Multiple Frames in Repeat Surveys

U.C. Sud, A.K. Srivastava and I.C. Sethi
Indian Agricultural Statistics Research Institute, New Delhi-12
(Received : May, 1999)

SUMMARY

The prediction approach has been used for estimation of parameters in sampling on two occasions when multiple frames are available on both the occasions. Predictors have primarily been developed for the current occasion. Gains through multiple frames are illustrated using 'a single frame' on the first occasion and 'two frames' on the second occasion.

Key words : Multiple frames, Prediction approach, Repeat surveys.

1. Introduction

Hartley [3] and [4] considered dual frame surveys in which there is a complete frame which is expensive to sample while another frame which is inexpensive to sample but is incomplete, is available. Vogel [11], Serruier and Philips [9], Armstrong [1] used the multiple frame technique in applied work. Rao [6] considered the problem of non-response in multiple frames. The theory of multiple frame was extended to two-stage sampling design as well as multiple characters by Saxena *et al.* ([7], [8]). Skinner and Rao [10] used pseudo maximum likelihood approach for estimation of domain sizes in multiple frames. In this paper the problem of prediction of finite population mean for a survey repeated on two occasions is attempted when multiple frames are available on both the occasions. Several alternatives have been considered – two frames on two occasions, one frame on first occasion and two frames on the second occasion and vice versa.

Two frame surveys are common in practice where list and area frames are available. For example in evaluating the impact of milk supply schemes on rural economy, impact studies are repeated over time, a tentative list of commercial milk producers normally supplying milk is available at the cooperative milk collection center and another updated list is obtained from the usual survey.

2. Multiple Frames on Both the Occasions

Let there be two overlapping frames A and B with sizes N_A and N_B which together constitute the entire population. This population can be classified into three domains (a), (ab) and (b) such that (a) consist of units belonging to frame A only, (ab) with units belonging to both A and B frames while (b) that of units from B frame only. Let N_a, N_{ab} and N_b be the respective domain sizes. Observe that the total population size is $N = N_a + N_b + N_{ab}$. We assume that the super population of which finite population is a realization is described by the relationship

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \tag{2.1}$$

where,

the random variable Y refers to the character under study, ε denotes the error terms, μ is the parameter of the super-population, i is the occasion identifier $\{i = 1, 2\}$, j is the domain identifier $\{j \in (a), (ab), (b)\}$, and k is the observation identifier, $\{k = 1, ..., N_j\}$.

Further,

$$E_m(\varepsilon_{ijk}) = 0$$

$$E_m(\varepsilon_{ijk}, \varepsilon_{i'j'k'}) = \begin{cases} \sigma^2 & \forall\, i = i', j = j', k = k' \\ \rho_a\,\sigma^2 & \forall\, i \neq i', j = j' \varepsilon (a), k = k \\ \rho_{ab}\,\sigma^2 & \forall\, i \neq i', j = j' \varepsilon (ab), k = k' \\ \rho_b\,\sigma^2 & \forall\, i \neq i', j = j' \varepsilon (b), k = k' \\ 0 & \forall\, i \neq i', j \neq j', k \neq k' \end{cases}$$

where, E_m denotes model based expectation.

Let the finite population mean on the second occasion be denoted by $\bar{y}_2$.

$$\bar{y}_2 = \frac{1}{N} \sum_{j \varepsilon (a),(ab),(b)} \sum_{k=1}^{N_j} y_{2jk}$$

On the first occasion, let independent sample of sizes n_A and n_B be drawn from frames A and B respectively. Then,

$$n_A = n_a + n_{ab};\; n_B = n_b + n_{ba}$$

where n_a, n_b are sample sizes from frames A and B belonging to (a) and (b) respectively while n_{ab} and n_{ba} belong to (ab) selected from frames A and B respectively. Random sub-samples of m_A and m_B units are retained for use on the second occasion from frames A and B respectively. Independent samples of sizes u_A and u_B are selected (unmatched with the first occasion) from A and B frames respectively.

For simplicity, we assume that the sample sizes are same on both the occasions.

Then $n_A = u_A + m_A ; n_B = u_B + m_B$ holds for both the occasions.

On the first occasion

$$u_A = u_{1a} + u_{1ab} ; m_A = m_a + m_{ab}$$

$$u_B = u_{1b} + u_{1ba} ; m_B = m_b + m_{ba}$$

while for the second occasion

$$u_A = u_{2a} + u_{2ab} ; m_A = m_a + m_{ab}$$

$$u_B = u_{2b} + u_{2ba} ; m_B = m_b + m_{ba}$$

where $u_{1a}, u_{1ab}, m_a, m_{ab}$ etc. are defined as above for the two occasions.

The model for the sampled data for both the occasions can be written in a compact form as

$$\underline{y} = \underset{\sim}{X} \underset{\sim}{\beta} + \underline{\varepsilon}$$

with

$$E_m\left(\underline{\varepsilon}\right) = 0, E_m\left(\underline{\varepsilon}\,\underline{\varepsilon}'\right) = \sigma^2 \underset{\sim}{\Sigma}$$

Here the vector $\underline{y}(2n \times 1)$ is a realization of the vector Y, X is the $2n \times 8$ matrix given by

$$\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{X_1} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{X_2} \end{bmatrix}$$

where $\underset{\sim}{0}$ is a null matrix of order $(n \times 4)$ and the matrices $\underset{\sim}{X_1}$ and $\underset{\sim}{X_2}$ are of the order $(n \times 4)$ defined by

$$\underset{\sim}{X_1} = \begin{bmatrix} \underline{1}_{u_{1a}} & \underline{0}_{u_{1a}} & \underline{0}_{u_{1a}} & \underline{0}_{u_{1a}} \\ \underline{0}_{u_{1b}} & \underline{1}_{u_{1b}} & \underline{0}_{u_{1b}} & \underline{0}_{u_{1b}} \\ \underline{0}_{u_{1ab}} & \underline{0}_{u_{1ab}} & \underline{1}_{u_{1ab}} & \underline{0}_{u_{1ab}} \\ \underline{0}_{u_{1ba}} & \underline{0}_{u_{1ba}} & \underline{0}_{u_{1ba}} & \underline{1}_{u_{1ba}} \\ \underline{1}_{m_a} & \underline{0}_{m_a} & \underline{0}_{m_a} & \underline{0}_{m_a} \\ \underline{0}_{m_b} & \underline{1}_{m_b} & \underline{0}_{m_b} & \underline{0}_{m_b} \\ \underline{0}_{m_{ab}} & \underline{0}_{m_{ab}} & \underline{1}_{m_{ab}} & \underline{0}_{m_{ab}} \\ \underline{0}_{m_{ba}} & \underline{0}_{m_{ba}} & \underline{0}_{m_{ba}} & \underline{1}_{m_{ba}} \end{bmatrix}, \quad \underset{\sim}{X_2} = \begin{bmatrix} \underline{1}_{m_a} & \underline{0}_{m_a} & \underline{0}_{m_a} & \underline{0}_{m_a} \\ \underline{0}_{m_b} & \underline{1}_{m_b} & \underline{0}_{m_b} & \underline{0}_{m_b} \\ \underline{0}_{m_{ab}} & \underline{0}_{m_{ab}} & \underline{1}_{m_{ab}} & \underline{0}_{m_{ab}} \\ \underline{0}_{m_{ba}} & \underline{0}_{m_{ba}} & \underline{0}_{m_{ba}} & \underline{1}_{m_{ba}} \\ \underline{1}_{u_{2a}} & \underline{0}_{u_{2a}} & \underline{0}_{u_{2a}} & \underline{0}_{u_{2a}} \\ \underline{0}_{u_{2b}} & \underline{1}_{u_{2b}} & \underline{0}_{u_{2b}} & \underline{0}_{u_{2b}} \\ \underline{0}_{u_{2ab}} & \underline{0}_{u_{2ab}} & \underline{1}_{u_{2ab}} & \underline{0}_{u_{2ab}} \\ \underline{0}_{u_{2ba}} & \underline{0}_{u_{2ba}} & \underline{0}_{u_{2ba}} & \underline{1}_{u_{2ba}} \end{bmatrix}$$

in which

$1_{\underline{m}_a}$ is the column vector of order $(m_a \times 1)$ with all the elements equal to 1.

$0_{\underline{m}_a}$ is a column vector of order $(m_a \times 1)$ with all the elements equal to 0.

Other terms in the matrices $\underset{\sim}{X_1}$ and $\underset{\sim}{X_2}$ can be similarly defined.

Also, $n = n_A + n_B$

$\underline{\beta}'$ is a 1×8 row vector of parameters having the structure

$$\left[\mu_{1a} \;\; \mu_{1ab} \;\; \mu_{1b} \;\; \mu_{1ba} \;\; \mu_{2a} \;\; \mu_{2ab} \;\; \mu_{2b} \;\; \mu_{2ba}\right]'$$

and $\underline{\varepsilon}$ is $2n \times 1$ vector of error terms.

$\underset{\sim}{\Sigma}$ is $2n \times 2n$ matrix having the structure

$$\underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{\Sigma_{11}} & \underset{\sim}{\Sigma_{12}} \\ \underset{\sim}{\Sigma'_{12}} & \underset{\sim}{\Sigma_{22}} \end{bmatrix}$$

where, $\underset{\sim}{\Sigma_{11}}$ and $\underset{\sim}{\Sigma_{22}}$ respectively are identity matrices of order $n \times n$. Similarly, $\underset{\sim}{\Sigma'_{12}}$ is of order $n \times n$ with the structure

$$\begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{\sigma_{12}} & \underset{\sim}{0} \end{bmatrix}$$

where

$$\underline{\sigma}_{12} = \begin{bmatrix} \underset{\sim}{\rho_a I_{m_a}} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\rho_{ab} I_{m_{ab}}} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{\rho_b I_{m_b}} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{\rho_{ab} I_{m_{ba}}} \end{bmatrix}$$

Estimate of $\underline{\beta}$ can be obtained by the generalised least squares technique. Let $\hat{\underline{\beta}}$ be the estimator of $\underline{\beta}$, then

$$\hat{\underline{\beta}} = \left(\underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{X} \right)^{-1} \underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1} \underline{y} \tag{2.2}$$

Consider the predictor of $\bar{y}_2$ defined by

$$\hat{\bar{y}}_2 = \frac{1}{N}\left[N_a \hat{\mu}_{2a} + N_{ab}\left(p\hat{\mu}_{2ab} + q\hat{\mu}_{2ba} \right) + N_b \hat{\mu}_{2b} \right] \tag{2.3}$$

such that $p + q = 1$

It can be seen that

$$EE_m \left(\hat{\bar{y}}_2 - \bar{y}_2 \right) = 0$$

where, E refers to unconditional design-based expectation.

The variance of $\left(\hat{\bar{y}}_2 - \bar{y}_2\right)$ (Cassel *et al.* [2]) can be obtained using

$$V\left(\hat{\bar{y}}_2 - \bar{y}_2\right) = EV_m\left(\hat{\bar{y}}_2 - \bar{y}_2\right) + VE_m\left(\hat{\bar{y}}_2 - \bar{y}_2\right)$$
$$= EV_m\left(\hat{\bar{y}}_2 - \bar{y}_2\right)$$

$V\left(\hat{\bar{y}}_2 - \bar{y}_2\right)$ for large 'N' can be shown equal to

$$\sigma^2\left[\frac{p_A^2}{n_A}\left\{(1-\alpha)\left(\frac{1-\phi_A\rho_a^2}{1-\phi_A^2\rho_a^2}\right) + p^2\alpha\left(\frac{1-\phi_A\rho_{ab}^2}{1-\phi_A^2\rho_{ab}^2}\right)\right\}\right.$$
$$\left.+\frac{p_B^2}{n_B}\left\{(1-\lambda)\left(\frac{1-\phi_B\rho_b^2}{1-\phi_B^2\rho_b^2}\right) + q^2\lambda\left(\frac{1-\phi_B\rho_{ab}^2}{1-\phi_B^2\rho_{ab}^2}\right)\right\}\right]$$

in which

$$\phi_A = \frac{u_A}{n_A};\ \phi_B = \frac{u_B}{n_B};\ p_A = \frac{N_A}{N};\ p_B = \frac{N_B}{N};\ \alpha = \frac{N_{ab}}{N_A};\ \lambda = \frac{N_{ab}}{N_B}$$

3. Optimization of Sample Sizes and Proportions

We consider the cost function

$$C = 2c'_A u_A + \left(c'_A + c'_{1A}\right)m_A + 2c'_B u_B + \left(c'_B + c'_{1B}\right)m_B$$

where c'_A and c'_B are the per unit costs of collecting information from frames A and B respectively, whereas c'_{1A} and c'_{1B} are the per unit costs of collecting information from frames A and B for the matched portion of the sample on the second occasion. Obviously,

$$\left(c'_{1A}, c'_{1B}\right) < \left(c'_A, c'_B\right)$$

C = total cost of survey operation.

For the sake of simplicity we assume that

$$c'_{1A} = c'_A \text{ and } c'_{1B} = c'_B$$

Then, the cost function reduces to

$$C = c_A n_A + c_B n_B \tag{3.1}$$

where

$$c_A = 2c'_A \text{ and } c_B = 2c'_B$$

Then $V(\hat{T} - T)$ may be minimised subject to the total cost C. For simplicity we consider 100% coverage by the frame A on both the occasions. For this case the variance expression reduces to

$$V(\hat{\bar{y}}_2 - \bar{y}_2) = \sigma^2 \left[\frac{p_A^2}{n_A}(1-\delta)\left(\frac{1-\phi_A \rho_a^2}{1-\phi_A^2 \rho_a^2}\right) + p^2 \frac{p_A^2}{n_A} \delta \left(\frac{1-\phi_A \rho_{ab}^2}{1-\phi_A^2 \rho_{ab}^2}\right) + q^2 \frac{p_B^2}{n_B}\left(\frac{1-\phi_B \rho_{ab}^2}{1-\phi_B^2 \rho_{ab}^2}\right)\right] \quad (3.2)$$

where, $\delta = \dfrac{N_B}{N_A}$

Minimising $V(\hat{\bar{y}}_2 - \bar{y}_2)$ subject to the cost function defined in (3.1), we obtain

$$p_0^2 = \frac{(1-\delta)K_1 K_3}{K_2(K_2\rho - \delta K_3)}$$

$$\phi_B = \frac{1}{1+\left(1-\rho_{ab}^2\right)^{1/2}}$$

$$n_A^2 = \frac{p_A^2}{\gamma c_A}\left[(1-\delta)K_1 + p^2 \delta K_2\right];\ n_B^2 = \frac{p_B^2}{\gamma c_B} q^2 K_3$$

$$\frac{(1-\delta)\rho_a^2\left[\phi_A^2 \rho_a^2 - 2\phi_A + 1\right]}{\left[1-\phi_A^2 \rho_a^2\right]^2} + p^2 \delta \rho_{ab}^2 \frac{\left[\phi_A^2 \rho_a^2 - 2\phi_A + 1\right]}{\left[1-\phi_A^2 \rho_{ab}^2\right]} = 0$$

where

$$K_1 = \left(\frac{1-\phi_A \rho_a^2}{1-\phi_A^2 \rho_a^2}\right);\ K_2 = \left(\frac{1-\phi_A \rho_{ab}^2}{1-\phi_A^2 \rho_{ab}^2}\right);\ K_3 = \left(\frac{1-\phi_B \rho_{ab}^2}{1-\phi_B^2 \rho_{ab}^2}\right)\ \rho = \frac{c_A}{c_B}$$

and γ is the Lagrange multiplier.

The resulting optimum variance is

$$V\left(\hat{\bar{y}}_2-\bar{y}_2\right)=\frac{1}{C}\left[p_A\, c_A^{1/2}\left\{(1-\delta)K_1+p^2\,\delta K_2\right\}+qp_B K_3^{1/2}c_B^{1/2}\right]^2 \tag{3.3}$$

To examine the gain if any due to use of multiple frame instead of a single frame we consider a predictor based on a sample of size n'_A from A frame which is assumed to consist of two poststrata of sizes N_a and N_{ab}.

The model for the sampled data now reduces to

$$\underset{\sim}{y_1}=\underset{\sim}{X_1}\,\underset{\sim}{\beta_1}+\underset{\sim}{\varepsilon_1};E_m\left(\underset{\sim}{\varepsilon_1}\right)=0;\,E_m\left(\underset{\sim}{\varepsilon_1}\,\underset{\sim}{\varepsilon'_1}\right)=\sigma^2\,\underset{\sim}{\Sigma_1}$$

where $\underset{\sim}{y_1}$ is a $2n'_A\times 1$ vector of observations on the study variable, $\underset{\sim}{X_1}$ is $2n'_A\times 4$ matrix having the elements 0 and 1.

$\underset{\sim}{\beta_1}$ is 4×1 vector of parameters having the structure

$$\underset{\sim}{\beta'_1}=\left[\mu'_{1a}\ \mu'_{1ab}\ \mu'_{2a}\ \mu'_{2ab}\right]'$$

$\underset{\sim}{\varepsilon_1}$ is $2n'_A\times 1$ vector of error terms.

$$\underset{\sim}{\Sigma_1}=\begin{bmatrix}\underset{\sim}{\Sigma_{111}} & \underset{\sim}{\Sigma_{112}}\\ \underset{\sim}{\Sigma'_{112}} & \underset{\sim}{\Sigma_{122}}\end{bmatrix}$$

where $\underset{\sim}{\Sigma_{111}}$ and $\underset{\sim}{\Sigma_{122}}$ are identity matrices of order $n'_A\times n'_A$.

The matrix $\underset{\sim}{\Sigma_{112}}$ is of order $n'_A\times n'_A$ and is given by

$$\underset{\sim}{\Sigma_{112}}=\begin{bmatrix}\underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0}\\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0}\\ \rho_a\underset{\sim}{I_{ma}} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0}\\ \underset{\sim}{0} & \rho_{ab}\underset{\sim}{I_{mab}} & \underset{\sim}{0} & \underset{\sim}{0}\end{bmatrix}$$

The vector $\underset{-}{\beta}_1$ is estimated as

$$\hat{\underset{-}{\beta}}_1 = \left(\underset{\sim}{x}_1' \, \underset{\sim}{\Sigma}_1^{-1} \, \underset{\sim}{x}_1 \right)^{-1} \underset{\sim}{x}_1' \, \underset{\sim}{\Sigma}_1^{-1} \, \underset{\sim}{y}_1 \tag{3.4}$$

Consider the predictor

$$\hat{\bar{y}}_2' = \frac{1}{N} \left[N_a \, \hat{\mu}_{2a}' + N_{ab} \, \hat{\mu}_{2ab}' \right] \tag{3.5}$$

where $\hat{\mu}_{2a}'$ and $\hat{\mu}_{2ab}'$, etc. can be obtained on the similar lines as described in (2.1) and (2.2).

The variance of $\left(\hat{\bar{y}}_2' - \bar{y}_2 \right)$ can be seen to be equal to

$$V\left(\hat{\bar{y}}_2' - \bar{y}_2 \right) = \sigma^2 \frac{p_A^2}{n_A} \left[(1-\delta) K_1 + K_2 \, \delta \right] \tag{3.6}$$

Assuming the total cost to be same as in multiple frame situation, we consider the cost function

$$C = c_A n_A'$$

The optimum variance in this case is

$$V\left(\hat{\bar{y}}_2' - \bar{y}_2 \right) = \sigma^2 \frac{p_A^2}{C} c_A \left[(1-\delta) K_1 + K_2 \, \delta \right] \tag{3.7}$$

A case of particular interest is described in Section 4.

4. Single Frame on the First Occasion and Multiple Frame on the Second Occasion

In this case a sample of size n_A is drawn on the first occasion from A frame. Random sub-sample of m_A units are retained for use on the second occasion. On the second occasion we assume that out of m_A units on the first occasion m_a units are common in the (a) domain and $m_A - m_a = m_{ab}$ units fall in the (ab) domain. Independent samples of sizes u_A and n_B are drawn from A and B frames respectively on the second occasion. In this case $V\left(\hat{\bar{y}}_2 - \bar{y}_2 \right)$ and $V\left(\hat{\bar{y}}_2' - \bar{y}_2 \right)$ reduce to

$$V\left(\hat{\bar{y}}_2 - \bar{y}_2 \right) = \sigma^2 \left[\frac{p_A^2}{n_A} \left\{ (1-\delta)\theta_1 + p^2 \delta \right\} + \frac{p_B^2}{n_B} q^2 \right] \tag{4.1}$$

$$V\left(\hat{\bar{y}}'_2 - \bar{y}_2\right) = \sigma^2\left[\frac{p_A^2}{n_A}\{(1-\delta)\theta_1 + p^2\delta\}\right] \tag{4.2}$$

where,

$$\theta_1 = \frac{\left(1-\rho_a^2\right)+(1-\delta)(1-\phi_A)\rho_a^2}{\left(1-\phi_A\,\rho_a^2\right)+\rho_a^2\,(1-\delta)\phi_A\,(1-\phi_A)} \tag{4.3}$$

For determination of optimum values we assume that on the first occasion information is available on the variable of interest. Thus, the cost function in this case reduces to

$$C_0 = c'_A\, n_A + c'_B\, n_B$$

Minimization of V ( ) subject to this cost function, we obtain

$$n_{Aopt} = C_0\,\frac{\left\{(1-\delta)\theta_{1opt} + p_0^2\,\delta\right\}^{1/2}}{(c'_A)^{1/2}\left[(c'_A)^{1/2}\left\{(1-\delta)\theta_{1opt} + p_0^2\,\delta\right\}^{1/2} + (c'_B)^{1/2}\,\delta q\right]}$$

$$n_{Bopt} = C_0\,\frac{\delta q}{(c'_B)^{1/2}\left[(c'_A)^{1/2}\left\{(1-\delta)\theta_{1opt} + p_0^2\,\delta\right\}^{1/2} + (c'_B)^{1/2}\,\delta q\right]}$$

$$p_0^2 = \frac{(1-\delta)\theta_{1opt}}{(\rho'-\delta)};\quad \rho' = \frac{c'_A}{c'_B};\quad \phi_{Aopt} = \frac{L_3}{\left[L_2 + \left(L_2^2 - L_1 L_3\right)^{1/2}\right]}$$

where

$$L_1 = P_1^2;\; L_2 = P_1\left[P_1 + 1 - \rho_a^2\right]$$

$$L_3 = P_1\left[2 + P_1 - 2\rho_a^2\right] - \rho_a^2\left(1-\rho_a^2\right);\text{ and } P_1 = \rho_a^2\,(1-\delta)$$

It can be seen that

(i) ϕ_A reduces to usual successive sampling formula for $\delta = 0$

(ii) $\phi_A = 1$ for $\rho_a = 1$

(iii) $\phi_A > 0$ provided $1 + \rho_a^2\delta^2 > 2\delta$

Table 1. Values of optimum sample sizes, proportions and reduction in variance for current estimate involving multiple frame

		δ															
		0.1				0.3				0.5				0.6			
ρ_A	CR ρ'^{-1}	ϕ_A	n_A	n_B	VR	ϕ_A	n_A	n_B	VR	ϕ_A	n_A	n_B	VR	ϕ_A	n_A	n_B	VR
0.1	0.1	0.446	97.731	22.686	0.952	0.288	92.432	75.679	0.844	0.003	85.58	144.2	0.717	–	–	–	–
	0.5	0.446	97.776	4.449	0.991	0.288	92.236	15.528	0.967	0.003	84.452	31.095	0.932	–	–	–	–
	1	0.446	99.986	0.014	1	0.288	99.946	0.054	1	0.003	99.875	0.125	1	–	–	–	–
0.3	0.1	0.458	97.695	23.054	0.95	0.303	92.29	77.099	0.839	0.024	85.25	147.5	0.708	–	–	–	–
	0.5	0.458	97.695	4.61	0.99	0.303	91.94	16.119	0.964	0.024	83.815	32.37	0.927	–	–	–	–
	1	0.458	99.87	0.13	1	0.303	99.502	0.498	1	0.024	98.848	1.152	1	–	–	–	–
0.6	0.1	0.506	97.546	24.537	0.944	0.365	91.725	82.748	0.819	0.111	83.967	160.33	0.672	–	–	–	–
	0.5	0.506	97.371	5.258	0.987	0.365	90.775	18.45	0.954	0.111	81.386	37.228	0.906	–	–	–	–
	1	0.506	99.404	0.596	1	0.365	97.758	2.242	0.999	0.111	94.987	5.013	0.997	–	–	–	–
0.8	0.1	0.583	97.324	26.755	0.934	0.464	90.901	90.99	0.789	0.25	82.174	178.26	0.624	0.063	76.291	237.09	0.532
	0.5	0.583	96.888	6.244	0.982	0.464	89.101	21.798	0.937	0.25	78.105	43.79	0.874	0.063	70.649	58.703	0.833
	1	0.583	98.713	1.287	0.999	0.464	95.284	4.716	0.995	0.25	89.898	10.102	0.99	0.063	85.848	14.152	0.987
0.9	0.1	0.663	97.111	28.892	0.924	0.566	90.128	98.724	0.76	0.393	80.564	194.36	0.582	0.241	74.101	258.99	0.488
	0.5	0.663	96.425	7.149	0.976	0.566	87.56	24.881	0.919	0.393	75.265	49.47	0.845	0.241	67.106	65.788	0.789
	1	0.663	98.053	1.947	0.997	0.566	93.036	6.964	0.989	0.393	85.605	14.395	0.98	0.241	80.338	19.662	0.975

– indicates infeasible values CR : Cost ratio

The optimum value of θ_1 (θ_{1opt}) can be obtained by substituting the value of ϕ_{Aopt} in (4.3).

The expressions for optimum variances are given as

$$V_{opt.}\left(\hat{\bar{y}}_2 - \bar{y}_2\right) = \sigma^2 \frac{p_A^2\left[\left(c'_A\right)^{1/2}\left\{(1-\delta)\theta_{1opt} + p_0^2\delta\right\}^{1/2} + \left(c'_B\right)^{1/2}\delta q\right]^2}{C_0}$$

$$V_{opt}\left(\hat{\bar{y}}'_2 - \bar{y}_2\right) = \sigma^2\left[\frac{p_A^2}{n'_{Aopt}}\left\{(1-\delta)\theta_{1opt} + \delta\right\}\right]$$

where, $\quad n'_{Aopt} = \dfrac{C_0}{c'_A}$

We denote by VR the ratio of variances of multiple frame predictor and the post-stratified predictor.

Thus,

$$VR = \frac{V_{opt}\left(\hat{\bar{y}}_2 - \bar{y}_2\right)}{V_{opt}\left(\hat{\bar{y}}'_2 - \bar{y}_2\right)}$$

Optimum values of n_A, n_B, ϕ_A and the ratio VR have been computed for different combinations of ρ_a, δ, and ρ'. The results are presented in Table 1.

It can be seen that the ratio VR decreases as both $\rho'$and δ increase. Also ϕ_A increases with increase in ρ_a but decreases with increase in δ.

Comparison between the multiple frame and post-stratified estimator for the current occasion when multiple frames are available on the first occasion and single frame on the second occasion has not been made. In this case both the predictors are equally precise. This is due to the fact that in this case no information is available on the second occasion from the B frame.

ACKNOWLEDGEMENTS

The authors are grateful to the referee's for improving the quality of the paper. The authors express sincere thanks to Shri D.P. Sharma, Technical Officer, for the preparation of this paper.

REFERENCES

[1] Armstrong, B. (1979). Test of multiple frame sampling techniques for agricultural surveys. *Sur. Meth.*, 178-199.

[2] Cassel, C.M., Sarndal, C.E. and Wretman, J.H. (1977). *Foundations of Inference in Survey Sampling*. John Wiley and Sons, New York.

[3] Hartley, H.O. (1965). Multiple frame surveys. *Proc. Soc. Statist. Sec. Amer. Statist. Assoc.*, Minneapolis, Minnestos, 203-208.

[4] Hartley, H.O. (1974). Multiple frame methodology and selected applications. *Sankhya*, **37**(**C**), 99-118.

[5] Patterson, H.D. (1950). Sampling on successive occasion with partial replacement of units. *J. Roy. Statist. Soc.*, **12**(**B**), 241-255.

[6] Rao, J.N.K. (1968). Some non-response sampling theory when the frame contains an unknown amount of duplication. *J. Amer. Statist. Assoc.*, **63**, 321.

[7] Saxena, B.C., Narain, P. and Srivastava, A.K. (1984). Multiple frame surveys in two-stage sampling. *Sankhya*, **46** (**B**), 75-82.

[8] Saxena, B.C., Narain, P. and Srivastava, A.K. (1985). Estimation of total for two characters in multiple frame surveys. *Sur. Meth.*, **11**(**2**), 119-132.

[9] Serrurier, D. and Phillips, J.E. (1976). Double frame Ontario pilot hog surveys. *Sur. Meth.*, **2**(**2**).

[10] Skinner, C.N. and Rao, J.N.K. (1996). Estimation in dual frame surveys with complex designs. *J. Amer. Statist. Assoc.*, **91**, 349-356.

[11] Vogal, F. A. (1975). Surveys with overlapping frames - problems in applications. *135th Annual Meeting of American Statist. Assoc.*, Atlanta, Georgia.