Journal of Reliability and Statistical Studies (ISSN: 0974-8024)
Vol. 2, Issue 1(2009): 28 - 40

UTILIZATION OF NON-RESPONSE AUXILIARY POPULATION MEAN IN IMPUTATION FOR MISSING OBSERVATIONS

D. Shukla [1], **Narendra Singh Thakur** [2] **and Dharmendra Singh Thakur** [3]
Deptt. of Maths and Stats., Dr. Harisingh Gour University, Sagar (M.P.), India.
E mail: 1. diwakarshukla@rediffmail.com, 2. nst_stats@yahoo.co.in

Abstract

This paper presents an imputation based factor-type class of estimation strategy for estimating population mean in presence of missing values of auxiliary variable. The non-sampled part of population is used as an imputation technique in the form of a proposed class of estimators. The bias and mean squared error of this class is obtained. Some special cases are discussed. A specific range of parameter is found where the proposed class is optimal. The efficiency of the proposed estimator is compared with similar non-imputed estimator and it is found useful under missing observations setup.

Keywords: Imputation, Non-response, Post-stratification, Simple Random Sampling Without Replacement (SRSWOR), Respondents (R).

1. Introduction

To estimate the population mean using auxiliary variable, many estimators are available in literature like-ratio, product, regression, dual-to-ratio estimator and so on. If some values of auxiliary variable are missing, none of the above estimators can be used. In sampling theory, the problem of mean estimation of a population is considered by many authors like Singh (1986), Singh and Singh (1991), Singh et al. (1994), Singh and Singh (2001). Sometimes, in survey situations, a small part of sample remains non-responded (or incomplete) due to many practical reasons. Techniques and estimation procedures are needed to develop for this purpose. The imputation is a well defined methodology by virtue of which this kind of problem could be partially solved. Ahmed et al. (2006), Rao and Sitter (1995), Rubin (1976) and Singh and Horn (2000) have given applications of various imputation procedures. Hinde and Chambers (1990) studied the non-response imputation with multiple sources of non-response. The non-response in sample surveys immensely looked into by Hansen and Hurwitz (1946), Lessler and Kalsbeek (1992), Khot (1994), Grover and Couper (1998) etc.

When the population is divided into two groups namely "response" and "non-response" then the procedure is known as post-stratification. Estimation problem in sample surveys, in the setup of post-stratification, under non-response situation is studied due to Shukla and Dubey (2001, 2004, and 2006). Some other useful contributions to this area are due to Smith (1991), Agrawal and Panda (1993), Shukla and Trivedi (1999, 2001, 2006), Wywial (2001) and Shukla et al. (2002, 2006). When a sample is full of response over study variable but some of auxiliary values are missing, it is hard to utilize the usual existing estimators. Traditionally, it is essential to estimate those missing observations first by some specific estimation techniques. One can think of utilizing the non-sampled part of the population in order to get estimates of missing observations in the sample. These estimates could be imputed into actual estimation procedures used for estimating the population mean. The content of this paper takes into account the similar aspect for non-responding values of the sample assuming post-stratified setup and utilizing the auxiliary source of data.

1.1 Symbols and Setup

Let $U = (U_1, U_2, ..., U_N)$ be a finite population of N units with Y as a study variable and X an auxiliary variable. The population has two types of individuals like N_1 as number of "respondents (R)" and N_2 "non-respondents (NR)", ($N = N_1+N_2$). Their population proportions are expressed like $W_1 = N_1/N$ and $W_2 = N_2/N$. Further, let $\bar{Y}$ and $\bar{X}$ be the population means of Y and X respectively. The following notations are used in this paper:

R-group	Respondents group (group of those who respond during survey.	NR-group	Non-respondents group or group of those who do not respond during survey.
$\bar{Y}_1$	Population mean of R-group of Y.	$\bar{Y}_2$	Population mean of NR-group of Y.
$\bar{X}_1$	Population mean of R-group of X.	$\bar{X}_2$	Population mean of NR-group of X.
S_{1Y}^2	Population mean square of R-group of Y.	S_{2Y}^2	Population mean square of NR-group of Y.
S_{1X}^2	Population mean square of R-group of X.	S_{2X}^2	Population mean square of NR-group of X.
C_{1Y}	Coefficient of Variation of Y in R-group.	C_{2Y}	Coefficient of Variation of Y in NR-group.
C_{1X}	Coefficient of Variation of X in R-group.	C_{2X}	Coefficient of Variation of X in NR-group.
ρ	Correlation Coefficient in population between X and Y.	n	Sample size from population of size N by SRSWOR.
n_1	Post-stratified sample size coming from R-group.	n_2	Post-stratified sample size from NR-group.
$\bar{y}_1$	Sample mean of Y based on n_1 observations of R-group.	$\bar{y}_2$	Sample mean of Y based on n_2 observations of NR-group.
$\bar{x}_1$	Sample mean of X based on n_1 observations of R-group.	$\bar{x}_2$	Sample mean of X based on n_2 observations of NR-group.
ρ_1	Correlation Coefficient between study variable Y and auxiliary variable X for R-group.	ρ_2	Correlation Coefficient between study variable Y and auxiliary variable X for NR-group.

Further, consider few more symbolic representations:

$$D_1 = E\left(\frac{1}{n_1}\right) = \left[\frac{1}{nW_1} + \frac{(N-n)(1-W_1)}{(N-1)n^2W_1^2}\right]; \quad D_2 = E\left(\frac{1}{n_2}\right) = \left[\frac{1}{nW_2} + \frac{(N-n)(1-W_2)}{(N-1)n^2W_2^2}\right] \tag{1.1}$$

$$\bar{Y} = \frac{N_1\bar{Y}_1 + N_2\bar{Y}_2}{N}; \quad \bar{X} = \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N} \tag{1.2}$$

2. Assumptions

The following assumptions are made before formulating an imputation based estimation procedure :

1. The values of N and n are known. Also, N_1 and N_2 are known by past data, past experience or guess of the investigator ($N_1 + N_2 = N$).
2. Other population parameters are assumed known, in either exact or in ratio form except the $\bar{Y}$, $\bar{Y}_1$ and $\bar{Y}_2$.
3. The population means $\bar{X}_1$ and $\bar{X}_2$ are known.
4. The sample of size n is drawn by SRSWOR and post-stratified into two groups of size n_1 and n_2 ($n_1 + n_2 = n$) according to R and NR group respectively.
5. The information about Y variable in sample is completely available.
6. The sample means $\bar{y}_1$ and $\bar{y}_2$ of both groups are known such that

   $\bar{y} = \dfrac{n_1\bar{y}_1 + n_2\bar{y}_2}{n}$ which is the sample mean on n units.
7. The sample mean $\bar{x}_1$ of auxiliary variable for R-group is known, but the information about $\bar{x}_2$ of NR-group is missing. Therefore, the value of $\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n}$ can not be obtained due to absence of $\bar{x}_2$.

3. Proposed Class of Estimation Strategy

To estimate population mean $\bar{Y}$ the usual ratio, product and regression estimators are not applicable when observations related to $\bar{x}_2$ are missing. Singh and Shukla (1987) have proposed a factor type estimator for estimating population mean $\bar{Y}$. Shukla et al. (1991), Singh and Shukla (1993), Shukla (2002) have also discussed properties of factor-type estimators applicable for estimating population mean under SRSWOR and Two-Phase Sampling. But all these cannot be useful due to unknown information $\bar{x}_2$. In order to solve this, an imputation $(\bar{x}_2^*)$ is adopted as:

$$\bar{x}_2^* = \left[\frac{N\bar{X} - n\bar{X}_2}{N-n}\right] \tag{3.1}$$

The logic for this imputation is to utilize the non-sampled part of the population of X for obtaining an estimate of missing $\bar{x}_2$ and generate $\bar{x}^*$ as describe below :

$$\bar{x}^* = \left[\frac{N_1\bar{x}_1 + N_2\bar{x}_2^*}{N_1 + N_2}\right] \tag{3.2}$$

A proposed and class of imputed factor-type estimation strategy for estimating $\bar{Y}$ is:

$$(\bar{y}_{FT})_k = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right)\left[\frac{(A+C)\bar{X} + f\,B\bar{x}^*}{(A+f\,B)\bar{X} + C\bar{x}^*}\right] \tag{3.3}$$

where $0 < k < \infty$ and k is a constant,

$A = (k-1)(k-2)$; $B = (k-1)(k-4)$; $C = (k-2)(k-3)(k-4)$; $f = n/N$

4. Large Sample Approximation

Consider the following for large n:

$$\bar{y}_1 = \bar{Y}_1(1+e_1); \quad \bar{y}_2 = \bar{Y}_2(1+e_2); \quad \bar{x}_1 = \bar{X}_1(1+e_3); \quad \bar{x}_2 = \bar{X}_2(1+e_4) \tag{4.1}$$

where, e_1, e_2, e_3 and e_4 are very small numbers and $|e_i| < 1$ (i = 1,2,3,4).

Using the basic concept of SRSWOR and the concept of post-stratification of the sample n into n_1 and n_2 [see Cochran (2005), Sukhatme et al. (1984)],we get

$$\left.\begin{array}{ll} \mathrm{E}(e_1)=\mathrm{E}\left[\mathrm{E}(e_1)\,|n_1\right]=0; & \mathrm{E}(e_2)=\mathrm{E}\left[\mathrm{E}(e_2)\,|n_2\right]=0 \\ \mathrm{E}(e_3)=\mathrm{E}\left[\mathrm{E}(e_3)\,|n_1\right]=0; & \mathrm{E}(e_4)=\mathrm{E}\left[\mathrm{E}(e_4)\,|n_2\right]=0 \end{array}\right\} \quad (4.2)$$

Assuming the independence of R-group and NR-group representation in the sample, the following expression could be obtained:

$$\mathrm{E}\left[e_1^2\right]=\mathrm{E}\left[\mathrm{E}(e_1^2)\,|n_1\right] \;=\mathrm{E}\left[\left\{\left(\frac{1}{n_1}-\frac{1}{N}\right)C_{1Y}^2\right\}\Big|n_1\right]$$

$$=\left[\left\{E\left(\frac{1}{n_1}\right)-\frac{1}{N}\right\}C_{1Y}^2\right]=\left[\left(D_1-\frac{1}{N}\right)C_{1Y}^2\right] \quad (4.3)$$

$$\mathrm{E}\left[e_2^2\right]=\mathrm{E}\left[E(e_2^2)\,|n_2\right] \;=\left[\left(D_2-\frac{1}{N}\right)C_{2Y}^2\right] \quad (4.4)$$

$$\mathrm{E}\left[e_3^2\right]=\mathrm{E}\left[\mathrm{E}(e_3^2)\,|n_1\right] \;=\left[\left(D_1-\frac{1}{N}\right)C_{1X}^2\right] \quad (4.5)$$

and

$$\mathrm{E}\left[e_4^2\right]=\left[\left(D_2-\frac{1}{N}\right)C_{2X}^2\right] \quad (4.6)$$

$$\mathrm{E}\left[e_1e_3\right]=\mathrm{E}\left[\mathrm{E}(e_1e_3)\,|n_1\right] \;=\mathrm{E}\left[\left\{\left(\frac{1}{n_1}-\frac{1}{N}\right)\rho_1C_{1Y}C_{1X}\right\}\Big|n_1\right]$$

$$=\left[\left(D_1-\frac{1}{N}\right)\rho_1C_{1Y}C_{1X}\right] \quad (4.7)$$

$$\mathrm{E}\left[e_1e_2\right]=\mathrm{E}\left[\mathrm{E}(e_1e_2)\,|n_1,n_2\right]\;=0 \quad (4.8)$$

$$\mathrm{E}\left[e_1e_4\right]=0 \quad (4.9)$$

$$\mathrm{E}\left[e_2e_3\right]=\mathrm{E}\left[\mathrm{E}(e_2e_3)\,|n_1,n_2\right]\;=0 \quad (4.10)$$

$$\mathrm{E}\left[e_2e_4\right]=\left(D_2-\frac{1}{N}\right)\rho_2C_{2Y}C_{2X} \quad (4.11)$$

$$\mathrm{E}\left[e_3e_4\right]=0 \quad (4.12)$$

The expressions (4.8), (4.9), (4.10) and (4.12) are true under the assumption of independent representation of R-group and NR-group units in the sample. This is introduced to simplify mathematical expressions.

Theorem 4.1: The estimator $\left(\bar{y}_{FT}\right)_k$ could be expressed under large sample approximations in following form:

$$\left(\bar{y}_{FT}\right)_k=\delta\,\bar{Y}\,[1+s_1W_1e_1+s_2W_2e_2][1+(\alpha-\beta)e_3-(\alpha-\beta)\beta\, e_3^2+(\alpha-\beta)\,\beta^2e_3^3-(\alpha-\beta)\,\beta^3e_3^4+...] \quad (4.13)$$

Proof: Rewrite $\bar{x}^*$ as in (3.2):

$$\bar{x}^*=\left[\frac{N_1\bar{x}_1+N_2(\bar{x}_2^*)}{N_1+N_2}\right] \qquad \text{where} \qquad \bar{x}_2^*=\left[\frac{N\bar{X}-n\bar{X}_2}{N-n}\right]$$

$$\Rightarrow \bar{x}^*=\frac{1}{N}\left[N_1\bar{x}_1+N_2\left\{\frac{N\bar{X}-n\bar{X}_2}{N-n}\right\}\right]=\bar{X}\left[W_1r_1+p(1-fr_2)+W_1r_1e_3\right]=\bar{X}[v+W_1r_1e_3] \quad (4.14)$$

where, $p = \frac{N_2}{N-n}$; $r_1 = \frac{\overline{X}_1}{\overline{X}}$; $r_2 = \frac{\overline{X}_2}{\overline{X}}$; $W_1 = \frac{N_1}{N}$; $f = \frac{n}{N}$; $v = W_1 r_1 + p(1 - f r_2)$.

Now, the estimator $\left(\overline{y}_{FT}\right)_k$ under large sample approximations (4.1) and using (4.12) will be

$$\left(\overline{y}_{FT}\right)_k = \left(\frac{N_1 \overline{y}_1 + N_2 \overline{y}_2}{N}\right)\left[\frac{(A+C)\overline{X} + f B \overline{x}^*}{(A + f B)\overline{X} + C\overline{x}^*}\right]$$

$$= \overline{Y}\left[1 + s_1 W_1 e_1 + s_2 W_2 e_2\right]\left[\frac{(A + fBv + C) + fBW_1 r_1 e_3}{(A + fB + Cv) + CW_1 r_1 e_3}\right]$$

$$= \overline{Y}\left[1 + s_1 W_1 e_1 + s_2 W_2 e_2\right]\left[\frac{\psi_1 + \psi_2 e_3}{\psi_3 + \psi_4 e_3}\right]$$

$$= \delta_2 \overline{Y}\left[1 + s_1 W_1 e_1 + s_2 W_2 e_2\right](1 + \alpha e_3)(1 + \beta e_3)^{-1} \tag{4.15}$$

where, $\psi_1 = A + fBv + C$; $\psi_2 = fBW_1 r_1$; $\psi_3 = A + fB + Cv$; $\psi_4 = C r_1 W_1$;

$s_1 = \frac{\overline{Y}_1}{\overline{Y}}$; $s_2 = \frac{\overline{Y}_2}{\overline{Y}}$; $\alpha = \frac{\psi_2}{\psi_1}$; $\beta = \frac{\psi_4}{\psi_3}$; $\delta = \frac{\psi_1}{\psi_3}$.

We can further express (4.15) as:

$$\left(\overline{y}_{FT}\right)_k = \delta\ \overline{Y}\ [1 + s_1 W_1 e_1 + s_2 W_2 e_2]\ [1 + (\alpha - \beta)e_3 - (\alpha - \beta)\ \beta\ e_3^2 + (\alpha - \beta)\ \beta^2\ e_3^3 - \ldots\ldots] \tag{4.16}$$

5. Bias and Mean Squared Error

Using E(.) for expectation, B(.) for bias and M(.) for mean squared error, we have to the first order of approximations for i, j = 1, 2, 3, ...

$$E\left[e_1^i e_2^j\right] = E\left[e_1^i e_3^j\right] = E\left[e_2^i e_3^j\right] = 0 \quad \text{when} \quad i + j > 2 \tag{5.1}$$

Theorem 5.1: To the first order of approximations, the bias of the estimator $\left(\overline{y}_{FT}\right)_k$ of $\overline{Y}$ is

$$B\left(\overline{y}_{FT}\right)_k = \overline{Y}\left[(\delta - 1) - \delta\ C_{1X}(\alpha - \beta)\left(D_1 - \frac{1}{N}\right)\left\{\beta\ C_{1X} - s_1 W_1 \rho_1 C_{1Y}\right\}\right] \tag{5.2}$$

Proof: $B\left(\overline{y}_{FT}\right)_k = E\left[\left(\overline{y}_{FT}\right)_k - \overline{Y}\right]$

Taking expectations in (4.16), we have

$$E\left(\overline{y}_{FT}\right)_k = \delta \overline{Y}\ E\ [1 + s_1 W_1 e_1 + s_2 W_2 e_2]\ [1 + (\alpha - \beta)e_3 - (\alpha - \beta)\ \beta\ e_3^2 + (\alpha - \beta)\ \beta^2\ e_3^3\ \ldots\ldots]$$

$$= \delta\ \overline{Y}\left[1 - (\alpha - \beta)\beta\left(D_1 - \frac{1}{N}\right)C_{1X}^2 + (\alpha - \beta)s_1 W_1\left(D_1 - \frac{1}{N}\right)\rho_1 C_{1Y} C_{1X}\right]$$

$$= \delta\ \overline{Y}\left[1 - (\alpha - \beta)\left(D_1 - \frac{1}{N}\right)C_{1X}\left\{\beta\ C_{1X} - s_1 W_1 \rho_1 C_{1Y}\right\}\right]$$

Therefore, $B\left(\overline{y}_{FT}\right)_k = \overline{Y}\left[(\delta - 1) - \delta\ C_{1X}(\alpha - \beta)\left(D_1 - \frac{1}{N}\right)\left\{\beta\ C_{1X} - s_1 W_1 \rho_1 C_{1Y}\right\}\right]$

Theorem 5.2: The mean squared error of $\left(\overline{y}_{FT}\right)_k$ is

$$M\left(\overline{y}_{FT}\right)_k = \overline{Y}^2\left[(\delta - 1)^2 + \left(D_1 - \frac{1}{N}\right)\left\{K_1 s_1^2 C_{1Y}^2 + K_2 C_{1X}^2 + 2K_3 s_1 \rho_1 C_{1Y} C_{1X}\right\} + \left(D_2 - \frac{1}{N}\right)\delta^2 s_2^2 W_2^2 C_{2Y}^2\right] \tag{5.3}$$

where, $K_1 = \delta_2^2 W_1^2$; $K_2 = \delta(\alpha - \beta)\{\delta(\alpha - \beta) - 2(\delta - 1)\beta\}$; $K_3 = W_1 \delta(2\delta - 1)(\alpha - \beta)$

Proof : $M\left(\overline{y}_{FT}\right)_k = E\left[\left(\overline{y}_{FT}\right)_k - \overline{Y}\right]^2$

$$= E\left[\delta\ \overline{Y}\ \{1 + s_1 W_1 e_1 + s_2 W_2 e_2\}\{1 + (\alpha - \beta)e_3 - (\alpha - \beta)\beta e_3^2 + (\alpha - \beta)\beta^2 e_3^3 \ldots\} - \overline{Y}\right]^2$$

Using large sample approximations of (5.1), we have

$$M\left(\bar{y}_{FT}\right)_k = \bar{Y}^2 E\left[(\delta-1) + \delta\{(\alpha-\beta)e_3 - (\alpha-\beta)\beta e_3^2 + (s_1W_1e_1 + s_2W_2e_2) + (\alpha-\beta)(s_1W_1e_1 + s_2W_2e_2)e_3\}\right]^2$$

$$= \bar{Y}^2\left[(\delta-1)^2 + \delta^2\{(\alpha-\beta)^2 E\left(e_3^2\right) + s_1^2W_1^2E\left(e_1^2\right) + s_2^2W_2^2E\left(e_2^2\right) + 2(\alpha-\beta)s_1W_1E(e_1e_3)\} + 2\delta(\delta-1)\{-(\alpha-\beta)\beta E(e_3^2) + (\alpha-\beta)s_1W_1E(e_1e_3)\}\right]$$

[Using (4.2), (4.7) & (4.8)]

$$= \bar{Y}^2\left[(\delta-1)^2 + \left(D_1 - \frac{1}{N}\right)\{\delta^2 s_1^2W_1^2C_{1Y}^2 + \delta(\alpha-\beta)\{\delta(\alpha-\beta) - 2(\delta-1)\beta\}C_{1X}^2 + 2\delta(2\delta-1)(\alpha-\beta)s_1W_1\rho_1C_{1Y}C_{1X}\} + \left(D_2 - \frac{1}{N}\right)\delta^2 s_2^2W_2^2C_{2Y}^2\right]$$

$$= \bar{Y}^2\left[(\delta-1)^2 + \left(D_1 - \frac{1}{N}\right)\{K_1s_1^2C_{1Y}^2 + K_2C_{1X}^2 + 2K_3s_1\rho_1C_{1Y}C_{1X}\} + \left(D_2 - \frac{1}{N}\right)\delta^2 s_2^2W_2^2C_{2Y}^2\right]$$

6. Some Special Cases

The term A, B and C are functions of k. In particular, there are some special cases:

Case I : $k = 1 \Rightarrow A = 0; B = 0; C = -6; \quad \psi_1 = -6; \ \psi_2 = 0; \quad \psi_3 = -6v; \quad \psi_4 = -6r_1w_1$

$\alpha = 0; \ \beta = \frac{r_1W_1}{v}; \ \delta = \frac{1}{v}; \quad K_1 = \frac{W_1^2}{v^2}; \ K_2 = \frac{r_1^2W_1^2(3-2v)}{v^4}; \ K_3 = \frac{r_1W_1^2(v-2)}{v^3};$

The estimator $\left(\bar{y}_{FT}\right)_k$ along with bias and m.s.e. under case I is:

$$\left(\bar{y}_{FT}\right)_{k=1} = \left[\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right]\left[\frac{\bar{X}}{\bar{x}^*}\right] \tag{6.1}$$

$$B\left(\bar{y}_{FT}\right)_{k=1} = \bar{Y}v^{-3}\left[(1-v)v^2 + \left(D_1 - \frac{1}{N}\right)r_1W_1^2C_{1X}\{r_1C_{1X} - vs_1\rho_1C_{1Y}\}\right] \tag{6.2}$$

$$M\left(\bar{y}_{FT}\right)_{k=1} = \bar{Y}^2v^{-4}\left[(1-v)^2v^2 + W_1^2\left(D_1 - \frac{1}{N}\right)\{v^2s_1^2C_{1Y}^2 + (3-2v)r_1^2C_{1X}^2 + 2(v-2)vr_1s_1\rho_1C_{1Y}C_{1X}\} + \left(D_2 - \frac{1}{N}\right)v^2W_2^2s_2^2C_{2Y}^2\right] \tag{6.3}$$

Case II : $k = 2 \Rightarrow A = 0; B = -2; C = 0; \psi_1 = -2fv; \psi_2 = -2fW_1r_1; \psi_3 = -2f; \psi_4 = 0;$

$\alpha = r_1W_1v^{-1}; \ \beta = 0; \ \delta = v; \quad K_1 = W_1^2v^2; \ K_2 = r_1^2W_1^2; \ K_3 = r_1W_1^2(2v-1);$

$$\left(\bar{y}_{FT}\right)_{k=2} = \left[\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right]\left[\frac{\bar{x}^*}{\bar{X}}\right] \tag{6.4}$$

$$B\left(\bar{y}_{FT}\right)_{k=2} = \bar{Y}\left[(v-1) + \left(D_1 - \frac{1}{N}\right)W_1^2r_1s_1\rho_1C_{1Y}C_{1X}\right] \tag{6.5}$$

$$M\left(\bar{y}_{FT}\right)_{k=2} = \bar{Y}^2\left[(v-1)^2 + W_1^2\left(D_1 - \frac{1}{N}\right)\{v^2s_1^2C_{1Y}^2 + r_1^2C_{1X}^2 + 2(2v-1)s_1r_1\rho_1C_{1Y}C_{1X}\} + \left(D_2 - \frac{1}{N}\right)v^2W_2^2s_2^2C_{2Y}^2\right] \tag{6.6}$$

Case III : $k = 3 \Rightarrow A = 2; B = -2; \ C = 0; \quad \psi_1 = 2(1-fv); \ \psi_2 = -2fW_1r_1; \quad \psi_3 = 2(1-f);$ $\psi_4 = 0; \alpha = \frac{-fW_1r_1}{1-fv}; \ \beta = 0; \delta = \frac{1-fv}{1-f}; \quad K_1 = \frac{(1-fv)^2W_1^2}{(1-f)^2}; \quad K_2 = \frac{f^2W_1^2r_1^2}{(1-f)^2};$

$K_3 = \frac{-fW_1^2 r_1\{1+f(1-2v)\}}{(1-f)^2}$;

$$\left(\bar{y}_{FT}\right)_{k=3} = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right)\left[\frac{\bar{X} - f\bar{x}^*}{(1-f)\bar{X}}\right] \tag{6.7}$$

$$\mathrm{B}\left(\bar{y}_{FT}\right)_{k=3} = \bar{Y} f(1-f)^{-1}\left[(1-v) - \left(D_1 - \frac{1}{N}\right)W_1^2 r_1 s_1 \rho_1 C_{1Y} C_{1X}\right] \tag{6.8}$$

$$\mathrm{M}\left(\bar{y}_{FT}\right)_{k=3} = \bar{Y}^2(1-f)^{-2}\left[f^2(1-v)^2\left(D_1 - \frac{1}{N}\right)W_1^2\{(1-fv)^2 s_1^2 C_{1Y}^2 + f^2 r_1^2 C_{1X}^2 - 2\{1+f(1-2v)\}fr_1 s_1 \rho_1 C_{1Y} C_{1X}\} + \left(D_2 - \frac{1}{N}\right)(1-fv)^2 W_2^2 S_2^2 C_{2Y}^2\right] \tag{6.9}$$

Case IV : $k = 4 \Rightarrow A = 6; B = 0; C = 0; \psi_1 = 6; \psi_2 = 0; \psi_3 = 6; \psi_4 = 0; \alpha = 0; \beta = 0;$

$\delta = 1$; $K_1 = W_1^2$; $K_2 = 0$; $K_3 = 0$;

$$\left(\bar{y}_{FT}\right)_{k=4} = \left[\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right] \tag{6.10}$$

$$\mathrm{B}\left(\bar{y}_{FT}\right)_{k=4} = 0 \tag{6.11}$$

$$\mathrm{V}\left(\bar{y}_{FT}\right)_{k=4} = \left(D_1 - \frac{1}{N}\right)W_1^2\bar{Y}_1^2 C_{1Y}^2 + \left(D_2 - \frac{1}{N}\right)W_2^2\bar{Y}_2^2 C_{2Y}^2 \tag{6.12}$$

7. Estimator Without Imputation

Throughout the discussion, the value of $\bar{x}_2$ is assumed unknown. This is imputed by the term $\bar{x}_2^*$ to provide the generation of $\bar{x}^*$. [See (3.1) & (3.2)]. Suppose $\bar{x}_2$ is known, then there is no need of imputation and the proposed estimators (3.2) and (3.3) reduce into :

$$\bar{x}^{(\#)} = \left(\frac{N_1\bar{x}_1 + N_2\bar{x}_2}{N}\right) \tag{7.1}$$

$$\left[\left(\bar{y}_{FT}\right)_w\right]_k = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right)\left(\frac{(A+C)\bar{X} + fB\bar{x}^{(\#)}}{(A+fB)\bar{X} + C\bar{x}^{(\#)}}\right) \tag{7.2}$$

where, k is a constant $(0 < k < \infty)$ and

$A = (k-1)(k-2)$; $B = (k-1)(k-4)$; $C = (k-2)(k-3)(k-4)$; $f = n/N$.

Theorem 7.1: The estimator $\left[\left(\bar{y}_{FT}\right)_w\right]_k$ is biased for $\bar{Y}$ with the amount of bias

$$\mathrm{B}\left[\left(\bar{y}_{FT}\right)_w\right]_k = \left(\psi_1' - \psi_2'\right)\left[\left(D_1 - \frac{1}{N}\right)W_1^2 r_1 C_{1X}\{s_1\rho_1 C_{1Y} - \psi_2' r_1 C_{1X}\} + \left(D_2 - \frac{1}{N}\right)W_2^2 r_2 C_{2X}\{s_2\rho_2 C_{2Y} - \psi_2' r_2 C_{2X}\}\right]$$

where $\psi_1' = fB/(A+fB+C)$; $\psi_2' = C/(A+fB+C)$. (7.3)

Proof: The estimator $\left[\left(\bar{y}_{FT}\right)_w\right]_k$ may be approximated as :

$$\left[\left(\bar{y}_{FT}\right)_w\right]_k = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right)\left(\frac{(A+C)\bar{X} + fB\bar{x}^{(\#)}}{(A+fB)\bar{X} + C\bar{x}^{(\#)}}\right)$$

$$= \left[\bar{Y} + W_1\bar{Y}_1 e_1 + W_2\bar{Y}_2 e_2\right]\left[1 + \psi_1'(W_1 r_1 e_3 + W_2 r_2 e_4)\right]\left[1 + \psi_2'(W_1 r_1 e_3 + W_2 r_2 e_4)\right]^{-1}$$

Expanding the above using binominal expansion, and ignoring $\left(e_i^k e_j^l\right)$ terms for $(k+l) > 2$, $(k, l = 0,1,2 \ldots)$, $(i, j = 1, 2, 3, 4)$; the estimator results into

$$\left[(\bar{y}_{FT})_w\right]_k = \bar{Y} + \bar{Y}(\Delta_1 - \Delta_2) + W_1\bar{Y}_1 e_1(1+\Delta_1-\Delta_2) + W_2\bar{Y}_2 e_2(1+\Delta_1-\Delta_2) \quad (7.4)$$

where, $\Delta_1 = (\psi_1^{'} - \psi_2^{'})(W_1 r_1 e_3 + W_2 r_2 e_4)$, $\Delta_2 = \psi_2^{'}(\psi_1^{'} - \psi_2^{'})(W_1 r_1 e_3 + W_2 r_2 e_4)^2$ and $W_1 r_1 + W_2 r_2 = 1$ holds.

Further, up to first order of approximation, one may derive the following:

(i) E [Δ_1] = 0; **(ii)** $E[\Delta_1^2] = (\psi_1^{'} - \psi_2^{'})^2\left[W_1^2 r_1^2\left(D_1 - \frac{1}{N}\right)C_{1X}^2 + W_2^2 r_2^2\left(D_2 - \frac{1}{N}\right)C_{2X}^2\right]$;

(iii) $E[\Delta_2] = \psi_2^{'}(\psi_1^{'} - \psi_2^{'})\left[W_1^2 r_1^2\left(D_1 - \frac{1}{N}\right)C_{1X}^2 + W_2^2 r_2^2\left(D_2 - \frac{1}{N}\right)C_{2X}^2\right]$;

(iv) $E[e_1\Delta_1] = (\psi_1^{'} - \psi_2^{'})\, W_1\, r_1\left(D_1 - \frac{1}{N}\right)\rho_1 C_{1X} C_{1Y}$;

(v) $E[e_2\Delta_1] = (\psi_1^{'} - \psi_2^{'})\, W_2\, r_2\left(D_2 - \frac{1}{N}\right)\rho_2 C_{2X} C_{2Y}$;

(vi) $E[e_1\Delta_2] = 0$ [under $o(n^{-1})$] ; **(vii)** $E[e_2\Delta_2] = 0$ [under $o(n^{-1})$] ;

The bias of the estimator without imputation is

$$B\left[(\bar{y}_{FT})_w\right]_k = E\left[\left\{(\bar{y}_{FT})_w\right\}_k - \bar{Y}\right] = E\left[\bar{Y}(\Delta_1-\Delta_2) + W_1\bar{Y}_1 e_1(1+\Delta_1-\Delta_2) + W_2\bar{Y}_2 e_2(1+\Delta_1-\Delta_2)\right]$$

$$= \bar{Y}(\psi_1^{'} - \psi_2^{'})\left[\left(D_1 - \frac{1}{N}\right)W_1^2 r_1 C_{1X}\left\{s_1\rho_1 C_{1Y} - \psi_2^{'} r_1 C_{2X}\right\} + \left(D_2 - \frac{1}{N}\right)W_2^2 r_2 C_{2X}\left\{s_2\rho_2 C_{2Y} - \psi_2^{'} r_2 C_{2X}\right\}\right]$$

Theorem 7.2 : The mean squared error of the estimator $\left[(\bar{y}_{FT})_w\right]_k$ is :

$$M\left[(\bar{y}_{FT})_w\right]_k = \bar{Y}^2\left[\left(D_1 - \frac{1}{N}\right)W_1^2\left\{s_1^2 C_{1Y}^2 + (\psi_1^{'} - \psi_2^{'})^2 r_1^2 C_{1X}^2 + 2s_1(\psi_1^{'} - \psi_2^{'})\, r_1\, \rho_1 C_{1Y} C_{1X}\right\}\right.$$
$$\left. + \left(D_2 - \frac{1}{N}\right)W_2^2\left\{s_2^2 C_{2Y}^2 + (\psi_1^{'} - \psi_2^{'})^2 r_2^2 C_{2X}^2 + 2s_2(\psi_1^{'} - \psi_2^{'})^2 r_2 \rho_2 C_{2Y} C_{2X}\right\}\right] \quad (7.5)$$

Proof : $M\left[(\bar{y}_{FT})_w\right]_k = E\left[\left\{(\bar{y}_{FT})_w\right\}_k - \bar{Y}\right]^2$

$$M\left[(\bar{y}_{FT})_w\right]_k = E\left[\bar{Y}(\Delta_1-\Delta_2) + W_1\bar{Y}_1 e_1(1+\Delta_1-\Delta_2) + W_2\bar{Y}_2 e_2(1+\Delta_1-\Delta_2)\right]^2$$

$$= \bar{Y}^2 E(\Delta_1^2) + W_1^2\bar{Y}_1^2 E(e_1^2) + W_2^2\bar{Y}_2^2 E(e_2^2) + 2W_1\,\overline{Y}\overline{Y}_1 E(e_1\Delta_1) + 2W_2\,\overline{Y}\overline{Y}_2 E(e_2\Delta_1) + 2W_1 W_2 \overline{Y}_1\overline{Y}_2 E(e_1 e_2)$$

$$M\left[(\bar{y}_{FT})_w\right]_k = \bar{Y}^2\left[\left(D_1 - \frac{1}{N}\right)W_1^2\left\{s_1^2 C_{1Y}^2 + (\psi_2^{'} - \psi_1^{'})^2 r_1^2 C_{1X}^2 + 2s_1(\psi_2^{'} - \psi_1^{'})\, r_1\, \rho_1 C_{1Y} C_{1X}\right\}\right.$$
$$\left. + \left(D_2 - \frac{1}{N}\right)W_2^2\left\{s_2^2 C_{2Y}^2 + (\psi_2^{'} - \psi_1^{'})^2 r_2^2 C_{2X}^2 + 2s_2(\psi_2^{'} - \psi_1^{'})^2 r_2 \rho_2 C_{2Y} C_{2X}\right\}\right]$$

Remark

At k = 1, k = 2, k = 3 and k = 4, the biases and mean squared errors of non-imputed estimators are given below :

Case I : k = 1 => A = 0 ; B = 0 ; C = **-** 6 ; $\psi_1^{'}$ = 0; $\psi_2^{'}$ = 1;

$$\left[(\bar{y}_{FT})_w\right]_{k=1} = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right)\left(\frac{\bar{X}}{\bar{x}^{(\#)}}\right) \quad (7.6)$$

$$B\left[(\bar{y}_{FT})_w\right]_{k=1} = -\bar{Y}\left[\left(D_1 - \frac{1}{N}\right)W_1^2 r_1 C_{1X}\left\{s_1\rho_1 C_{1Y} - r_1 C_{1X}\right\} + \left(D_2 - \frac{1}{N}\right)W_2^2 r_2 C_{2X}\left\{s_2\rho_2 C_{2Y} - r_2 C_{2X}\right\}\right] \quad (7.7)$$

$$M\left[(\bar{y}_{FT})_w\right]_{k=1} = \bar{Y}^2\left[\left(D_1 - \frac{1}{N}\right)W_1^2\left\{s_1^2 C_{1Y}^2 + r_1^2 C_{1X}^2 - 2s_1 r_1\, \rho_1 C_{1Y} C_{1X}\right\} + \left(D_2 - \frac{1}{N}\right)W_2^2\left\{s_2^2 C_{2Y}^2 + r_2^2 C_{2X}^2 - 2s_2 r_2\, \rho_2 C_{2Y} C_{2X}\right\}\right] \quad (7.8)$$

Case II : $k = 2 \Rightarrow A = 0;\ B = -2;\ C = 0;\ \psi_1^{'} = 1;\ \psi_2^{'} = 0;$

$$\left[(\bar{y}_{FT})_w\right]_{k=2} = \left(\frac{N_1\bar{y}_2 + N_2 y_2}{N}\right)\left(\frac{\bar{x}^{(\#)}}{\bar{X}}\right) \tag{7.9}$$

$$B\left[(\bar{y}_{FT})_w\right]_{k=2} = \bar{Y}\left[\left(D_1 - \frac{1}{N}\right)W_1^2 s_1 r_1 \rho_1 C_{1X} C_{1Y} + \left(D_2 - \frac{1}{N}\right)W_2^2 s_2 r_2 \rho_2 C_{2X} C_{2Y}\right] \tag{7.10}$$

$$M\left[(\bar{y}_{FT})_w\right]_{k=2} = \bar{Y}^2\left[\left(D_1 - \frac{1}{N}\right)W_1^2\left\{s_1^2 C_{1Y}^2 + r_1^2 C_{1X}^2 + 2 s_1 r_1 \rho_1 C_{1Y} C_{1X}\right\}\right.$$
$$\left. + \left(D_2 - \frac{1}{N}\right)W_2^2\left\{s_2^2 C_{2Y}^2 + r_2^2 C_{2X}^2 + 2 s_2 r_2 \rho_2 C_{2Y} C_{2X}\right\}\right] \tag{7.11}$$

Case III : $k = 3 \Rightarrow A = 2;\ B = -2;\ C = 0;\ \psi_1^{'} = -f(1-f)^{-1};\ \psi_2^{'} = 0;$

$$\left[(\bar{y}_{FT})_w\right]_{k=3} = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right)\left(\frac{\bar{X} - \bar{x}^{(\#)}}{(1-f)\bar{X}}\right) \tag{7.12}$$

$$B\left[(\bar{y}_{FT})_w\right]_{k=3} = -\bar{Y}f(1-f)^{-1}\left[\left(D_1 - \frac{1}{N}\right)W_1^2 r_1 s_1 \rho_1 C_{1Y} C_{1X} + \left(D_2 - \frac{1}{N}\right)W_2^2 r_2 s_2 \rho_2 C_{2Y} C_{2X}\right] \tag{7.13}$$

$$\mathrm{M}\left[(\bar{y}_{FT})_w\right]_{k=3} = \bar{Y}^2\left[\left(D_1 - \frac{1}{N}\right)W_1^2\left\{s_1^2 C_{1Y}^2 + (1-f)^{-2} f^2 r_1^2 C_{1X}^2 - 2(1-f)^{-1} f\, s_1 \rho_1 C_{1Y} C_{1X}\right\}\right.$$
$$\left. + \left(D_2 - \frac{1}{N}\right)W_2^2\left\{s_2^2 C_{2Y}^2 + (1-f)^{-2} f^2 r_2^2 C_{2X}^2 - 2(1-f)^{-1} f\, s_2 r_2 \rho_2 C_{2Y} C_{2X}\right\}\right] \tag{7.14}$$

Case IV: $k = 4 \Rightarrow A = 6;\ B = 0;\ C = 0;\ \psi_1^{'} = 0;\ \psi_2^{'} = 0;$

$$\left[(\bar{y}_{FT})_w\right]_{k=4} = \left(\frac{N_1\bar{y}_1 + N_2\bar{y}_2}{N}\right) \tag{7.15}$$

$$B\left[(\bar{y}_{FT})_w\right]_{k=4} = 0; \tag{7.16}$$

$$V\left[(\bar{y}_{FT})_w\right]_{k=4} = \left[\bar{Y}^2\left(D_1 - \frac{1}{N}\right)W_1^2 s_1^2 C_{1Y}^2 + \left(D_2 - \frac{1}{N}\right)W_2^2 s_2^2 C_{2Y}^2\right] \tag{7.17}$$

8. Numerical Illustration

Consider two artificial populations I and II (given in Appendix A and B), each of which is divided into R and NR-groups of sizes N_1 and N_2 respectively. Let the samples of sizes 40 and 30 respectively from populations I and II drawn with SRSWOR be post-stratified into R and NR-groups. Then, we have

Population I	**Population II**
$n_1 = 28;\ n_2 = 12;\ n = 40;\ f = 0.22$	$n_1 = 20;\ n_2 = 10;\ n = 40;\ f = 0.20$

The values of population parameters for the two populations are given in Table 8.1 and the values of bias and MSE are shown in Table 8.2.

	Entire Population		R-group		NR-group	
	I	II	I	II	I	II
Size	180	150	100	90	80	60
Mean *Y*	159.03	63.77	173.60	66.33	140.81	59.92
Mean *X*	113.22	29.20	128.45	30.72	94.19	26.92
M.S. *Y*	2205.18	299.87	2532.36	349.33	1219.90	206.35
M.S. *X*	1972.61	110.43	2300.86	112.67	924.17	100.08
C.V. *Y*	0.295	0.272	0.290	0.282	0.248	0.240
C.V. *X*	0.392	0.360	0.373	0.345	0.323	0.372
Cor.Coeff.	0.897	0.809	0.857	0.805	0.956	0.808

Table 8.1: Parameters of Populations – I & II given in Appendix A & B.

Type of Estimator	Description of the estimator	Population-I		Population-II	
		Bias	MSE	Bias	MSE
$(\bar{y}_{FT})_k$	$(\bar{y}_{FT})_{k=1}$	-2.0	17.1255	-2.3386	8.025
	$(\bar{y}_{FT})_{k=2}$	1.6628	228.6822	2.6018	49.8306
	$(\bar{y}_{FT})_{k=3}$	1.4183	28.0158	-0.6512	7.0054
	$(\bar{y}_{FT})_{k=4}$	0	43.64	0	9.2662
$(\bar{y}_{FT})_w$	$[(\bar{y}_{FT})_w]_{k=1}$	0.1433	12.9589	0.1095	6.0552
	$[(\bar{y}_{FT})_w]_{k=2}$	0.3141	216.3024	0.1599	46.838
	$[(\bar{y}_{FT})_w]_{k=3}$	-0.5962	24.327	-0.031	5.2423
	$[(\bar{y}_{FT})_w]_{k=4}$	0	43.64	0	9.2662

Table 8.2: Bias and M.S.E. Comparisons of $(\bar{y}_{FT})_k$ and $(\bar{y}_{FT})_w$

The m.s.e. of the proposed imputed estimator is higher than that of non-imputed estimator but both are very close. Obviously, the non-imputed estimator will be better than the imputed estimator due to complete availability of information. The proposed one is very near to the non-imputed estimator showing utility due to new estimation technique in missing observation environment.

Define a term *LI* as "percentage loss due to imputation" with formulation.

$$(LI)_k = \left[\frac{MSE(\bar{y}_{FT})_k}{MSE(\bar{y}_{FT})_w}\right] \times 100 \qquad (8.1)$$

The table 8.3 shows the variation of *LI* over *k*.

k	***(LI)$_k$***	
	Population –I	**Population-II**
k = 1	132.1524203	132.5307
k = 2	105.7233762	106.3893
k = 3	115.1633987	133.6322
k = 4	100	100

Table 8.5: Variation of *LI* over k

The percentage loss in MSE due to imputation is small and accomodable over suitable choice of *k*. But at the same time, proposed one tackles and solves the problem of missing observations also.

9. Conclusions

As per table 8.2 and 8.3, the imputed class performs closer to the non-imputed class of estimators over suitable choice of *k*. The over all comparative procedure shows almost a closed performance of imputed factor-type estimator to the same without imputation. The imputed factor-type class of estimators reveals a good potential for utilizing the information $\bar{X}_2$ in place of missing $\bar{x}_2$. The class presents efficient member when $k = 1$ and $k = 3$. The *LI* comparison shows that with a little loss, one can handle the non-responded observations effectively. Actually, the best choice of *k* is suppose to be near to $k = 1$ or near to $k = 3$. It is worthwhile to say that the proposed class contains estimators is effective for mean estimation even when some observations of auxiliary variable *X* are missing (or non-responded).

Acknowledgement

Authors are thankful to the referee for his valuable comments and useful suggestions which improved the quality of the original manuscript.

References

1. Agrawal, M.C. and Panda, K.B. (1993). An efficient estimator in post-stratification, Metron, 51, 3-4, 179-187.

2. Ahmed, M.S., Al-titi, Omar, Al-Rawi, Ziad and Abu-dayyeh, Walid (2006). Estimation of a population mean using different imputation method, Statistics in Transition, 7(6), 1247-1264.

3. Cochran, W.G. (2005). Sampling Techniques, Third Edition, John Willey & Sons, New Delhi.

4. Grover and Couper (1998). Non-response in household surveys, John Wiley and Sons, New York.

5. Hansen, M.H. and Hurwitz, W.N. (1946). The problem of non-response in sample surveys, Jour. Amer. Stat. Asso., 41, 517-529.

6. Hinde, R.L. and Chambers, R.L. (1990). Non-response imputation with multiple source of non-response, Jour. Off. Statistics, 7, 169-179.

7. Khot, P.S. (1994). A note on handling non-response in sample surveys, Jour. Amer. Stat. Assoc., 89, 693-696.

8. Lessler and Kalsbeek (1992). Non-response error in surveys, John Wiley and Sons, New York.

9. Rao, J.N.K. and Sitter, R.R. (1995). Variance estimation under two phase sampling with application to imputation for missing data, Biometrika, 82, 453-460.

10. Rubin, D.B. (1976). Inference and missing data, Biometrika, 63, 581-593.

11. Shukla, D. (2002). F-T estimator under two phase sampling, Metron, 59, 1-2, 253-263.

12. Shukla, D. and Dubey, J. (2001). Estimation in mail surveys under PSNR sampling scheme, Ind. Soc. Ag. Stat., 54, 3, 288-302.

13. Shukla, D. and Dubey, J. (2004). On estimation under post-stratified two phase non-response sampling scheme in mail surveys, International Jour. of Management and Systems, 20, 3, 269-278.

14. Shukla, D. and Dubey, J. (2006). On earmarked strata in post-stratification, Statistics in Transition, 7, 5, 1067-1085

15. Shukla, D. and Trivedi, M. (1999). A new estimator for post-stratified sampling scheme, Proceedings of NSBA-TA (Bayesian Analysis), Eds. Rajesh Singh, 206-218.

16. Shukla, D. and Trivedi, M. (2001). mean estimation in deeply stratified population under post-stratification. Ind. Soc. Ag. Stat., 54(2), 221-235.

17. Shukla, D. and Trivedi, M. (2006). Examining stability of variability ratio with application in post-stratification, International Jour. of Management and Systems, 22, 1, 59-70.

18. Shukla, D., Bankey, A. and Trivedi, M. (2002). Estimation in post-stratification using prior information and grouping strategy, Ind. Soc. Ag. Stat., 55, 2, 158-173.

19. Shukla, D., Singh, V.K. and Singh, G.N. (1991). On the use of transformation in factor type estimator, Metron, 49 (1-4), 359-361.

20. Shukla, D., Trivedi, M. and Singh, G.N. (2006). Post-stratification two-way deeply stratified population, Statistics in Transition, 7, 6, 1295-1310.

21. Singh G.N. and Singh V.K. (2001). On the use of auxiliary information in successive sampling, Jour. Ind. Soc. Ag. Stat., 54(1), 1-12.

22. Singh, H.P. (1986). A generalized class of estimators of ratio product and mean using supplementary information on an auxiliary character in PPSWR scheme, Guj. Stat. Rev., 13, 2, 1-30.

23. Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling, Metrika, 51, 266-276.

24. Singh, V.K. and Shukla, D. (1987). One parameter family of factor type ratio estimator, Metron, 45 (1-2), 273-283.

25. Singh, V.K. and Shukla, D. (1993). An efficient one-parameter family of factor type estimator in sample survey, Metron, 51, 1-2, 139-159.

26. Singh, V.K. and Singh, G.N. (1991). Chain type estimator with two auxiliary variables under double sampling scheme, Metron, 49, 279-289.

27. Singh, V.K., Singh, G. N. and Shukla, D. (1994). A class of chain ratio estimator with two auxiliary variables under double sampling scheme, Sankhya, Ser. B., 46(2), 209-221.

28. Smith, T.M.F. (1991). Post-stratification, The Statisticians, 40, 323-330.

29. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Ashok, C. (1984). Sampling Theory of Surveys with Applications, Iowa State University Press, I.S.A.S. Publication, New Delhi.

30. Wywial, J. (2001). *Stratification of population after sample selection,* Statistics in Transition, 5, 2, 327-348.

Appendix A: Population I (N= 180)

R-group: (N_1=100)

Y:	110	75	85	165	125	110	85	80	150	165	135	120	140	135	145
X:	80	40	55	130	85	50	35	40	110	115	95	60	70	85	115
Y:	200	135	120	165	150	160	165	145	215	150	145	150	150	195	190
X:	150	85	80	100	25	130	135	105	185	110	95	75	70	165	160
Y:	175	160	165	175	185	205	140	105	125	230	230	255	275	145	125
X:	145	110	135	145	155	175	80	75	65	170	170	190	205	105	85
Y:	110	110	120	230	220	280	275	220	145	155	170	195	170	185	195
X:	75	80	90	165	160	205	215	190	105	115	135	145	135	110	145
Y:	180	150	185	165	285	150	235	125	165	135	130	245	255	280	150
X:	135	110	135	115	125	205	100	195	85	115	75	190	205	210	105
Y:	205	180	150	205	220	240	260	185	150	155	115	115	220	215	230
X:	110	105	110	175	180	215	225	110	90	95	85	75	175	185	190
Y:	210	145	135	250	265	275	205	195	180	115					
X:	170	85	95	190	215	200	165	155	150	175					

NR-group: (N_2=80)

Y:	85	75	115	165	140	110	115	13.5	120	125	120	150	145	90	105
X:	55	40	65	115	90	55	60	65	70	75	80	120	105	45	65
Y:	110	90	155	130	120	95	100	125	140	155	160	145	90	90	95
X:	70	60	85	95	80	55	60	75	90	105	125	95	45	55	65
Y:	115	140	180	170	175	190	160	155	175	195	90	90	80	90	80
X:	75	105	120	115	125	135	110	115	135	145	45	55	50	60	50
Y:	105	125	110	120	130	145	160	170	180	`145	130	195	200	160	110
X:	65	75	70	80	85	105	110	115	130	95	65	135	130	115	55
Y:	155	190	150	180	200	160	155	170	195	200	150	165	155	180	200
X:	115	130	110	120	125	145	120	105	100	95	90	105	125	130	145
Y:	160	155	170	195	200										
X:	120	115	120	135	150										

Appendix B : Population II (N=150)

R-group (N_1=90)

Y:	90	75	70	85	95	55	65	80	65	50	45	55	60	60	95
X:	30	35	30	40	45	25	40	50	35	30	15	20	25	30	40
Y:	100	40	45	55	35	45	35	55	85	95	65	75	70	80	65
X:	50	10	25	25	10	15	10	25	35	55	35	40	30	45	40
Y:	90	95	80	85	55	60	75	85	80	65	35	40	95	100	55
X:	40	50	35	45	35	25	30	40	25	35	10	15	45	45	25
Y:	45	40	40	35	55	75	80	80	85	55	45	70	80	90	55
X:	15	15	20	10	30	25	30	40	35	20	25	30	40	45	30
Y:	65	60	75	75	85	95	90	90	45	40	45	55	60	65	60
X:	25	40	35	30	40	35	40	35	15	25	15	30	30	25	20
Y:	75	70	40	55	75	45	55	60	85	55	60	70	75	65	80
X:	25	20	35	30	45	10	30	25	40	15	25	30	35	30	45

NR-group (N_2=60)

Y:	40	90	95	70	60	65	85	55	45	60	65	60	55	55	45
X:	10	30	30	30	25	30	40	25	15	20	30	30	35	25	20
Y:	65	80	55	65	75	55	50	55	60	45	40	75	75	45	70
X:	35	45	30	30	40	15	15	20	30	15	10	40	45	10	30
Y:	65	70	55	35	35	50	55	35	55	60	30	35	45	55	65
X:	30	40	30	10	15	25	30	15	20	30	10	20	15	30	30
Y:	75	65	70	65	70	45	55	60	85	55	60	70	75	65	80
X:	30	35	40	25	45	10	30	25	40	15	25	30	35	30	45