Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online): 2229-5666
Vol. 11, Issue 2 (2018): 57-68

MODEL BASED PREDICTION OF FINITE POPULATION TOTAL UNDER SUPER POPULATION MODEL

Shweta Chauhan and B.V.S. Sisodia*
Department of Agricultural Statistics
N.D. University of Agriculture & Technology
Kumarganj, Faizabad, U.P.
E Mail: sstat@gmail.com; bvssisodia@gmail.com
***Corresponding author**

Received November 03, 2017
Modified September 18, 2018
Accepted October 28, 2018

Abstract

In the present paper, we have proposed a predictor of finite population total under super population model when the study variate is subjected to measurement error. The results of the simulation study revealed that the percent loss in precision of the predictor varied between 26% and 40% depending upon the sample size and ratio of variance of measurement error to model error variance. The robustness of the predictor has also been examined when the assumed model deviates with measurement error in y.

Key Words: Predictor, Robustness, Super Population Model, Finite Population, Overbalanced sample, Measurement Error.

1. Introduction

In finite population survey sampling, the data on the study variate (y) are generally collected by personally interviewing the respondents selected in the sample on recall basis or sample units are measured by certain instrument to obtain the values of the variables of interest. Therefore, there is likelihood of response error or measurement error in sample surveys. If the data collected are subject to response error or measurement error, the precision of the estimator of finite population parameters is expected to be inflated in practice.

If the estimation procedure is model based/model assisted, the data on the auxiliary variable x related to y may also be subject to measurement error. Bolfarine (1991) dealt with prediction in finite population under error-in-variables super population models where he assumed that samples come from bivariate normal population and variance of error terms are constant. Mukhopadhyay (1994) has developed predictor for finite population mean under simple linear regression model with constant error variance when study variate y and an auxiliary variable x are subject to measurement error in sample surveys. Chattopadhyay and Datta (1994) extended the work of Bolfarine (1991) to stratified sampling under the location error-in-variables super population model. Some contribution on this aspect have been made by Eltinge (1994), Stefanski (2000), Ghosh and Sinha (2007), Arima *et al.* (2012) etc.

In most of the economic surveys, the variance of y is generally a function of the auxiliary variable x related to y, i.e. $V(y) = \sigma^2 x^g$, $1/2 \le g \le 2$, where σ^2 is model-error variance (see the work of Smith; 1938, Jessen; 1942, Desh Raj; 1958, Rao and Bayless; 1969, Bayless and Rao; 1970). In view of this fact, Royall and Herson (1970, 1971, 1973a) developed prediction approach based estimation of finite population total in finite population survey sampling under the following super population model

$$y_i = \beta x_i + e_i [v(x_i)]^{1/2}, \ \text{i}= 1,2,\ldots,\text{N}, \qquad V(y_i) = \sigma_e^2 v(x_i) \tag{1.1}$$

where y_i's are independent random variables, e_i's are error terms distributed independently with mean zero and variance σ_e^2. β is model parameter. They denoted this model as $\xi[0,1 : v(x)]$.They showed that the predictor $\hat{T} = \sum_{i \in s} y_i + \sum_{i \in \bar{s}} \hat{y}_i$ of finite population total $T = \sum_{i \in s} y_i + \sum_{i \in \bar{s}} y_i$ is best linear unbiased estimator (blue) under the model (1.1) for a given sample s of size n from the population consisting of N units. $\hat{y}_i$ is predicted value of y_i for non sampled units of the population through fitting of the model(1.1) by least square theory. $\bar{s}$ is the complement of s. It was shown by them that the estimator $\hat{T}$ reduces to the usual ratio estimator for v(x)=x. It was also shown by them that the unbiasedness of the ratio estimator can be preserved under the general polynomial regression model of degree J, i.e. $\xi[\delta_o, \delta_1, \ldots., \delta_J : v(x)]$ by choice of a balanced sample. The symbol δ_j is indicator variable and takes values 1 or zero according to presence or absence of the term $\beta_j x^j$, j=0,1,2,…,J, respectively, in the model. They further demonstrated that the ratio estimator remained blue even under the model $\xi(\delta_o, \delta_1, \ldots., \delta_J : x)$ for a balanced sample. Royall and Herson (1973b) extended this work to stratified sampling. Scott *et al.* (1978) extended the work of Royall and Herson (1973a) and showed that the estimator $\hat{T}$ under the model (1.1) remained blue under the model $\xi[\delta_o, \delta_1, \ldots., \delta_J : v^*(x)]$ for any variance function $v^*(x) = v(x) \sum_{j=0}^{J} \delta_j a_j x_i^{j-1}$, if the selected sample s satisfies the following condition

$$\frac{\sum_{\bar{s}} x_i^j}{\sum_{\bar{s}} x_i} = \frac{\sum_{s} x_i^{j+1} / v(x_i)}{\sum_{s} x_i^2 / v(x_i)}, \quad \text{j} = 0,1,2,\ldots,\text{J}$$

The sample satisfies the above condition was referred to as overbalanced sample by them. They further showed that under the model $\xi(0,1:x^2)$, the estimator $\hat{T}$ reduces to

$$\hat{T} = \sum_{i\in s} y_i + \frac{1}{n}\sum_{i\in s}\frac{y_i}{x_i}\sum_{i\in \bar{s}} x_i \ , \quad \text{with model variance}$$

$$V\left[\hat{T}\right] = \sigma_e^2 (N-n)\left[\left(\bar{x}_{\bar{s}}^{(2)} + \frac{(N-n)}{n}\bar{x}_{\bar{s}}^2\right)\right] \tag{1.2}$$

where $$\bar{x}_{\bar{s}}^{(2)} = \frac{1}{N-n}\sum_{\bar{s}} x_i^2 \quad and \quad \bar{x}_{\bar{s}}^2 = \left(\frac{1}{N-n}\sum_{\bar{s}} x_i\right)^2$$

Sisodia *et.al.* (2015) and Singh *et al.* (2018) have developed model based estimators of finite population total under the following error-in-variables super population model when study variate is only subject to measurement error in stratified and single phase sampling, respectively.

$$y_i = \beta x_i + e_i x_i^{1/2}, \ \text{i=1, 2, …, N}$$
$$Y_i = y_i + u_i, \ V(y_i) = \sigma_e^2 x_i \tag{1.3}$$

where y_i's are independent random variables. Y_i and y_i are observed and true value of y, respectively. The model error e_i and measurement error u_i are mutually independently distributed with mean zero and variances σ_e^2 and σ_v^2, respectively. They demonstrated through simulation study with hypothetical and real data that the standard error of the estimator got inflated by 4 to 8 percent due to measurement error in y. Chauhan and Sisodia (2018) and Chauhan *et al.* (2018) have studied the robustness of the estimators developed by Sisodia *et al.* (2015) Singh *et. al.* (2018).

In view of the above discussion on the recent literature on the topic and motivated by the model considered by Scott et al(1978) , an attempt has been made in the present paper to develop predictor for finite population total under the model (1.3) replacing the error term by $e_i x_i$ and $V(y_i) = \sigma^2 x_i$ by $V(y_i) = \sigma_e^2 x_i^2$, i.e. the model now becomes $\xi(0,1:x^2)$, when study variate is subject to measurement error. Robustness of the predictor has also been examined if the assumed model deviates. Some simulation studies have been conducted to examine the extent of loss in precision of the predictor due to measurement error in y.

2. Predictor of finite population total under the model $\xi(0,1:x^2)$ when the study variate is subject to measurement error

We consider the following error-in-variable super population model

$$y_i = \beta x_i + e_i x_i, \quad Y_i = y_i + v_i, \quad v(y_i) = \sigma_e^2 x_i^2 \ , \ \text{i=1,2,…,N} \tag{2.1}$$

where notations and assumptions are mentioned in (1.3).

The objective is to predict $T = \sum_{i=1}^{N} y_i$. Consider that a sample s of size n, by whatever manner not necessarily by probability sampling, is drawn from the finite population consisting of N units. The population total T can be decomposed as $T = \sum_{i \in s} y_i + \sum_{i \in \bar{s}} y_i$, where S contains the units in a given sample and s contains the non- sampled units from the population. A predictor of T is, therefore, given by

$$\hat{T}_1 = \sum_{i \in s} Y_i + \hat{\beta} \sum_{i \in \bar{s}} x_i \tag{2.2}$$

where $\hat{\beta} = \frac{1}{n} \sum_{i \in s} \frac{Y_i}{x_i}$, which is obtained after the fitting of the model (2.1) with data contained in s by least square technique. Taking the model expectation of the predictor $\hat{T}_1$, we get

$$E_\xi\left[\hat{T}_1 - T\right] = E_\xi\left[\sum_{i \in s} Y_i + \frac{1}{n} \sum_{i \in s} \frac{Y_i}{x_i} \sum_{i \in \bar{s}} x_i - T\right]$$

$$= E_\xi\left[\sum_{i \in s} (\beta x_i + e_i x_i + v_i) + \frac{1}{n} \sum_{i \in s} \frac{(\beta x_i + e_i x_i + v_i)}{x_i} \sum_{i \in \bar{s}} x_i - \sum_{i=1}^{N} (\beta x_i + e_i x_i)\right]$$

$$E_\xi\left[\hat{T}_1 - T\right] = \left[\beta\left(\sum_{i \in s} x_i + \sum_{i \in \bar{s}} x_i\right) - \beta \sum_{i=1}^{N} x_i\right] = 0 \tag{2.3}$$

This shows that estimator is model unbiased estimator of T, even if there is measurement error in observing y_i. We derive the model variance of $\hat{T}_1$ as follows

$$V\left[\hat{T}_1\right] = E_\xi\left[\hat{T}_1 - T\right]^2$$

$$= E_\xi\left[\sum_{i \in s} Y_i + \frac{1}{n} \sum_{i \in s} \frac{Y_i}{x_i} \sum_{i \in \bar{s}} x_i - T\right]^2$$

$$= E_\xi\left[\sum_{i \in s} (\beta x_i + e_i x_i + v_i) + \frac{1}{n} \sum_{i \in s} \frac{(\beta x_i + e_i x_i + v_i)}{x_i} \sum_{i \in \bar{s}} x_i - \sum_{i=1}^{N} (\beta x_i + e_i x_i)\right]^2$$

$$V\left[\hat{T}_1\right] = (N-n)\sigma_e^2\left[\bar{x}_{\bar{s}}^{(2)} + \frac{(N-n)}{n}\bar{x}_{\bar{s}}^2\right] + \sigma_v^2\left[n + \frac{(N-n)}{n}\left\{(N-n)\bar{x}_{\bar{s}}^2\bar{x}_s^{(-2)} + 2n\bar{x}_{\bar{s}}\bar{x}_s^{(-1)}\right\}\right] \tag{2.4}$$

where $\quad \bar{x}_s^{(-2)} = \frac{1}{n}\sum_s \frac{1}{x_i^2}, \bar{x}_{\bar{s}} = \frac{1}{N-n}\sum_{\bar{s}} x_i$ and $\bar{x}_s^{(-1)} = \frac{1}{n}\sum_s \frac{1}{x_i}$

For $\delta = \sigma_v^2 / \sigma_e^2$, the above expression (2.4) can alternatively be expressed as

$$V\left[\hat{T}_1\right] = \sigma_e^2\left[(N-n)\left(\bar{x}_{\bar{s}}^{(2)} + \frac{(N-n)}{n}\bar{x}_{\bar{s}}^2\right)\right] + \delta\left[n + \frac{(N-n)}{n}\left\{(N-n)\bar{x}_{\bar{s}}^2\bar{x}_s^{(-2)} + 2n\bar{x}_{\bar{s}}\bar{x}_s^{(-1)}\right\}\right] \tag{2.5}$$

It can be observed from the variance expression of $\hat{T}_1$ given in (2.4) that its first term is the variance of $\hat{T}$ when there is no measurement error in y_i. The second term of (2.4) will be always positive and this is due to measurement error. It shows that if there is measurement in y, there will be loss in precision of the model based estimator $\hat{T}_1$, although it remains model unbiased. The extent of loss in precision, however, depends on the magnitude of $\delta = \sigma_v^2 / \sigma_e^2$. Therefore, two important conclusions are as follows :(i) If $\sigma_v^2 = 0$, i.e., there is no variation in measurement error in y, there will be no loss in precision. That means if same magnitude of measurement error is committed in observing $y_i \in s$, then the estimator $\hat{T}_1$ will remain unbiased and it will have same variance as it is in case of $V\left[\hat{T}\right]$ given in (1.2) indicating thereby no loss in precision, and (ii) if σ_v^2 is relatively small in comparison to σ_e^2 (model-error), then the loss in precision will be relatively smaller. The extent of loss in precision will be studied in later section by conducting a limited simulation study.

3. Robustness of $\hat{T}_1$ when model $\xi(0,1:x^2)$ deviates and the study variate is subject to measurement error

A model based unbiased predictor $\hat{T}_1$ has been developed under the model (2.1) in the preceding section, which is reproduced here with variance

$$\hat{T}_1 = \sum_{i \in s} Y_i + \frac{1}{n}\sum_{i \in s}\frac{Y_i}{x_i}\sum_{i \in \bar{s}} x_i \tag{3.1}$$

$$V\left[\hat{T}_1\right] = (N-n)\sigma_e^2\left[\bar{x}_{\bar{s}}^{(2)} + \frac{(N-n)}{n}\bar{x}_{\bar{s}}^2\right] + \sigma_v^2\left[n + \frac{(N-n)}{n}\left\{(N-n)\bar{x}_{\bar{s}}^2\bar{x}_s^{(-2)} + 2n\bar{x}_{\bar{s}}\bar{x}_s^{(-1)}\right\}\right] \tag{3.2}$$

Suppose that model $\xi(0,1:x^2)$ is not true but the true model is $\xi\left(1,1:x^2\right)$,i.e.

$y_i = \alpha + \beta x_i + e_i x_i$, $Y_i = y_i + v_i$, i=1,2…….,N (3.3)

$v(y_i) = \sigma_e^2 x_i^2$, $v(e_i) = \sigma_e^2$ and $E(e_i) = E(v_i) = 0$ for all i where other notations and terms in the aforesaid model are already defined in (1.3) and α is y-intercept. We wish to examine the property of $\hat{T}_1$ under the model $\xi(1,1:x^2)$. We derive the model expectation of $\hat{T}_1$ under the model $\xi(1,1:x^2)$ as follows

$$E_\xi\left[\hat{T}_1 - T\right] = E_\xi\left[\sum_{i\in s} Y_i + \frac{1}{n}\sum_{i\in s}\frac{y_i}{x_i}\sum_{i\in \bar{s}} x_i - T\right]$$

$$= E_\xi\left[\sum_{i\in s}(\alpha+\beta x_i + e_i x_i + v_i) + \frac{1}{n}\sum_{i\in s}\frac{(\alpha+\beta x_i + e_i x_i + v_i)}{x_i}\sum_{i\in \bar{s}} x_i - \sum_{i=1}^{N}(\alpha+\beta x_i + e_i x_i)\right]$$

$$= E_\xi\left[(n-N)\alpha + \beta\sum_{i\in s} x_i + \sum_{i\in s}(e_i x_i + v_i) + \beta\sum_{i\in \bar{s}} x_i + \frac{1}{n}\alpha\sum_{i\in s}\frac{1}{x_i} + \frac{1}{n}\sum_{i\in s}\frac{(e_i x_i + v_i)}{x_i}\sum_{i\in \bar{s}} x_i - \beta\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} e_i x_i\right]$$

$$= (N-n)\alpha\left[\bar{x}_s^{(-1)}\bar{x}_{\bar{s}} - 1\right] \tag{3.4}$$

Obviously $\hat{T}_1$ is biased estimator if it is used in $\xi(1,1:x^2)$. The mean square error(MSE) of $\hat{T}_1$ under the model $\xi(1,1:x^2)$ is obtained as

$$MSE\left[\hat{T}_1\right] = E_\xi\left[\hat{T}_1 - T\right]^2$$

$$= E_\xi\left[\sum_{i\in s} Y_i + \frac{1}{n}\sum_{i\in s}\frac{Y_i}{x_i}\sum_{i\in \bar{s}} x_i - T\right]^2$$

$$= E_\xi\left[\sum_{i\in s}(\alpha+\beta x_i + e_i x_i + v_i) + \frac{1}{n}\sum_{i\in s}\frac{(\alpha+\beta x_i + e_i x_i + v_i)}{x_i}\sum_{i\in \bar{s}} x_i - \sum_{i=1}^{N}(\alpha+\beta x_i + e_i x_i)\right]^2$$

$$= E_\xi\left[(n-N)\alpha + \beta\sum_{i\in s} x_i + \sum_{i\in s}(e_i x_i + v_i) + \beta\sum_{i\in \bar{s}} x_i + \frac{1}{n}\alpha\sum_{i\in s}\frac{1}{x_i} + \frac{1}{n}\sum_{i\in s}\frac{(e_i x_i + v_i)}{x_i}\sum_{i\in \bar{s}} x_i - \beta\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} e_i x_i\right]^2$$

$$= \left[(N-n)\alpha\left(\bar{x}_s^{(-1)}\bar{x}_{\bar{s}} - 1\right)\right]^2 + \sigma_e^2(N-n)\left[\bar{x}_{\bar{s}}^{(2)} + \frac{(N-n)}{n}\bar{x}_{\bar{s}}^2\right] + \sigma_v^2\left[n + \frac{(N-n)}{n}\left\{(N-n)\bar{x}_{\bar{s}}^2\bar{x}_s^{(-2)} + 2n\bar{x}_{\bar{s}}\bar{x}_s^{(-1)}\right\}\right] \tag{3.5}$$

It is evident from the expression (3.4) that the bias is zero if $\bar{x}_s^{(-1)}\bar{x}_{\bar{s}}$ =1. That means a given sample s satisfies the overbalancing condition as described in section-1 for j=0and 1.

Therefore, if the drawn sample s is overbalanced, then $\hat{T}_1$ will be unbiased with variance given in (2.5) even if the true model is $\xi\left(1,1:x^2\right)$.This result is summarized in the following theorem.

Theorem 3.1: For overbalanced sample s satisfying the criteria given in the Section-1, if the estimator $\hat{T}_1$ developed in the model $\xi\left(0,1:x^2\right)$ is used in $\xi\left(1,1:\mathrm{x}^2\right)$, it remains unbiased with variance given in (2.5).

Under the overbalanced sample condition given in the section-1 for j=0,1, the variance of $\hat{T}$ without measurement error in y will remain same, denoted as

$$V_1\left[\hat{T}\right]=\sigma_e^2\,\frac{N(N-n)}{n}\left[f\bar{x}_{\bar{s}}^{(2)}+(1-f)\bar{x}_{\bar{s}}^2\right]\quad ,\text{f=n/N} \tag{3.6}$$

The relative efficiency of $\hat{T}_1$ with measurement error in y as compared to $\hat{T}$ without measurement is given by

$$E=\frac{V_1\left[\hat{T}\right]}{V\left[\hat{T}_1\right]}$$

$$=\frac{\sigma_e^2\,\frac{N(N-n)}{n}\left[f\bar{x}_{\bar{s}}^{(2)}+(1-f)\bar{x}_{\bar{s}}^2\right]}{\sigma_e^2\,\frac{N(N-n)}{n}\left[f\bar{x}_{\bar{s}}^{(2)}+(1-f)\bar{x}_{\bar{s}}^2\right]+\sigma_v^2\,\frac{N(N-n)}{n}\left[\frac{f(2-f)}{(1-f)}+(1-f)\bar{x}_{\bar{s}}^2\bar{x}_{\bar{s}}^{(-2)}\right]}$$

$$E=\frac{\left[f\bar{x}_{\bar{s}}^{(2)}+(1-f)\bar{x}_{\bar{s}}^2\right]}{\left[f\bar{x}_{\bar{s}}^{(2)}+(1-f)\bar{x}_{\bar{s}}^2\right]+\delta\left[\frac{f(2-f)}{(1-f)}+(1-f)\bar{x}_{\bar{s}}^2\bar{x}_{\bar{s}}^{(-2)}\right]} \tag{3.7}$$

Oviously,the value of E will be less than one indicating that there is loss in precision in the estimate due to measurement error in y. The % loss in precision (% LP) is worked out as

$$\%LP=(1-E)\times 100$$, i.e.

$$\%LP=\frac{\delta\left[\dfrac{f(2-f)}{(1-f)}+(1-f)\bar{x}_{\bar{s}}^{2}\bar{x}_{\bar{s}}^{(-2)}\right]}{\left[f\bar{x}_{\bar{s}}^{(2)}+(1-f)\bar{x}_{\bar{s}}^{2}\right]+\delta\left[\dfrac{f(2-f)}{(1-f)}+(1-f)\bar{x}_{\bar{s}}^{2}\bar{x}_{\bar{s}}^{(-2)}\right]}\times 100 \tag{3.8}$$

It is evident that % LP will depend upon the sampling fraction f and different sample and non-sample means of auxiliary variable x. A limited simulation study has been carried out to find out the extent of loss in precision.

4. Empirical study

Two simulation study have been conducted by generating hypothetical data through the model for examining the loss in precision of the estimator due to measurement error in y.

Simulation study-1

A limited simulation study has been conducted by generating hypothetical population of size N= 500 using the following model

$$y_i=\beta x_i+e_i x_i\,,\ \ v(y_i)=\sigma_e^2 x_i^2\,,\qquad i=1,2,\ldots,N \tag{4.1}$$

We have assumed that $\beta=1.50$, the error term follows normal distribution with mean zero and variance $\sigma^2=2$. It is assumed that x follows chi-square distribution with 5 degree of freedom. A population of Chi-Square of size N= 500 with 5 degree of freedom has been generated. Similarly, 500 values of e_i using normal distribution with mean 0 and variance 2 were generated. Using these values in model (4.1), a hypothetical population of N= 500 values of y were generated. Random sample of size n= 75 and n= 100 were drawn by simple random sampling without replacement (SRSWOR). This process of selection of sample for each size of 75 and 100 were repeated 20,000 times. That means 20,000 samples of each of the size 75 and 1000 were drawn from the population of size N= 500. R-software has been used for the simulation study. The variance of $\hat{T}$ given in (1.2), denoted as V_1 and variance of $\hat{T}_1$ given in (2.5), denoted as V_2, were computed for each sample of size n=75 and n=100 for $\delta=\dfrac{\sigma_v^2}{\sigma_e^2}$ = 0.75, 1.00 and 1.25. The average variance is computed as

$$V_1'=\frac{1}{s}\sum_{i=1}^{S}V_{1i}\ ,\ \ V_2'=\frac{1}{s}\sum_{i=1}^{S}V_{2i}\quad i=1,2,\ldots,s$$

where s= 20000 denotes the simulation run.

Percent loss in precision (% LP) was worked out as $\%LP=\left(1-\dfrac{V_1'}{V_2'}\right)\times 100$ and the simulation results are presented in the Table 4.1.

Variance	Sample size	δ		
		0.75	1.00	1.25
V_1'	n=75	29760.65	29760.65	29760.65
	n=100	27928.77	27928.77	27928.77
V_2'	n=75	42053.04	46150.51	50247.97
	% LP	29.23	35.51	40.77
	n=100	37921.46	41252.36	44583.26
	% LP	26.35	32.29	37.35

Table 4.1: The average estimate of variance for n=75, 100 and $\delta = \frac{\sigma_v^2}{\sigma_e^2}$ =0.75, 1.00 and 1.25

It is very obvious from the results of the Table (4.1) that there is considerable percent loss in precision varying from 26.35 to 40.75 percent depending upon the values of n and δ, when there is measurement error in y. It is also evident that the % loss in precision depends on δ. For lower value of δ, the % LP is smaller and it increases when δ increases. Therefore, if the variability in measurement error is relatively small as compared to model error variability σ_e^2, the % LP is expected to be smaller.

The results, of course, suggest that a caution has to be taken up by the investigator/ survey statisticians to collect/ measure reliable data from the sampled units from the population. A rigorous training needs to be provided to the field investigators in this regard in order to get precise estimate of population parameters in finite population survey sampling.

Simulation study-II

Using the same procedure of generation of hypothetical data under the model (4.1), we have generated the population of N= 500 values of y and x. 20,000 samples of each size n=75 and 100 were drawn from the population satisfying the criteria of overbalance sampling, i.e. $\bar{x}_s^{(-1)}\bar{x}_{\bar{s}} = 1$ by simple random sampling without replacement. R-software was used for the simulation. % LP was computed using the formula given in equation (3.8) for individual sample and for different values of $\delta = 0.75$, 1.00 and 1.25. The average value of % LP was obtained as

$$\% \text{ ALP} = \frac{1}{s}\sum_{i=1}^{S} \%\text{LP}$$

Here s=20,000. The value of % ALP are presented in the Table 4.2.

Sample size	δ		
	0.75	1.00	1.25
n=75	10.75	11.18	14.16
n=100	8.76	9.75	13.25

Table 4.2: Percent Average loss in precision (% ALP) for different values of n=75, and 100 and δ =0.75, 1.00 and 1.25

It is evident from the result of the Table 4.2 that % ALP increases with increase in the value of δ. Percent loss in precision decreased with increases in sample size. If sample selected is overbalanced, then by comparing the results of the Table 4.1 and 4.2 we find that % ALP (S) where of low order in later case, i.e. 10.75 to 14.16 % for n=75 and 8.76 to 13.25 % for n=100.

5. Concluding Remarks

A predictor of population total under the model $\xi(0,1 : x^2)$ has been developed with measurement error in y. Simulation results have shown substantial loss in precision that varied between 26 to 40 percent depending on sample size and $\delta = \sigma_v^2 / \sigma_e^2$. If the sample selected is overbalanced, then theoretical findings and results of another simulation study show that choosing a overbalanced sample enable us two-fold advantages: (i) it reduces the loss in precision considerably due to measurement error in y and (ii) it also protects the property of the predictor against the deviation of the model.

Acknowledgment

Authors are thankful to the referee for the valuable suggestions and corrections that has improved the earlier version of the manuscript of the paper. The authors are thankful to Mr. Sandeep Kumar, a researcher scholar, in Department of Agricultural Statistics, NDUAT, Kumarganj, Faizabad, for his help in simulation study.

References:

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

2. 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, Journal of the American Statistical Association, 83, p.28-36.
3. 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), Journal of the American Statistical Association, 65, p.1645-1667.
4. Bolfarine, H. (1991). Finite population prediction under error-in-variables super population models, The Canadian Journal of Statistics, 19 (2), p.191-207.
5. Chattopadhyay, M. 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, Series B, 3, p. 415-433.
6. Chauhan, Shweta and Sisodia, B.V.S. (2018). Robust estimation in finite population under error-in-variables super population model, International Journal of Agricultural and Statistical Sciences, 14(1), p. 35-41.
7. Chauhan, Shweta and Sisodia, B.V.S. and Singh, Dhirendra (2018). Roubest estimation in stratified sampling under error–in-variable super population model, Indian Society of Agriculture Statistics, 72 (1), p. 33-37.
8. Eltinge L. Jhon (1994). A Finite population approach to approximating small measurement error effects in regression, Sankhya: The Indian Journal of Statistics, 56, Series(B), pt 2; p.234-250.
9. Ghosh, M. and Sinha, K. (2007). Empirical bayes estimation in finite population sampling under functional measurement error models, Journal of Statistical Planning and Inference, 137, p. 2759 – 2773.
10. Jessen, R. (1942). Statistical investigation of a sample survey for obtaining farm facts, Iowa Agricultural Experiment Station Research Bulletin, 41, p.104.
11. Mukhopadhyay, P. (1994). Prediction in finite population under error-in-variables superpopulation models, Journal of Statistical Planning and Inference, 41, p.151-161.
12. Raj, Des (1958). On the relative accuracy of some sampling techniques, Journal of the American Statistical Association, 53, p.98-101.
13. 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, Journal of the American Statistical Association, 64, p.540-559.
14. Royall, R.M. (1970). On finite population sampling theory under certain linear regression models, Biometrika, 57(2), p.377−387.
15. Royall, R.M. (1971). Linear regression models in finite population sampling theory, in Godambe, V.P. and Sprott, D.A. eds. foundations of statistical inference Toronto Holt, Rinehart and Winston of Canada Ltd.
16. Royall, R. M. and Herson, J. (1973a). Robust estimation in finite populations, Journal of the American Statistical Association, 68, p.880-889.
17. Royall, R. M. and Herson, J. (1973b). Robust estimation in finite populations II. Stratification on a size variable, Journal of the American Statistical Association, 68, p. 890-893.
18. Scott, A. J., Brewer, K. R. W. and Ho, E. W. H. (1978). Finite population sampling and robust estimation, Journal of the American Statistical Association, 73, p.359-361.

19. Singh, A, Sisodia, B.V.S. and Mourya, K.K.(2018). Estimation of finite population total under error-in-variables super population models, Journal of Indian Society of Agricultural Statistics 72 (1), p.7-13.
20. Smith, H. Fairfield (1938). An experimental law desribing heterogeneity in the yields of agricultural crops, Journal of Agricultural science, 28, p.1-23.
21. Stefanski, L.A.(2000). Measurement error model, Journal of the American Statistical Association*,* 95, p.1353-1358.