ESTIMATING THE VARIANCE IN THE PRESENCE OF IMPUTATION USING A RESIDUAL

Steven Kaufman, National Center for Education Statistics
Room 422d, 555 New Jersey Ave. N. W., Washington, D.C. 20208

Key Words: Imputation, Variance, Simulation, Half-Sample Replication

1. Introduction

With most surveys there is always some item nonresponse. One way of dealing with item nonresponse is to include imputations for the missing data. This makes it easy to compute appropriate population estimates. However, the variance will be underestimated in the presence of imputation. This follows from the main assumption with design based estimation; the only random variable is the variable specifying which units are in sample. When an imputation is generated from a sample, the imputation is not known given the unit requiring imputation. However, since it is assumed that the imputed value is known (i.e., it has no variability), standard variance estimation can not properly reflect the imputation variance.

A number of methodologies have been proposed to estimate imputation variance. Multiple imputation (Rubin,1978), a modified jacknife (Rao, 1992) and a model assisted methodology (Särndal, 1990) are some of the approaches that have been proposed. Rao's and Särndal's methodologies require special software to estimate the variance. Rubin's methodology requires the computation of the variance estimate associated with the average of the multiple imputation estimates. With respect to additional software requirements, the multiple imputation variance estimate, although not as complicated as the other methods, is still more complicated then standard variance estimation.

The methodology proposed in this paper is a mixture of the methodologies described above. Like multiple imputation, two potential imputations are assigned to each nonrespondent. The difference between these values is a residual that is added to appropriate data elements to reflect the imputation variance. The residual term differs from the one used in the modified jacknife, in that the modified jacknife residual introduces the variability replicate by replicate, while here the variability is introduced to the actual data elements. Once the residuals have been added, standard replication programs can compute the total variance. The questions this paper addresses are: 1) when will this process appropriately measure the total variance; and 2) when the process is not appropriate, what must be added to provide appropriate estimates.

A simulation study, modeled after NCES's Schools and Staffing Survey (SASS), which has a complex sample design, will demonstrate the proposed procedures. A nearest neighbor imputation will be used. The nonresponse will be modeled assuming unequal nonresponse rates per cell and that larger units are more likely to be nonrespondents.

2. Imputations

The nearest neighbor imputations used in this paper are done within imputation cells, after the schools have been sorted by the number of student per school. The imputation cells are state/school level/urbanicity. There are three school levels - elementary, secondary and combined schools. There are three levels of urbanicity - central city, urban fringe/large town and rural/small town. After the file is sorted, it is accessed sequentially using the nearest responding school as the donor for a nonresponding school. Two imputations will be determined for each nonrespondent. One where the file is sorted in ascending order and another where the file is sorted in descending order. A random selection is used to determine which imputation is used in the estimate of interest ($\hat{y}_{\bullet}$).

3. Definitions

r is the set of responding units

nr is the set of nonresponding units

3.1 Terms defining the imputation and residual

$$\tilde{y}_k = \begin{cases} y_k & \text{if } k \in \mathrm{r} \\ y_{k1} I_k + y_{k2}(1 - I_k) & \text{if } k \in \mathrm{nr} \end{cases}$$

I_k represents an independent selection within each unit k with probability of .5 for a value of 1, and 0 otherwise

y_k is the response from unit k

y_{k1} is the nearest neighbor imputation in ascending order

y_{k2} is the nearest neighbor imputation in descending order

$\hat{y}_{\bullet} = \sum_{k \in s} w_k \tilde{y}_k$, the main estimate of interest

w_k is the sampling weight for unit k

s is a probability sample of units

$\hat{y} = \sum_{k \in s} w_k y_k$, the estimate with complete response

$$\tilde{d}_k^R = \begin{cases} 0 \text{ if } k \in \text{r} \\ \hat{y}_{j_k 2} - \hat{y}_{j_k 1} \text{ if } k \in \text{nr and } y_{k1} \text{ is used in } \hat{y}_\bullet \\ \hat{y}_{j_k 1} - \hat{y}_{j_k 2} \text{ if } k \in \text{nr and } y_{k2} \text{ is used in } \hat{y}_\bullet \end{cases}$$

where:

j_k is a unit independently and randomly selected from $\{j \mid j \in \text{nr and within } k\text{'s imputation cell}\}$.

The selection is done proportional to the w_j's.

In addition, R will be used to represent the selection of unit j_k.

$$\hat{d}^R = \sum_{k \in s} w_k \tilde{d}_k^R$$

3.2 Terms used in section 4

$$\tilde{d}_{ki}^T = \begin{cases} 0 \text{ if } k \in \text{r} \\ y_{m_k 2} - y_{m_k 1} \text{ if } k \in \text{nr and } i = 1 \\ y_{m_k 1} - y_{m_k 2} \text{ if } k \in \text{nr and } i = 2 \end{cases}$$

m_k is defined like j_k above, independently for $i = 1$ and 2.

T is used to represent the m_k selection.

$$\hat{d}_i^T = \sum_{k \in s} w_k \tilde{d}_{ki}^T = \hat{y}_{2i}^T - \hat{y}_{1i}^T$$

$$\tilde{y}_{ki} = \begin{cases} y_k \quad \text{if } k \in \text{r} \\ y_{k1} \quad \text{if } k \in \text{nr and } i = 1 \\ y_{k2} \quad \text{if } k \in \text{nr and } i = 2 \end{cases}$$

$$\hat{y}_i = \sum_{k \in s} w_k \tilde{y}_{ki},$$

$$\bar{y}_\bullet = 1/2(\hat{y}_1 + \hat{y}_2)$$

3.3 Terms used in section 5

$$\tilde{d}_k = \begin{cases} 0 \text{ if } k \in \text{r} \\ y_{k2} - y_{k1} \text{ if } k \in \text{nr and } y_{k1} \text{ is used in } \hat{y}_\bullet \\ y_{k1} - y_{k2} \text{ if } k \in \text{nr and } y_{k2} \text{ is used in } \hat{y}_\bullet \end{cases}$$

y_k' is the term subtracted in $\tilde{d}_k$

$$\hat{y}' = \sum_{k \in s} w_k y_k'$$

$$\hat{d} = \sum_{k \in s} w_k \tilde{d}_k$$

4. Measuring the Imputation Variance

The goal of this paper is to present a methodology which measures the imputation variance using standard replication variance software packages, in a simple manner. One way of doing this is adding an independent residual term ($\tilde{d}_k^R$) to the data elements ($\tilde{y}_k$), so that an appropriate amount of imputation variance is added. The estimate ($\hat{y}_\bullet$) is transformed into $\hat{Y} = \hat{y}_\bullet + \hat{d}^R$. An appropriate constant is added to the $\tilde{d}_k^R$'s , so that $\hat{Y} = \hat{y}_\bullet$. Now, the question is when does $V(\hat{Y})$ appropriately measure the true variance. When $V(\hat{Y})$ is not appropriate, what must be done to make it appropriate. The first step is to compute $V(\hat{Y})$.

4.1 Computation of Variance of $\hat{Y}$

Let $\hat{Y} = \sum_{k \in s} w_k (\tilde{y}_k + \tilde{d}_k^R) = \hat{y}_\bullet + \hat{d}^R$.

$$V(\hat{Y}) = \underset{1}{V}\underset{2}{E}(\hat{Y}) + \underset{1}{E}\underset{2}{V}(\hat{Y}), \qquad (1)$$

where: 1 represents the selection of the sample s and 2 represents the imputation selection of I_k and residual selection R.

4.2 Computation of $\underset{1}{V}\underset{2}{E}(\hat{Y})$

$$\underset{2}{E}(\hat{Y}) = 1/2(\hat{y}_1 + \hat{d}_1^T) + 1/2(\hat{y}_2 + \hat{d}_2^T)$$

d_i^T has been normalized to equal zero

Looking at the right hand term first:

$$1/4\underset{1}{V}(\hat{y}_1 + \hat{d}_1^T) = 1/4\underset{1}{V}(\hat{y}_1 + \hat{y}_{21}^T - \hat{y} + \hat{y} - \hat{y}_{11}^T)$$

$$= 1/4[\underset{1}{V}(\hat{y}_1) + \underset{1}{V}(\hat{y}_{21}^T - \hat{y}) + \underset{1}{V}(\hat{y} - \hat{y}_{11}^T)$$

$$+ 2\underset{1}{\text{cov}}(\hat{y}_{21}^T - \hat{y}, \hat{y} - \hat{y}_{11}^T)], \text{ since } \text{cov}(\hat{y}_1, \hat{d}_1^T) = 0$$

$$= 1/4[\underset{1}{V}(\hat{y}_1 - (\hat{y}_{21}^T - \hat{y})) + \underset{1}{V}(\hat{y} - \hat{y}_{11}^T) \qquad (2)$$

$+ 2\underset{1}{\text{cov}}(\hat{y}_{21}^T - \hat{y}, \hat{y})]$, assuming $\hat{y}_{11}^T$ is distributed as $\hat{y}_1$

Likewise,

$$1/4\underset{1}{V}(\hat{y}_2 + \hat{d}_2^T) = 1/4[\underset{1}{V}(\hat{y}_2 - (\hat{y}_{22}^T - \hat{y})) + \underset{1}{V}(\hat{y} - \hat{y}_{12}^T)$$

$$+ 2\underset{1}{\text{cov}}(\hat{y}_{22}^T - \hat{y}, \hat{y})] \qquad (3)$$

Combining 2 and 3 gives:

$$\underset{1}{V}\underset{2}{E}(\hat{Y}) = \underset{1}{V}(\hat{y}) + \underset{1}{V}(\hat{y} - \bar{y}_\bullet) + 2\underset{1}{\text{cov}}(\bar{y}_\bullet - \hat{y}, \hat{y}), \quad (4)$$

assuming that $\hat{y}_{21}^T$ and $\hat{y}_{12}^T$ are distributed as $\hat{y}_2$; and $\hat{y}_{11}^T$ and $\hat{y}_{22}^T$ are distributed as $\hat{y}_1$;.

$\text{cov}(\hat{y} + \hat{y}_1 - \hat{y}_{21}^T, \hat{y} + \hat{y}_2 - \hat{y}_{22}^T)$ and $\text{cov}(\hat{y} - \hat{y}_{11}^T, \hat{y} - \hat{y}_{12}^T)$ are zero by the independence in the T selection

4.3 Computation of $\underset{2}{V}(\hat{Y})$

$$\underset{2}{V}(\hat{Y}) = \underset{2}{V}(\hat{y}_{\bullet}) + \underset{2}{V}(\hat{d}^R)$$
$$= \underset{2}{V}(\hat{y}_{\bullet}) + \underset{2}{V}^{\text{ppz}}(\hat{d}^R) \quad (5)$$

$\underset{2}{V}^{\text{ppz}}$ is the variance associated with probability proportionate to size sampling with replacement. The size is the w_k's for $k \in \text{nr}$ within each imputation cell.

4.4 Computation of $V(\hat{Y})$

Putting 1, 4 and 5 together gives:

$$V(\hat{Y}) = \underset{1}{V}(\hat{y}) + \underset{1}{V}(\hat{y} - \bar{y}_{\bullet}) + 2\underset{1}{\text{cov}}(\bar{y}_{\bullet} - \hat{y}, \hat{y})$$
$$+ \underset{1}{E}\underset{2}{V}(\hat{y}_{\bullet}) + \underset{1}{E}\underset{2}{V}^{ppz}(\hat{d}^R)$$
$$= V(\hat{y}) + V(\hat{y} - \hat{y}_{\bullet}) + 2\,\text{cov}(\bar{y}_{\bullet} - \hat{y}, \hat{y})$$
$$+ \underset{1}{E}\underset{2}{V}^{ppz}(\hat{d}^R)$$

As formulated by Särndal (1994), $V(\hat{y})$ is the variance assuming no nonresponse and $V(\hat{y} - \hat{y}_{\bullet})$ is the imputation variance. The sum of these components equals the total variance, assuming the $\text{cov}(\hat{y}, \hat{y} - \hat{y}_{\bullet}) = 0$. The simulation done in Särndal (1994), using nearest neighbor imputation, indicates that the covariance term is small and can be approximated by zero.

$2\,\text{cov}(\bar{y}_{\bullet} - \hat{y}, \hat{y})$ and $\underset{1}{E}\underset{2}{V}^{ppz}(\hat{d}^R)$ are terms not included in the estimate of the true variance. There are two ways of handling these terms. The first is to estimate the terms and subtract them from $V(\hat{Y})$. The second way is realizing that the only non-zero values in these terms comes from nonrespondents and if the item response rate is relatively high, these terms should be small and can be ignored.

5. **Computation of** $\hat{\text{cov}}(\bar{y}_{\bullet} - \hat{y}, \hat{y})$

Since $\text{cov}(\bar{y}_{\bullet} - \hat{y}, \hat{y})$ requires y_k for $k \in \text{nr}$, it can not be computed with the given sample. However, it can be approximated using $\tilde{d}_k$. If $\tilde{d}_k = y_{k2} - y_{k1}$, and given how $\tilde{d}_k$ is formed, y_{k2} is the nearest neighbor imputation for y_{k1}. Therefore, $\tilde{d}_k$ can be viewed as the difference between the nearest neighbor imputation for a known value and the known value. Since $\tilde{d}_k$, in terms of the nearest neighbor imputation, is close to unit k, the covariance might be approximated by $\text{cov}(\hat{d}, \hat{y}')$.

6. **Computation of** $\underset{2}{V}^{ppz}(\hat{d}^R)$

$\underset{1}{E}\underset{2}{V}^{ppz}(\hat{d}^R)$ is estimated from a sample by computing $\underset{2}{V}^{ppz}(\hat{d}^R)$ for that sample. This is done by either using an exact formula (see Cochran, 1977) or by replication given a set of replicate weights designed for the conditional variance. For this report the exact formula will be used.

7. **Simulations**

In order to determine how well the proposed imputations work, a simulation study is performed for two variations of the proposed variance methodology, using four states within the SASS sample design and performing 5,000 simulations.

For each simulation sample, a nearest neighbor imputation ($\tilde{y}_k$) is performed on the selected sample of nonrespondents. Three populations of nonrespondents with 5%, 15% and 30% nonresponse will be generated and compared.

The simulation estimates are based on the design and collection variables that can be found on the frame. In this way, estimates for any selected sample, as well as, estimates of the true variance, can be computed.

7.1 Population of Nonrespondents

In order to do the analysis, a population of nonrespondents must be defined. A school k is chosen to be a member of the nonrespondent population by independently selecting k, proportional to P_k^{NR}. P_k^{NR} is determined to obtain an expected X% unweighted nonresponse rate in s.

$$P_k^{NR} = \left(p_k \,/ \sum_{k \in c} p_k^2\right) \times N_c^{NR}$$

p_k is the probability of selecting k in s

c is an imputation cell

N_c^{NR} is the expected number of nonrespondents in c

$$N_c^{NR} = \left(en_c r_c N_l^{NR}\right) / \left(\sum_{c \in l} en_c r_c\right)$$

en_c is the expected sample size in c (i.e., $\sum_{k \in c} p_k$)

r_c is the relative rate of nonresponse in c (i.e., 2 for central city schools, 1 otherwise)

N_l^{NR} is the expected number of nonrespondents required for the analysis in stratum l

$$N_l^{NR} = \left(\sum_{c \in l} en_c\right) \times (\text{X}/100)$$

It is easy to verify: 1) the expected nonresponse rate for s will be X%; 2) the relative rate of nonresponse in central city cells is twice that of other cells; and 3) larger schools, as measured by p_k, will have a higher probability of being selected as nonrespondents.

7.2 Simulation Sample Design

The Schools and Staffing Survey (SASS) is a stratified probability proportionate to size sample of elementary, secondary and combined schools. The selection is done systematically using the square root of the number of teachers per school as the measure of size. State by school level cells define the stratification. Before selection, schools are sorted to provide a good geographic distribution. Estimates are designed to provide state estimates.

In order to assure unbiased variance estimation using half-sample replication, the simulation design has been slightly altered. Each state/school level stratum has been split into a number of strata so that exactly two schools are selected within each stratum, while maintaining the original sample size. Another modification is that the selection within stratum is done with replacement.

7.3 Estimates

Three estimates per state are computed:

$\hat{y}_e = \sum_{k \in s} S_k p_{ke}$; $\hat{y}_m = \sum_{k \in s} S_k p_{km}$; and

$\hat{y}_h = \sum_{k \in s} S_k p_{kh}$.

S_k is the known number of students in school k

p_{ke} is the proportion of the students in school k in grades pre-kindergarten to 6

p_{km} is the proportion of the students in school k in grades 7 to 9

p_{kh} is the proportion of the students in school k in grades 10 to 12

It is assumed that S_k is known for all k and that only the p's require collection. Therefore, when k requires imputation, a nearest neighbor donor's p will be applied to S_k. This is a common SASS imputation that can be duplicated from the frame.

7.4 Simulated Variance Estimates

All variances below, except for $V_T(\hat{y}_\bullet)$ and $\underset{2}{V}^{ppz}(\hat{d}^R)$, are estimated using a fully balanced half-sample replication variance estimator (Wolter,1985).

The four variance estimates computed within each sample and averaged across samples are:

$\hat{V}_R = \hat{V}(\hat{Y})$

$\hat{V}_{EC}(\hat{Y}) = \hat{V}(\hat{Y}) - 2\,\hat{\mathrm{cov}}(\bar{y}_\bullet - \hat{y}, \hat{y}) - \underset{2}{V}^{ppz}(\hat{d}^R)$

$\hat{V}(\hat{y})$ is the variance estimate when all sample cases respond

$\hat{V}(\hat{y}_\bullet)$ is the variance estimate of $\hat{y}_\bullet$

The following estimates can not be computed from a single sample. However, they can be computed in the simulation setting and are used for comparison purposes. The first is the variance estimate using a true estimate of the covariance and the second is an estimate of the true variance.

$\hat{V}_{TC}(\hat{Y}) = \hat{V}(\hat{Y}) - 2\,\mathrm{cov}(\bar{y}_\bullet - \hat{y}, \hat{y}) - \underset{2}{V}^{ppz}(\hat{d}^R)$

$V_T(\hat{y}_\bullet) = 1/n \sum_{s=1}^{n} (\hat{y}_{\bullet s} - \bar{y}_{\bullet s})^2$

$\hat{y}_{\bullet s}$ and $\bar{y}_{\bullet s}$ are the value of $\hat{y}_\bullet$ for the s^{th} simulation and the average of the $\hat{y}_{\bullet s}$, respectively.

8. **Analysis Statistics**

To evaluate the imputation methodology relative bias of the estimated standard error (RB) and coverage rates (C_t and C_m) are computed.

$RB = \left(\sqrt{V_i(\hat{Y})} - \sqrt{V_T(\hat{y}_\bullet)}\right) / \sqrt{V_T(\hat{y}_\bullet)}$

$V_i(\hat{Y})$ is one of the variance estimates defined above

C_t is the coverage rate testing whether the true estimate with complete response is in the 95% confidence interval

C_m is the coverage rate testing whether the average of the simulated estimates ($\bar{y}_{\bullet s}$) is in the 95% confidence interval

Since the estimates with imputation are biased, the levels for C_t are unknown. One however, expects them to be smaller than 95%. Since C_m should be closer to 95%, it is used for comparison.

9. **Results**

Tables 1-3 provide results for the populations of nonrespondents, 30%, 15% and 5%, respectively. By comparing C_t and C_m for $\hat{V}(\hat{y}_\bullet)$, it can be seen that the bias in $\hat{y}_\bullet$, relative to the coverage rates, is small. The only exception is state 24's $\hat{y}_h$ in table 1, where the coverage rates differ by 9 points (90-81).

The half-sample coverage rates (C_t's for $\hat{V}(\hat{y})$) are usually less than 95%. Since C_t for $\hat{V}(\hat{y})$ provides the coverage rate with complete response, it will be assumed that the best imputation coverage rate, instead of being the one closest to 95%, will be the coverage rate coming from the procedure that gets closest to the C_t's for $\hat{V}(\hat{y})$.

From table 1, it can be seen that when the nonresponse rate is high $\hat{V}_{EC}(\hat{Y})$ is superior to $\hat{V}_R(\hat{Y})$. This is true for both relative bias and coverage rates. $\hat{V}_{EC}(\hat{Y})$ has better coverage rates 11 times, while $\hat{V}_R(\hat{Y})$ is better only once; $\hat{V}_{EC}(\hat{Y})$ has smaller relative biases 9 times, while $\hat{V}_R(\hat{Y})$ is better only 3 times. Therefore, $\hat{V}_{EC}(\hat{Y})$ should be used when the item nonresponse rate is high.

When the nonresponse rate is moderate in size, tables 2 show that $\hat{V}_R(\hat{Y})$ and $\hat{V}_{EC}(\hat{Y})$ are comparable with respect to coverage rates. $\hat{V}_{EC}(\hat{Y})$ is better or equal to $\hat{V}_R(\hat{Y})$ 7 times, while $\hat{V}_R(\hat{Y})$ is also better or equal to $\hat{V}_{EC}(\hat{Y})$ 7 times. With respect to relative bias, $\hat{V}_R(\hat{Y})$ is slightly better than $\hat{V}_{EC}(\hat{Y})$. $\hat{V}_R(\hat{Y})$ has smaller or equal relative biases 8 times, compared to $\hat{V}_{EC}(\hat{Y})$ being better or equal 5 times. Hence, when computation simplicity is important, $\hat{V}_R(\hat{Y})$ can be used. However, it may still be safer to use $\hat{V}_{EC}(\hat{Y})$.

When the nonresponse rate is low, table 3 shows that $\hat{V}_{EC}(\hat{Y})$ has better coverage rates than $\hat{V}_R(\hat{Y})$. $\hat{V}_{EC}(\hat{Y})$ is better or equal to $\hat{V}_R(\hat{Y})$ 10 times, while $\hat{V}_R(\hat{Y})$ is better or equal to $\hat{V}_{EC}(\hat{Y})$ 7 times. However, $\hat{V}_R(\hat{Y})$ has smaller relative biases. $\hat{V}_R(\hat{Y})$ is better or equal 9 times, while $\hat{V}_{EC}(\hat{Y})$ is better or equal 4 times. This probably means that it is safe to assume the terms subtracted in $\hat{V}_{EC}(\hat{Y})$ are close to zero and $\hat{V}_R(\hat{Y})$ can safely be used to estimate the total variance. However, since $\hat{V}_{EC}(\hat{Y})$ has better coverage rates, it should still be considered.

Tables 1-3 shows that $\hat{V}_{EC}(\hat{Y})$ and $\hat{V}_{TC}(\hat{Y})$ work equally well for relative biases. This means that for estimating the total variance, $\hat{\text{cov}}(\hat{d},\hat{y}')$ is a good approximation for $\text{cov}(\bar{y}_{\bullet}-\hat{y},\hat{y})$.

10. **Conclusions**

For the design and imputations simulated, $\hat{V}_{EC}(\hat{Y})$ should be used when the item nonresponse rate is high, although $\hat{V}_R(\hat{Y})$ is still an improvement over $\hat{V}(\hat{y}_{\bullet})$. When the item response rate is moderate or low $\hat{V}_R(\hat{Y})$ provides good results and given its simplicity, can be used to estimate the total variance.

Table - 1 Relative Bias (RB) and Coverage Rates (C_T and C_m) for Population with 30% Nonresponse

30% Nonresponse		Percent RB (C_T)		Percent RB (C_m)			
State	Estimate	$\hat{V}(\hat{y})$	$\hat{V}(\hat{y}_{\bullet})$	$\hat{V}(\hat{y}_{\bullet})$	$\hat{V}_{TC}(\hat{Y})$	$\hat{V}_{EC}(\hat{Y})$	$\hat{V}_R(\hat{Y})$
2	$\hat{y}_e$	0 (94)	-13 (90)	-13 (90)	0 (91)	9 (94)	2 (93)
	$\hat{y}_m$	1 (93)	-25 (81)	-25 (83)	-3 (90)	9 (93)	-3 (90)
	$\hat{y}_h$	1 (91)	-28 (80)	-28 (81)	-11 (85)	-5 (89)	-14 (86)
9	$\hat{y}_e$	1 (94)	-28 (83)	-28 (83)	2 (94)	0 (94)	-13 (89)
	$\hat{y}_m$	0 (94)	-29 (82)	-29 (82)	1 (93)	0 (93)	-14 (88)
	$\hat{y}_h$	1 (94)	-15 (88)	-15 (89)	3 (94)	-1 (93)	-7 (91)
10	$\hat{y}_e$	1 (91)	-22 (85)	-22 (85)	-7 (90)	-5 (91)	-13 (88)
	$\hat{y}_m$	0 (92)	-26 (83)	-26 (83)	-6 (91)	-3 (91)	-13 (88)
	$\hat{y}_h$	1 (87)	-31 (72)	-31 (73)	-1 (88)	-10 (83)	-21 (77)
24	$\hat{y}_e$	-1 (94)	-30 (82)	-30 (81)	-3 (93)	-1 (93)	-15 (88)
	$\hat{y}_m$	0 (93)	-29 (81)	-29 (81)	-2 (93)	2 (94)	-13 (88)
	$\hat{y}_h$	0 (93)	-13 (81)	-13 (90)	2 (93)	15 (96)	2 (93)

Table - 2 Relative Bias (RB) and Coverage Rates (C_T and C_m) for Population with 15% Nonresponse

15% Nonresponse		Percent RB (C_T)		Percent RB (C_m)			
State	Estimate	$\hat{V}(\hat{y})$	$\hat{V}(\hat{y}_\bullet)$	$\hat{V}(\hat{y}_\bullet)$	$\hat{V}_{TC}(\hat{Y})$	$\hat{V}_{EC}(\hat{Y})$	$\hat{V}_R(\hat{Y})$
2	$\hat{y}_e$	0 (94)	-6 (92)	-6 (92)	2 (93)	5 (95)	1 (94)
	$\hat{y}_m$	1 (93)	-9 (89)	-9 (91)	2 (92)	8 (93)	0 (92)
	$\hat{y}_h$	1 (91)	-1 (91)	-1 (92)	2 (91)	9 (93)	3 (92)
9	$\hat{y}_e$	1 (94)	-10 (90)	-10 (91)	4 (94)	2 (94)	-4 (92)
	$\hat{y}_m$	0 (94)	-11 (89)	-11 (90)	6 (95)	4 (94)	-4 (92)
	$\hat{y}_h$	1 (94)	-7 (92)	-7 (92)	7 (95)	3 (94)	-2 (93)
10	$\hat{y}_e$	1 (91)	-9 (89)	-9 (90)	2 (92)	3 (93)	-4 (91)
	$\hat{y}_m$	0 (92)	-10 (89)	-10 (91)	1 (93)	1 (94)	-5 (92)
	$\hat{y}_h$	-1 (87)	-10 (82)	-10 (86)	7 (90)	-3 (86)	-9 (86)
24	$\hat{y}_e$	-1 (94)	-14 (90)	-14 (90)	7 (94)	8 (94)	-4 (92)
	$\hat{y}_m$	0 (93)	-13 (90)	-13 (90)	8 (95)	11 (95)	-1 (93)
	$\hat{y}_h$	0 (93)	-5 (92)	-5 (91)	5 (93)	5 (93)	0 (93)

Table - 3 Relative Bias (RB) and Coverage Rates (C_T and C_m) for Population with 5% Nonresponse

5% Nonresponse		Percent RB (C_T)		Percent RB (C_m)			
State	Estimate	$\hat{V}(\hat{y})$	$\hat{V}(\hat{y}_\bullet)$	$\hat{V}(\hat{y}_\bullet)$	$\hat{V}_{TC}(\hat{Y})$	$\hat{V}_{EC}(\hat{Y})$	$\hat{V}_R(\hat{Y})$
2	$\hat{y}_e$	0 (94)	-4 (92)	-4 (93)	-1 (94)	-1 (94)	-2 (93)
	$\hat{y}_m$	1 (93)	-7 (88)	-7 (91)	11 (95)	8 (93)	0 (92)
	$\hat{y}_h$	1 (91)	-6 (88)	-6 (90)	-6 (87)	3 (91)	-2 (91)
9	$\hat{y}_e$	1 (94)	-3 (93)	-3 (93)	6 (95)	5 (95)	-1 (94)
	$\hat{y}_m$	0 (94)	-5 (92)	-5 (92)	2 (94)	3 (94)	-2 (93)
	$\hat{y}_h$	1 (94)	-1 (94)	-1 (94)	0 (93)	0 (93)	-1 (93)
10	$\hat{y}_e$	1 (91)	-2 (91)	-2 (91)	5 (92)	2 (91)	-1 (91)
	$\hat{y}_m$	0 (92)	-1 (92)	-1 (92)	2 (92)	2 (93)	0 (92)
	$\hat{y}_h$	-1 (87)	-2 (87)	-2 (87)	-2 (88)	-2 (87)	-2 (87)
24	$\hat{y}_e$	-1 (94)	-7 (91)	-7 (92)	1 (95)	1 (94)	-4 (93)
	$\hat{y}_m$	0 (93)	-5 (90)	-5 (92)	4 (94)	4 (93)	-2 (92)
	$\hat{y}_h$	0 (93)	1 (93)	1 (93)	1 (93)	2 (93)	1 (93)

References

Cochran, W.G. (1977), *Sampling Techniques*, New York: John Wiley & Sons, p253.

Rao, J.N.K., Shao, J (1992), "Jacknife Variance Estimation with Survey Data Under Hot Deck Imputation", *Biometrika*,79, pp. 811-822.

Rubin, D.B. (1978), *Multiple Imputation for Nonresponse in Surveys*, New York: John Wiley & Sons.

Särndal, C.E (1990), "Methods for Estimating the Precision of Survey Estimates When Imputation Has Been Used", *Proceedings of Statistics Canada's Symposium '90: Measurement and Improvement of Data Quality*, pp.369-380.

Särndal, C. E. (1994), "Estimation of the Variance in the Presents of Nearest Neighbor Imputation", *Proceedings of the Section on Survey Research Methods*, ASA Annual Meetings, pp. 888-893.

Wolter, K.M. (1985), *Introduction to Variance Estimation*, New York: Springer-Verlag.