Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 69(2) 2015 127-133

Estimation of Finite Population Total in Stratified Sampling Under Error-in-variables Super Population Model

B.V.S. Sisodia, Amar Singh, K.K. Mourya and V.N. Rai
N.D. University of Agriculture & Technology
Kumarganj, Faizabad, U.P.

Received 01 July 2014; Revised 16 April 2015; Accepted 21 April 2015

SUMMARY

In the present paper, an attempt has been made to examine the effect of measurement error in the study variate on the efficiency of the model-based estimators of finite population total under super population model when variance of the study variate, y, is a function of the auxiliary variable x, related to y, and included as an independent variable in the model. Simulation results show that there is considerable loss in the precision of the estimators due to measurement error. However, such losses are marginal if the variability in the measurement errors as compared to variability in model errors is small.

Keywords: Super population model, Measurement errors, Regression, Finite population, Prediction.

1. INTRODUCTION

Bolfarine (1991) considered finite population prediction under error-in-variables super population model. He considered two models: (i) location error-in-variables super population model, *i.e.*

$$y_i = \mu + e_i,\ i = 1, 2, ..., N$$
$$Y_i = y_i + v_i \tag{1.1}$$

where Y_i and y_i are the observed and true value of y, respectively, $e_i \sim iidN(0, \sigma_e^2)$, $v_i \sim iidN(0, \sigma_v^2)$, $\text{COV}(e_i, v_i) = 0$ and $V(y_i) = \sigma_e^2$ and (ii) simple regression model with measurement error as

$$y_i = \beta_0 + \beta_1 x_i + e_i,\ i = 1, 2, \ldots, N$$
$$v(y_i) = \sigma_e^2$$
$$Y_i = y_i + v_i \text{ and } X_i = x_i + u_i \tag{1.2}$$

where Y_i and X_i are observed values and, y_i and x_i are true values of y and x, respectively. $e_i \sim iidN(0, \sigma_e^2)$, $v_i \sim iidN(0, \sigma_v^2)$ and $u_i \sim iidN(0, \sigma_u^2)$. e_i, v_i and u_i are mutually independent. However, he assumed that the random sample comes from a bi-variate normal population of (y, x). Mukhopadhyay (1994) developed prediction estimators for finite population mean $\overline{Y}$ and S_y^2 (population mean square) under the model (1.1). He developed the optimal predictor for $\overline{Y}$ under the model (1.2). Chattopdhyay and Datta (1994) have extended the work of Bolfarine (1991) to the stratified sampling under the location error-in-variables super-population model given in (1.1). Various authors have made contribution on this aspect in recent past. Notably among them are Battese *et al.* (1988), Eltinge (1994), Stefanski (2000), Ghosh and Sinha (2007), Ma and Li (2010), West (2010), Arima *et al.* (2012) *etc*.

When we consider the model-based/model assisted estimation of finite population total or mean of the study variate y, it has been found generally in most of socio-economic surveys that variance of y is a function

Corresponding author: B.V.S. Sisodia
E-mail address: bvssisodia@gmail.com

of the auxiliary variable x related to y. The structure of the variance function is generally obtained as $V(y_i) = \sigma^2 x_i^g$, $1/2 \le g \le 2$, σ^2 is variance of error term in the model, for most of the data encountered in practice (see the work of Rao and Bayless 1969, Bayless and Rao 1970 *etc*). Also for instance, Royall (1970, 1971) and Royall and Herson (1973a, 1973b) have assumed $g = 1$. Scott *et al.* (1978) have considered $g = 2$. Royall and Herson (1973b) have extended their work of 1973a to the stratified sampling when stratification is done on size variable (x) by assuming the following super population model

$$y_{hi} = \beta_h x_{hi} + e_{hi} x_{hi}^{1/2},\ i = 1, 2, \ldots N_h,\ h = 1, 2, \ldots, H$$

$$V(y_{hi}) = \sigma_e^2 x_{hi} \tag{1.3}$$

They referred the model (1.3) as ξ - model, denoted as $\xi(0, 1; x_h)$. The strata were formed as follows: The N_1 units whose x-values are smallest form stratum 1, the next N_2 smallest units form stratum 2, and so on, such that $\sum_{h=1}^{H} N_h = N$. Thus, no unit in stratum h is larger than any units in stratum $h + 1$. They suggested a model-based separate ratio estimator of $T = \sum_{h=1}^{H}\sum_{i=1}^{N_h} y_{hi}$, the finite population total, under the model (1.3). The estimator of T they developed is

$$\hat{T} = \sum_{h=1}^{H} \hat{T}_h \tag{1.4}$$

where, $\hat{T}_h = \sum_{i \in s_h} y_{hi} + \hat{\beta}_h \sum_{i \in \bar{s}_h} x_{hi}$, $\hat{\beta}_h = \sum_{i \in s_h} y_{hi} / \sum_{i \in s_h} x_{hi}$, a best linear unbiased estimator (BLUE) of β_h, s_h is a sample of n_h units and $\bar{s}_h$ is compliment of s_h in stratum h such that $s_h \cup \bar{s}_h = N_h$ and $\sum_{h=1}^{H} n_h = n$. The model variance of $\hat{T}$ is given by

$$V(\hat{T}) = \sigma^2 \sum_{h=1}^{H} \frac{\sum_{i \in \bar{s}_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \sum_{i=1}^{N_h} x_{hi} \tag{1.5}$$

They also proved a theorem, which is stated below without proof

Theorem 1. If $n_h \propto N_h \bar{x}_h^{1/2}$, *i.e.* $n_h = \dfrac{n N_h \bar{x}_h^{1/2}}{\sum_{h=1}^{H} N_h \bar{x}_h^{1/2}}$, where $\bar{x}_h$ is stratum mean of x in stratum h, then under the general polynomial model $\xi(\delta_0, \delta_1, \ldots, \delta_J : x_h)$ of degree J, the strategy $\left[s^+(J), \hat{T}\right]$ is more efficient than the strategy $\left[s(J), \hat{T}(0,1:x)\right]$, where $\hat{T}(0,1:x)$ is the estimator of T under the model $\xi(0, 1 : x)$ when there is no stratification of the population. Note that $s(J)$ and $s^+(J)$ are simple balance and stratified balance sample (see, Royall and Herson 1973a, 1973b).

Kaushal (2007) considered the following super population model for stratified sampling where the slope (regression coefficient) is common across strata.

$$y_{hi} = \beta x_{hi} + e_{hi} x_{hi}^{1/2},\ i = 1, 2, \ldots, N_h,\ h = 1, 2, \ldots, H$$

$$V(y_{hi}) = \sigma_e^2 x_{hi} \tag{1.6}$$

He developed the estimator of T as

$$\hat{T}' = \sum_{h=1}^{H}\sum_{i \in s_h} y_{hi} + \hat{\beta} \sum_{h=1}^{H}\sum_{i \in \bar{s}_h} x_{hi} \tag{1.7}$$

where $\hat{\beta} = \sum_{h=1}^{H}\sum_{i \in s_h} y_{hi} / \sum_{h=1}^{H}\sum_{i \in s_h} x_{hi}$, as BLUE of β, with model variance

$$V(\hat{T}') = \sigma_e^2 \frac{\sum_{h=1}^{H}\sum_{i \in \bar{s}_h} x_{hi}}{\sum_{h=1}^{H}\sum_{i \in s_h} x_{hi}} X,\quad X = \sum_{h=1}^{H}\sum_{i=1}^{N_h} x_{hi} \tag{1.8}$$

The research question addressed in this paper is that when there is measurement error in y and x how it affects the precision of the estimators developed under model (1.3) and (1.6). However, it has been generally conceived that the auxiliary information on the auxiliary variables (x) related to the study variate (y) are obtained from administrative records and various other sources in most of the socio-economic surveys. Therefore, error in the magnitude of x may not be always found and even if there is little bit error, it may not pose serious problems, particularly, in estimating model parameters. But response error or measurement error in the study variable would certainty affects the estimate of finite population parameters as well as its standard error. Therefore, in the present paper, an attempt has been made to develop model based estimators for the finite population total when only the study variate is subject to measurement error under super population model (1.3) and (1.6). Section 2 deals with the development of model-based estimators and derivation of model variance of the estimators, etc. A limited simulation study to examine the effect of measurement error on the precision of the estimators under these two models has been conducted in Section 3. The concluding remarks are given in Section 4.

2. ESTIMATION OF POPULATION TOTAL IN STRATIFIED SAMPLING WHEN STUDY VARIATE IS SUBJECT TO MEASUREMENT ERROR

We consider that the stratification of the population is carried out on the basis of size variable (x). The strata are formed as described in Section 1. We consider the following error-in-variables super population model, referred to as model-I.

$$y_{hi} = \beta_h x_{hi} + e_{hi} x_{hi}^{1/2},\ i = 1, 2, ..., N_h,\ h = 1, 2, ...\ H$$

$$Y_{hi} = y_{hi} + v_{hi},\ V(y_{hi}) = \sigma_e^2 x_{hi} \tag{2.1}$$

$E(e_{hi}) = E(v_{hi}) = 0$, $V(v_{hi}) = \sigma_v^2$, $V(e_{hi}) = \sigma_e^2$ and, $COV(e_{hi}, v_{hi})$, $COV(v_{hi}, v_{hj})$ and $COV(e_{hi}, e_{hj})$ are zero for $i \neq j = 1, 2, ..., N_h$. β_h is the model parameter and Y_{hi} and y_{hi} are the observed and true value of y, respectively. The model (2.1) is also denoted as $\xi(0, 1: x)$. The objective is to estimate $T = \sum_{h=1}^{H}\sum_{i=1}^{N_h} y_{hi}$, the finite population total of y.

For a given sample s_h of size n_h units from stratum h, the stratum total T_h is given by

$$T_h = \sum_{i \in s_h} y_{hi} + \sum_{i \in \bar{s}_h} y_{hi}$$

A model based estimator of T_h is given by

$$\hat{T}_h = \sum_{i \in s_h} Y_{hi} + \sum_{i \in \bar{s}_h} \hat{y}_{hi} \tag{2.2}$$

where $\hat{y}_{hi} = \hat{\beta}_h x_{hi}$ and $\hat{\beta}_h = \dfrac{\sum_{i \in s_h} Y_{hi}}{\sum_{i \in s_h} x_{hi}}$, which is least square estimate of β_h under model (2.1).

An estimator of T is, therefore, given by

$$\hat{T}_1 = \sum_{h=1}^{H} \hat{T}_h \tag{2.3}$$

The model expectation of the estimator $\hat{T}_1$ is

$$E(\hat{T}_1 - T) = E\left[\sum_{h=1}^{H}\hat{T}_h - \sum_{h=1}^{H} T_h\right]$$

$$= E\left[\sum_{h=1}^{H}\left(\sum_{i \in s_h} Y_{hi} + \frac{\sum_{i \in s_h} Y_{hi}}{\sum_{i \in s_h} x_{hi}} \sum_{i \in \bar{s}_h} x_{hi} - \sum_{i=1}^{N_h} y_{hi}\right)\right]$$

$$= \sum_{h=1}^{H} E\left[\sum_{i \in s_h} Y_{hi}\left(\frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}}\right) - \sum_{i=1}^{N_h} y_{hi}\right]$$

$$= \sum_{h=1}^{H} E\left[\sum_{i \in s_h}\left(\beta x_{hi} + e_{hi} x_{hi}^{1/2} + v_{hi}\right)\left(\frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}}\right) - \sum_{i=1}^{N_h}\left(\beta_h x_{hi} + e_{hi} x_{hi}^{1/2}\right)\right]$$

= 0, as per underlying assumptions of model (2.1)

This shows that the estimator $\hat{T}_1$ is a model unbiased estimator of T.

The model variance of $\hat{T}_1$ is derived, which is obtained as follows.

$$V(\hat{T}_1) = \sum_{h=1}^{H} E(\hat{T}_1 - T)^2$$

$$= \sigma_e^2 \sum_{h=1}^{H} \left[\frac{\sum_{h=1}^{N_h} x_{hi} \sum_{i \in \bar{s}_h} x_{hi}}{\sum_{i \in s_h} x_i} + n_h \delta \left(\frac{\sum_{h=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \right)^2 \right] \quad (2.4)$$

where $\delta = \frac{\sigma_v^2}{\sigma_e^2}$

2.1 Optimum Allocation of Sample Sizes under Model 2.1

Under proportional allocation, $n_h = \frac{nN_h}{N}$, the variance expression of $\hat{T}_1$ given in (2.4) reduces to

$$V(\hat{T}_1)_P = \sigma_e^2 \sum_{h=1}^{H} \left[\frac{N-n}{n} \frac{X_h \bar{x}_{\bar{s}_h}}{\bar{x}_{s_h}} + \delta \frac{N}{nN_h} \frac{X_h^2}{\bar{x}_{s_h}^{\,2}} \right] \quad (2.1.1)$$

where X_h is stratum total of x; and $\bar{x}_{s_h}$ and $\bar{x}_{\bar{s}_h}$ are sample and non sample mean of x in h^{th} stratum, respectively.

Subject to fixed total cost $c = c_0 + \sum_{h=1}^{H} c_h n_h$, c_0 is the overhead cost and c_h is the cost of enumeration for each unit sampled in the stratum h, the variance (2.4) is minimized. For this, we consider the function

$$\phi = \sum_{h=1}^{H} \left[\sigma_e^2 \sum_{i=1}^{N_h} x_{hi} \left(\frac{\sum_{i=1}^{N_h} x_{hi} - \sum_{i \in s_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \right) + n_h \left(\frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \right)^2 \sigma_v^2 \right]$$

$$+\lambda\left(c - c_0 - c_h n_h\right) \quad (2.1.2)$$

where λ is Lagrangian multiplier.

For stratified balanced sampling, *i.e.* $\bar{x}_{s_h} = \bar{x}_h = \bar{x}_{\bar{s}_h}$, the function ϕ in (2.1.2) reduces to

$$\phi = \sum_{h=1}^{H} \left[N_h \bar{x}_h \sigma_e^2 \left(\frac{N_h \bar{x}_h - n_h \bar{x}_h}{n_h \bar{x}_h} \right) + n_h \left(\frac{N_h}{n_h} \right)^2 \sigma_v^2 \right]$$

$$+\lambda(c - c_o c_h n_h) \quad (2.1.3)$$

Differentiating ϕ in (2.1.3) with respect to n_h and equating it zero and after solving it, we get

$$n_h = n \frac{N_h \sqrt{\sigma_e^2 \bar{x}_h + \sigma_v^2} / \sqrt{c_h}}{\sum_{h=1}^{H} N_h \sqrt{\sigma_e^2 \bar{x}_h + \sigma_v^2} / \sqrt{c_h}} \quad (2.1.4)$$

If $c_h = c$ for all $h = 1, 2, \ldots, H$, then (2.1.3) reduces to

$$n_h = n \frac{N_h \sqrt{\bar{x}_h + \delta}}{\sum_{h=1}^{H} N_h \sqrt{\bar{x}_h + \delta}} \quad (2.1.5)$$

The allocation in (2.1.5) is also similar to Neyman allocation. Thus, when the allocation (2.1.4) is used the variance (2.4) simplifies to

$V(\hat{T}_1)_N$

$$= \sum_{h=1}^{H} \left[\sigma_e^2 \frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \sum_{i \in \bar{s}_h} x_{hi} + n \frac{N_h \sqrt{\sigma_e^2 \bar{x}_h + \sigma_v^2}}{\sum_{h=1}^{H} N_h \sqrt{\sigma_e^2 \bar{x}_h + \sigma_v^2}} \left(\frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \right)^2 \sigma_v^2 \right] \quad (2.1.6)$$

The above expression of the variance can also be written as

$V(\hat{T}_1)$

$$= \sigma_e^2 \sum_{h=1}^{H} \left[\frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \sum_{i \in \bar{s}_h} x_{hi} + n\delta \frac{N_h \sqrt{\bar{x}_h + \delta}}{\sum_{h=1}^{H} N_h \sqrt{\bar{x}_h + \delta}} \left(\frac{\sum_{i=1}^{N_h} x_{hi}}{\sum_{i \in s_h} x_{hi}} \right)^2 \right] \quad (2.1.7)$$

where $\delta = \sigma_v^2 / \sigma_e^2$, and n_h involved in sample and non sample total will be replaced by n_h obtained from allocation (2.1.5).

2.2. A Special Case when Slopes (Regression Coefficients) are Common Across the Strata

Under this situation, the model (2.1) reduces to

$$y_{hi} = \beta x_{hi} + e_{hi} x_{hi}^{1/2}, \quad i = 1, 2, \ldots, N_h,$$
$$h = 1, 2, \ldots, H$$

$$Y_{hi} = y_{hi} + v_{hi},\ V(y_i) = \sigma_e^2 x_{hi} \quad (2.2.1)$$

where other notations and assumptions stand as usual as in the model (2.1).

The estimator $\hat{T}_1$ in (2.3) reduces to

$$\hat{T}_2 = \sum_{h=1}^{H} \sum_{i \in s_h} Y_{hi} + \sum_{h=1}^{H} \sum_{i \in \bar{s}_h} \hat{y}_{hi} \quad (2.2.2)$$

where $\hat{\beta} = \dfrac{\sum_{h=1}^{H} \sum_{i \in s_h} Y_{hi}}{\sum_{h=1}^{H} \sum_{i \in s_h} x_{hi}}$, which is obtained after the fitting of the model (2.2.1) with data contained in $s = \bigcup_{h=1}^{H} s_h$ by least square technique.

It can easily be verified that $\hat{T}_2$ is also model unbiased estimator of T with variance

$$V\left(\hat{T}_2\right) = \sigma_e^2 \left[\frac{X \sum_{h=1}^{H} \sum_{i \in \bar{s}_h} x_{hi}}{\sum_{h=1}^{H} \sum_{i \in s_h} x_{hi}} \right] + \left[n\sigma_v^2 \left(\frac{X}{\sum_{h=1}^{H} \sum_{i \in s_h} x_{hi}} \right)^2 \right] \quad (2.2.3)$$

which can further be expressed as

$$V\left(\hat{T}_2\right) = \sigma_e^2 \frac{\sum_{h=1}^{H} (N_h - n_h) \bar{x}_{\bar{s}_h}}{\sum_{h=1}^{H} n_h \bar{x}_{s_h}} X + n \frac{\sigma_v^2 (X)^2}{\left(\sum_{h=1}^{H} n_h \bar{x}_{s_h} \right)^2} \quad (2.2.4)$$

For stratified balanced sample (Royall and Herson, 1973b), we have $\bar{x}_{\bar{s}nh} = \bar{x}_{s_h} = \bar{x}_h$ and in this case the expression (2.2.4) reduces to

$$V(\hat{T}_2) = \sigma_e^2 X \left[\left(\frac{\sum_{h=1}^{H} N_h \bar{x}_h}{\sum_{h=1}^{H} n_h \bar{x}_h} - 1 \right) + n\delta \frac{X}{\left(\sum_{h=1}^{H} n_h \bar{x}_h \right)^2} \right] \quad (2.2.5)$$

Subject to fixed total cost $c = c_0 + \sum_{h=1}^{H} c_h n_h$, as mentioned in sub-section (2.1), the optimum value of n_h is obtained as

$$n_h = \frac{n N_h \bar{x}_h^{1/2} / \sqrt{c_h}}{\sum_{h=1}^{H} N_h \bar{x}_h^{1/2} / \sqrt{c_h}} \quad (2.2.6)$$

If $c_h = c$ for all h = 1, 2, …, H, then

$$n_h = \frac{n N_h \bar{x}_h^{1/2}}{\sum_{h=1}^{H} N_h \bar{x}_h^{1/2}} \quad (2.2.7)$$

The allocation given in (2.2.7) is similar to Neyman allocation.

The variance of $\hat{T}_2$ given in (2.2.5) under allocation (2.2.7) is obtained as

$$V(\hat{T}_2)_N = \frac{\sigma_e^2}{n} \frac{\sum_{h=1}^{H} N_h \bar{x}_h}{\sum_{h=1}^{H} N_h \bar{x}_h^{3/2}} \times \left[\left\{ \left(\sum_{h=1}^{H} N_h \bar{x}_h \right) \left(\sum_{h=1}^{H} N_h \bar{x}_h^{1/2} \right) - \sum_{h=1}^{H} N_h \bar{x}_h^{3/2} \right\} + N\delta \sum_{h=1}^{H} N_h \bar{x}_h^{1/2} \right] \quad (2.2.8)$$

The variance of $\hat{T}_2$ given in (2.2.5) under proportional allocation, *i.e.* $n_h = n.N_h/N$, is obtained as

$$V(\hat{T}_2)_P = \frac{\sigma_e^2}{n} \left[(N - n) \sum_{h=1}^{H} N_h \bar{x}_h + N^2 \delta \right] \quad (2.2.9)$$

3. A LIMITED SIMULATION STUDY

In this section we provide simulation results illustrating percent relative increase in standard error of the estimators when the study variate is subject to measurement error. The population data were generated using the model (2.1) and (2.2.1). The values for x were generated assuming x to follow Chi-square distribution with 5 degree of freedom. It is assumed that e_{hi} were independently normally distributed with mean zero and variance 2. So e_{hi}'s were generated by normal distribution with mean 0 and variance 2. The value of

β_h were assumed to be 0.5, 0.8, 0.9, 1.3 and 1.4 for model (2.1) and value of β is assumed to be 0.5 for model (2.2.1). The population for $N = 1500$ of y were generated using these models. Then, five strata of size 100, 200, 300, 400 and 500 were formed according to the increasing values of x. A stratified sample of size $n = 150$ with samples of sizes 10, 20, 30, 40 and 50 from respective stratum were drawn by simple random sampling without replacement (SRSWOR) using proportional allocation. We also draw a sample of size $n = 150$ with samples of sizes 13, 13, 24, 38 and 62 for model (2.1) under allocation (2.1.5) and with samples of sizes 5, 13, 24, 40 and 68 for model (2.2.1) under (2.2.7) by SRSWOR. Each scenario was independently replicated 50000 times. For each scenario (simulation run $i = 1, 2, \ldots, 50000$), variances of the estimators were computed. Let V_i and V_i' be the variance of the estimators without measurement error and with measurement error, respectively, for i^{th} simulation. Let $\bar{V}$ and $\bar{V}'$ be the average of V_i and V_i'. The percent relative increase (%R.I.) in standard error of the estimator due to measurement error has been computed as follows

$$\%R.I. = \frac{\sqrt{\bar{V}'} - \sqrt{\bar{V}}}{\sqrt{\bar{V}}} \times 100 \tag{3.1}$$

The simulation is carried out using R-software. The simulation results under model (2.1) with proportional allocation and allocation (2.1.5) are presented in Table 3.1.

Table 3.1. Variance of the estimator and %R.I. in standard error with proportional allocation and allocation (2.1.5) under model (2.1).

S.No.	Allocation	$\bar{V}$	$\bar{V}'$ and %R.I. for different values of δ		
			0.75	1.00	1.25
1	Proportional	139910.68	162534.43	170075.68	177616.93
		(374.05)	(403.16)	(412.40)	(421.45)
		%RI	7.78	10.25	12.67
2	Allocation (2.1.4)	135102.12	151232.75	155006.75	158780.80
		(367.56)	(388.89)	(393.71)	(398.47)
		%RI	5.80	7.11	8.41

NB: The figures in the parentheses denote the average standard error of the estimators.

It is obvious from the results of Table 3.1 that the precision of the estimator of the finite population total gets affected considerably when the study variate is subject to measurement error or response error. The percent relative increase in standard error of the estimator is relatively less if allocation (2.1.5) is employed instead of proportional allocation in stratified sampling. The%RI depends on the values of $\delta = \sigma_v^2 / \sigma_e^2$, and it decreases if δ decreases. That means if the variability in the measurement error (σ_v^2) is relatively small as compared to variability in the model error, then %RI is expected to be marginal one.

The simulation results under model (2.2.1) are presented in Table 3.2.

Table 3.2. Variance of the estimator and %R.I. in standard error with proportional allocation and allocation (2.2.7) under model (2.2.1).

S.No.	Allocation	$\bar{V}$	$\bar{V}'$ and %R.I. for different values of δ		
			0.75	1.00	1.25
1	Proportional	139793.73	162346.68	169864.33	177381.98
		(373.89)	(402.92)	(412.15)	(421.17)
		%RI	7.76	10.23	12.64
2	Allocation (2.2.7)	118449.92	135387.82	141033.78	146679.75
		(344.17)	(367.95)	(375.54)	(382.99)
		%RI	6.90	9.11	11.28

NB: The figures in the parentheses denote the average standard error of the estimators.

The results of Table 3.2 reflect the same trend in terms of variances and %RI in standard error of the estimate due to measurement error as it has been observed in Table 3.1. The %RI in standard errors of the estimators with proportional allocation under both the models have been obtained almost similar. However, with allocation (2.2.7), %RI under model (2.1) has been found quite smaller than that of under model (2.2.1). Note that the slope under the model (2.2.1) is assumed to be same across the strata. Therefore, the model (2.1) is preferable to the model (2.2.1), and it is also justifiable that the slope may not be same across the strata in practice. In case the slope is common across the strata, a single super population can be assumed for the development of the estimator of finite population total with and without measurement error in the study variate (see, the work of Singh *et al.* 2014). It may also be noted that the estimator of population total under model (2.1) is similar to separate ratio estimator (Royall and Herson 1973b).

4. CONCLUDING REMARKS

An attempt has been made in the present paper to examine the effect of measurement error in the study variate on the precision of the estimators developed in stratified sampling under the super population models when (i) slopes varies from stratum to stratum and (ii) slope is common across strata. Minimum loss in the precision has been obtained under model (2.1) with allocation (2.1.5) and (2.2.7). However, %RI in standard error of the estimator is still substantial, which affects largely the precision of the estimator. Therefore, it is imperative for survey statisticians that they provide proper training and guidance to the field investigators responsible for survey work for colleting reliable data with maximum probability of without collecting response error/measurement error in order to obtain precise estimate of the population parameters, otherwise standard error of the estimate would be inflated.

ACKNOWLEDGEMENTS

The authors are highly thankful to the referee of the paper for his critical suggestions which have improved the earlier version of the paper. The authors are also thankful to the Coordinating Editor of JISAS for his fruitful suggestions.

REFERENCES

Arima, S., Datta, G.S. and Liseo, B. (2012). Objective bayesian analysis of a measurement error small area model. *Bayesian Anal.*, **7(2)**, 363-384.

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error components models for predicition of county crop areas using survey and satellite data. *J. Amer. Statist. Assoc.,* **83**, 28-36.

Bayless, D.L. and Rao, J.N.K. (1970). An empirical study of stabilities of estimators and variance estimators in unequal probabilities sampling (n = 3 or 4). *J. Amer. Statist. Assoc.*, **65,** 645-667.

Bolfarine, H. (1991). Finite population prediction under error-in-variables super population models. *Canad. J. Statist.,* **19(2)**, 191-207.

Chattopadhyay, Manas K. and Datta, G.S. (1994). Finite population prediction for stratified sampling under error-in-variables superpopulation models. *Sankhya* : *The Indian Journal of Statistics*, **56(B)**,. 415-433.

Eltinge L. Jhon (1994). A finite population approach to approximating small measurement error effects in regression. *Sankhya, The Indian Journal of Statistics.* **56(B)**, 234-250.

Ghosh, M. and Sinha, K. (2007). Empirical bayes estimation in finite population sampling under functional measurement error models. *J. Statist. Plann. Inf.,* **137,** 2759-2773.

Kaushal, R.P. (2007). Robust Estimation in Stratified Sampling under Super-population Model. (Ph.D. Thesis), N.D.U.A.&T. Kumarganj, Faizabad.

Ma, Yanyuan and Li, Runze (2010). Variable selection in measurement error models. *Bernoulli*, **16(1)**, 274-300.

Mukhopadhyay, P. (1994). Prediction in finite population under error-in-variables superpopulation models. *J. Statist. Plann. Inf.*, **41,** 151-161.

Rao, J.N.K. and Bayless, D.L. (1969). An empirical study of stabilities of estimators and variance estimators in unequal probability sampling of two units per stratum. *J. Amer. Statist. Assoc.*, **64**, 540-559.

Royall, R.M. (1970). On finite population sampling theory under certain linear regression models. *Biometrika*, **57(2)** 377-387.

Royall, R.M. (1971). Linear regression models in finite population sampling theory. In: *Foundations of Statistical Inference Toronto Holt*, Godambe, V.P. and Sprott, D.A. Eds., Rinehart and Winston of Canada Ltd.

Royall, R.M. and Herson, J. (1973a). Robust estimation in finite populations. *J. Amer. Statist. Assoc.*, **68,** 880-889.

Royall, R.M. and Herson, J. (1973b). Robust estimation in finite populations II. Stratification on a size variable. *J. Amer. Statist. Assoc.*, **68,** 890-893.

Scott, A.J., Brewer, K.R.W. and Ho, E.W.H. (1978). Finite population sampling and robust estimation. *J. Amer. Statist. Assoc.*, **73**, 359-361.

Singh, A., Sisodia, B.V.S., and Mourya, K.K. (2014) ,Estimation of finite population total under error-in-variables super population. *Statistics* (Submitted for publication).

Stefanski, L.A. (2000). Measurement error model. *J. Amer. Statist. Assoc.*, **95(452),** 1353-1358.

West, T. Brady (2010). The impact of measurement error in auxiliary variables on model-based estimation of finite population totals: A simulation study. *Social Statist. Sec.-JSM.*