14)$$

$$\min V(t_h^{(d)}) = \min V(t_{sp})$$

$$\min V(t_h^{(k)}) = \min V(t_{ss})$$

The MVB estimators of t_{sp} or $t_h^{(d)}$ and t_{ss} or $t_h^{(k)}$ are also respectively given by

$$t_{RG}^{(sp)} = \bar{y} - \frac{1}{n}\sum_{i \in s} u_i \beta_{iyx}(\bar{x}_i - \bar{X}_i) - \beta_{byx}(\bar{x}' - \bar{X})$$

$$t_{RG}^{(ss)} = \bar{y} - \frac{1}{n}\sum_{i \in s} u_i \beta_{iyz}(\bar{z}_i - \bar{Z}_i) - \beta_{byx}(\bar{x}' - \bar{X})$$

From (8), (13) and (14) we have

$$\min V(t_h) \le \min V(t_{sp}) \Rightarrow V(t_{RG}^0) \le V(t_{RG}^{(sp)})$$

and $\min V(t_h) \le \min V(t_{ss}) \Rightarrow V(t_{RG}^0) \le V(t_{RG}^{(ss)})$

Hence, we may conclude that t_h is superior to t_{sp}, t_{ss}, $t_h^{(d)}$ and $t_h^{(k)}$ on the ground of MVB criterion.

6. NUMERICAL STUDY

To study precision of the suggested methodology numerically, we consider data of two populations as described below.

Population 1. Consists of 198 blocks (ssu) divided into $N = 27$ wards of Berhampur city of Orissa. The number of blocks (M_i) of 27 wards are 6, 6, 12, 5, 6, 6, 10, 5, 6, 6, 6, 6, 6, 12, 6, 7, 7, 7, 10, 6, 6, 7, 10, 11, 9, 8 and 6. The three variables *viz*., number of educated females, female population and number of households are used as y, x and z respectively, data on which are available in Census of India (1971) document. We have taken $n = 9$ and m_i = 2, 2, 4, 2, 2, 2, 3, 2, 2, 2, 2, 2, 2, 4, 2, 2, 2, 2, 3, 2, 2, 2, 3, 4, 3, 3 and 2 respectively.

Population 2. MU284 population available in Sarndal *et al.* (1992, p. 660, Appendix C). It consists of 284 municipalities (ssu) divided into 50 clusters (fsu) with three variables *viz*., Revenue from the 1985 municipal taxation as y, 1975 population as x and 1985 population as z. We consider $n = 12$, and $m_i = 2$ for every i.

Relative precision of different MVB estimators compared to the simple expansion estimator $\overline{y}$, are compiled in Table 1. The estimator t_{RG}^{0} attains the maximum precision for both populations. Thus, our numerical study shows that the new methodology developed here to create predictive estimators may be useful for many practical situations.

Table 1. Relative precision of different estimators

Pop. No.	Estimators				
	$\overline{y}$	$t_{RG}^{(s)}$	$t_{RG}^{(sp)}$	$t_{RG}^{(ss)}$	t_{RG}^{0}
1	100	148	184	175	195
2	100	947	3579	3366	3725

REFERENCES

Basu, D. (1971). An essay on the logical foundations of survey sampling, Part I. In: *Foundations of Statistical Inference*, V.P. Godambe and D.A. Sprott (eds), Holt, Rinehart and Winston, Toronto, Canada, 203-242.

Sahoo, L.N. and Panda, P. (1997). A class of estimators in two-stage sampling with varying probabilities. *South African Statist. J.*, **31**, 151-160.

Sahoo, L.N. and Sahoo, R.K. (2005). On the construction of a class of estimators in two-stage sampling. *J. Appl. Statist. Sci.*, **14**, 317-322.

Sarndal, C.E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer-Verlag.

Srivastava, S.K. (1980). A class of estimators using auxiliary information in sample surveys. *Canad. J. Stat.*, **8**, 253-254.