Use of Non-response and Measurement error for Mean Estimation

QUAID-I-AZAM UNIVERSITY

ISLAMABAD

By

Aneesa Rehman

Department of Statistics
Faculty of Natural Sciences
Quaid-i-Azam University, Islamabad

2019

بسم الله الرحمن الرحيم

"IN THE NAME OF ALLAH, THE MOST GRACIOUS AND THE MOST MERCIFUL"

Use of Non-response and Measurement error for Mean Estimation

By

Aneesa Rehman

A THESIS SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF PHILOSOPHY IN STATISTICS

Supervised By

Prof. Dr. Javid Shabbir

Department of Statistics
Faculty of Natural Sciences
Quaid-i-Azam University, Islamabad
2019

CERTIFICATE

Use of Non-response and Measurement errors for Mean Estimation

By

Aneesa Rehman

(Reg. No. 02221711007)

A THESIS SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF M.PHIL. IN

STATISTICS

We accept this thesis as conforming to the required standards

1.____________________ 19/7/2019
Prof. Dr. Javid Shabbir
(Supervisor)

2.____________________
Dr. Masood Anwar
(External Examiner)

3.____________________ 22/07/19
Dr. Ijaz Hussain
(Chairman)

DEPARTMENT OF STATISTICS
QUAID-I-AZAM UNIVERSITY
ISLAMABAD, PAKISTAN
2019

Declaration

I 'Aneesa Rehman' hereby solemnly declare that this thesis entitled "Use of Non-response and Measurement error for Mean Estimation" submitted by me for the partial fulfillment of Master of Philosophy in Statistics is the original work and has not been submitted concomitantly or latterly to this or any other university for any other Degree.

Dated: 18-7-2019 Signature: ______

Dedication

I am feeling great honor and pleasure to dedicate this research work to

My Whole Family and Teachers

Whose endless affection, prayers and wishes have been a great source of comfort for me during my whole education period

Acknowledgments

I bestow my humblest praise for Almighty Allah, Who blessed me with courage, knowledge and resilience to complete this thesis. The completion of this thesis has been an expedition of personal growth that could not have happened without the guidance and support of many.

I would like to express my heartiest gratitude to my venerated supervisor Prof. Dr. javid Shabbir for his consistent support and supervision in every step of research. I am indebted to his guidance specifically throughout my thesis. I am very privileged and blessed to his guidance in my corner. I owe him many thanks.

I offer my deepest sense of gratitude, profound respect and tribute to all my teachers at the Department of Statistics, Quaid-i-Azam University.

My sincere thanks go to my whole family especially to my elder brothers Yasir Rehman, Nasir Rehman and my sister Fouzia Rehman for their love and support throughout my life. Without their support and encouragement it would have not been possible for me to complete this work. I would like to say thanks to all my friends and classmates for their cooperation and help especially Misbah Iftikhar, Waseem Abbas, Naila Altaf, Rida Zainab, Gazifa Azher, Sidra Shad, Amreen Qamar, Samar un Nisa Chaudhary. Finally, I am thankful to Mustansar Aatizaz Amjad and Shakeel Ahmad who have been very supportive and cooperative during my research work.

Aneesa Rehman

Abstract

We proposed an improved class estimators for the finite population mean utilizing information on two auxiliary variables, under two phase sampling technique by using simple random sampling without replacement (SRSWOR). The proposed class estimators is modified for handling the problem of measurement error and non-response. It is assumed that the problem of measurement error and nonresponse are present in second phase sample. The expressions for bias and MSE, are drived up-to- first order of approximation. Two real data set are used to investigate the performance of the proposed estimator. The proposed estimator is compared with usual mean estimator, traditional ratio estimator, exponential ratio type estimator, traditional regression estimator, Kirgyera(1984) type estimator, Sahoo (1930) regression type estimator, and Roy(2003) regression type estimator. It is found that proposed estimators are more efficient than all other the considered existing estimators by observing the MSE values.

Contents

List of Tables

Chapter 1

Introduction

A sampling method in which a portion of the population is selected in a way that the selected samples express the entire population. Everyone of us use it in day-to-day life. For example, to check whether the rice is cooked or not, one or two grains of rice are taken from the cooking pan. In order to decide by a quality controller whether the lot is according to the desired description or not few items are to be taken. To test for any change in blood of the entire body than normal, a pathologist takes few drops of blood. In all these circumstances, sampling is necessary and gives suitable results.To get data by sampling on the average yield of a crop, it is inadequate to include fields in the sample from the various parts of the country, the sample may perhaps consists of a fraction of fields of a specific class like growing improved variety, irrigated, manured, than is present in the population. The sample will stop to signify the entire population if any class is constantly preferred at the cost of the other. Even if a part of selected sample under various classes correspond with those in the population, the sample still may not imply the population. In sampling method, all features of the population must be reflected in the sample as closely as the sample size will allow, so that the consistent estimates of the population characteristics can be shaped from the sample.

1.1 Simple random sampling

Simple random sampling is an elementary approch of selection a sample from a population. In simple random sampling every individual has equal opportunity of being included in the sample. One method of attaining is to allot a number in the sampling frame to every individual and then use random numbers to choose an appropriate sample. One can generate random numbers by using calculator, a spreadsheet, random number table and more conventional techniques of drawing pieces of paper from a cap, flipping coins or rolling a die.

1.2 Auxiliary variable

The accuracy of an estimator for different population parameters can be increased by utilizing the auxiliary variable along with information of the study variable. Neyman (1938) used such information for the first time in estimating the population mean and got some improved results. This type of information can be used either at designing stage or at estimation stage or at both stages to enhance the precision of an estimator. At estimation stage it can be used in ratio, product, exponential-ratio, exponential-product and methods of regression estimation under different sampling schemes. Ratio and ratio type estimators are quite effective when the association is positive between the study variable and the auxiliary variable. On the other hand product and product type estimators are effective only if correlation between the the study variable and the auxiliary variable is negative. The auxiliary information can be in the form of quantitative and qualitative. Consider the following examples, when auxiliary information is available for estimating the finite population mean. (i) Let y be the I.Q of child and x be the age of child (ii) Let y be the amount of milk produced and x be the weight of the cow. (iii) Let y be the savings of a person and x be his income.

1.3 Double sampling

Ratio, regression and product estimators demand prior information on the auxiliary variable (x) to enhance the accuracy of an estimator. When such information are not available then taking a large primary sample of size n' is relatively cheaper for estimating population mean of the auxiliary variate to be used at estimation stage and a sample of size n ($n < n'$) is obtained for both the study (y) and the auxiliary (x) variate at 2nd phase. When a sampling is done in three or more phases, it is known as multi-phase sampling.

1.4 Non-response

In almost every survey, the sampling units which may be persons or households are not conducted. There are many reasons for non-contact. For example a person may be away from home or wrong addresses and telephone numbers. Even if sampling units are conducted, they may not respond to one or more questions in the survey. A portion of the units selected in the sample may not answer to the whole survey and some may give just few answers in a survey. These two cases are classified as unit non-response. If non-response is ignored and estimator is based on only information from respondents only then the estimator produces bias results and by increasing the sample size, mean square error (MSE) of the estimator decreases. Response rate depend upon very much on the subject matter of the survey. If topic of the survey is common interest of the people, then response rate will be high but if people were asked to provide their personal information like saving, income etc., then response rate generally will be low. There are many strategies to deal the problem of non-response like imputation of missing data, call backs and sub-sampling the non-respondents. In ratio method, missing values are imputed by the ratio of sum of the responding values of the study variable to the sum of the responding values of the auxiliary variable. In mean method, missing values are imputed by the mean of all responding values for that variable. Method of call back is due to Deming (1953). As indicated by Deming (1953) if the selected sampling unit is unavailable at home or absent to take part in the interview during call time, then it is recommended for the interviewer to call back. Hansen and Hurwitz (1964) suggested of taking a random sample from non-respondents with some extra effort and sources. Now a days this technique is commonly used to address the problem of non-response.

1.5 Measurement error

In many statistical analyses we assume that observations are noted accurately. This assumption is not satisfied in many situations and the observations are collected or recorded with measurement errors. Results become invalid in presence of measurement error. Measurement error is the difference between observed value and actual value of the variable, due to imperfection in the way that the data is collected. Phenomenon, therefore, measurement error are present in almost all the survey situations. Without taking into account the amount and effect of these errors, inferences drawn may be misleading. The data we obtained for statistical analysis is considered to be exact and free of measurement error. But practically it is far from reality. Some reasons of measurement errors are; inadequate sampling frame, incorrect identification of target population, non-response, incorrect questionnaire design, interviewer bias, respondent bias. We can avoide measurement error by correct identification of target population. Using an up to date sampling frame, revisiting to unavailable respondents, cautious questionnaire design, providing complete training for interviewers and processing staff.

Objectives of the study

Following are the objectives of our study that is;

1. To enhance precision of the estimators by taking advantages of the correlation between the study variable and auxiliary variables when measurement error and non-response exists on both study and auxiliary variables under simple random sampling scheme.

2. To obtain the properties of estimators up to first order of approximation under measurement error and nonresponse.

3.To make a comparisons of estimators by using the criteria of MSE

4. Two data set are used to support the theoretical findings.

1.6 Literature review

To enhance the proficiency of estimators, we use the auxiliary information. Neyman (1938) initially, used the auxiliary information for estimating the finite population mean. After his proposal, many authors have contributed in improved (in terms of bias and efficiency) estimation of mean. For example, Srivastava (1971) considered a class of ratio type estimators for estimating the population mean using multi-auxiliary information. Generally such type of information may not available prior to the conduction of survey but, in certain situations, it is easy as well as cheap to obtain information on the auxiliary variable from large sample in advance. Then the information on the study variable as well as from the auxiliary variable(s) are collected from the second sample which is smaller than the first phase sample. Using two-phase sampling in the presence of one auxiliary variable, Singh and Espejo (2007) , Vishwakarma and Singh (2012) proposed various estimators. Kiregyera (1984) constructed two regression type estimators with two auxiliary variables under two-phase sampling scheme for estimation of population mean. Sahoo et al. (1993) used two phase sampling mechanism and suggested a regression type estimator in the presence two auxiliary variables. They showed that the suggested estimator is found to be more accurate than usual two phase regression estimator and regression type estimators proposed by kiregyera (1984). Khare and Rehman (2013) proposed chain type estimators for population mean using auxiliary and additional auxiliary variables under double sampling scheme and then their properties are also studied. They concluded that the suggested estimators are also found to more accurate than the relevant chain type estimators for population mean proposed by Chand (1975) and Kiregyera (1980, 1984). In favour of the suggested estimators, an empirical study has been given. Pradhan (2005) suggested chain regression type estimators using three auxiliary variables under two-phase sampling scheme for estimating the population mean. Singh et al. (2008) proposed improved chain-ratio type estimator to estimate the population mean in double sampling scheme. They compared proposed estimator with two phase ratio type estimator and some other chain type estimators. With a numerical illustration, the performances of the proposed estimators have been made upto first order of approximation. In order to estimate the unknown population parameters and reduce bias due to non-response, a technique of sub-sampling of non-respondents was introduced by Hansen and Hurwitz (1946). It is assumed that on second call, all respondents will give full response. By using some extra sources and efforts, a random sample from non-respondents will be taken. In order to improve the efficiency of estimators, Hansen and Hurwitz (1946) have not used any kind of auxiliary information. Whereas Neyman (1938) for the first time and Cochran (1977) in case of non-response used the auxiliary information for estimating the population mean. Khare and Srivastava (1997), Rao (1986) further extend this work. When the information on auxiliary variable is not available. it is cheap to collect information on the auxiliary variable from large sample. Then the information on the study variable as well as on the auxiliary variable is collected, called double or two phase sampling. Technique introduced by Hansen and Hurwitz (1946) is further discussed by, Khare and Srivastava (1995) to deal with non-response in two-phase sampling for the estimating population mean. Okafor and Lee (2000), Singh and Karpe (2009) further extend this work. When population mean of the auxiliary variable is not known, Tabasum and Khan (2004) considered ratio estimator for population mean under double sampling in case of non-response and found optimal value for sample of both phases and sub-sampling fraction in such a way that it minimize the cost for specified precision. The problem of estimating the population mean $\bar{Y}$ of the study variate y with two auxiliary variates x and z in the presence of non-response is discussed by Singh and Kumar (2010). A comparison between suggested class of estimator and usual unbiased estimator reported by Hansen and Hurwitz (1946), Khare and Srivastava (1995) , Tabasum and Khan (2004, 2006) and Okafor and Lee (2000) were made and it is found that under certain conditions, proposed estimators perform better than other considered estimators. Data is collected on more than one auxiliary variable on a large scale survey. When information is available on more than one auxiliary variable, a general class of estimators is suggested which uses known value of the auxiliary variable. Olkin (1958), Rao and Mudholkar (1967), proposed a general class of estimators using the multi-auxiliary information for population mean. But it is possible when population mean of the auxiliary variable is unknown. So, we hereby use double sampling scheme to enhance the estimate of population mean of auxiliary variable, suggested by Srivastava (1981). But in case of non-response from sampling units, bias may occur and then non-respondents are again contact to avoid the bias. In case of non-response, Khare and Sinha (2007) suggested a general class of estimators utilizing the multi-auxiliary information. Singh and Kumar (2010) improved went a step further by proposing class of estimators under different situation i.e double-sampling scheme. They identified asymptotically optimum estimator for each class and also obtained empirical study in order to have an efficient estimator.

Several authors have worked on the problem of measurement error. Singh et al. (2014)

proposed difference-type estimator in the presence of measurement error using the auxiliary information for estimation of population mean. The proposed estimator has been recognized along with its mean square error formula. They showed that the proposed estimator performs better than other existing estimators. Singh and Karpe (2009) proposed a class of estimators of the population variance in the presence of measurement error using the auxiliary information. They obtained the bias and mean squared errors of the recommended class of estimators up to the terms of order $O(n^{-1})$. Kumar (2011) considered the problem of estimating population mean in the presence of measurement error using auxiliary information. A comparative study is done among proposed estimators, the ratio estimator and mean per unit estimator in presence of measurement error. Sharma and Singh (2013) generalized the auxiliary information by considering the problem of estimating population variance in the presence of measurement error. They supposed that measurement error is presented both on the study and the auxiliary variable. In the presence of measurement errors, a numerical study is made to compare the performances of the proposed estimator with variance per unit estimator and other estimators. Singh and Singh (2002) inspected the effect of measurement errors on usual linear regression estimator. In the presence of measurement errors, a comparative study is carried out among the mean per unit estimator, ratio estimator and linear regression estimator.

Chapter 2

Estimation of Population Mean Under Two Phase Sampling

Outline

In this Chapter we have proposed a new estimator for estimating the population mean using two auxiliary variables under double sampling scheme by applying Searls (1964) technique. Bias and MSE expression of proposed estimator are derived up to first oder of approximation and proposed estimator is compared with existing estimators to observe the efficiency. Some real datasets are used to observe the performances of the estimators.

2.1 Introduction

Sometimes population mean of the auxiliary variable is not available then we adopt the procedure of two phase sampling. We select a large sample of size n' in first phase by adopting simple random sampling without replacement scheme(SRSWOR). In the second phase, we select the sub-sample of size n, is drawn from the first sample $(n')$and observed the study variable y as well as the auxiliary variables (x, z). When population mean $\bar{X}$ is not unknown but information on second variable z is available whose correlation with study variable (y) is less, then we adopt two phase sampling. Assume finite population $S = \{S_1, S_2, \ldots, S_N\}$ of size N and let the observations on the i^{th} unit of the study variable y and the auxiliary variables (x, z) are y_i, x_i and z_i respectively. By using simple random sampling without replacement scheme, a large sample of size n' is selected from the population and variables x and z are measured. Than a second phase sample of size $n(n < n')$ is selected from the first phase sample by using simple random sampling without replacement technique. On this sample, variables y, x and z are measured. Let $\bar{x}'$ and $\bar{z}'$ be the first phase sample means of variables x and z respectively. Let $\bar{y}$, $\bar{x}$ and $\bar{z}$ be the second phase sample means of variables y, x and z respectively. Let $\bar{Y}$, $\bar{X}$ and $\bar{Z}$ be the population means of variables y, x and z respectively.

2.2 Notations and symbols

We define the following notations and symbols.

$$\textit{Let} \quad e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}},\ e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}},\ e_1' = \frac{\bar{x}' - \bar{X}}{\bar{X}},\ e_2 = \frac{\bar{z} - \bar{Z}}{\bar{Z}},\ e_2' = \frac{\bar{z}' - \bar{Z}}{\bar{Z}}$$

$$e_3 = \frac{s_{yx} - S_{yx}}{S_{yx}},\ e_4 = \frac{s_x^2 - S_x^2}{S_x^2},\ e_5 = \frac{s_{yz} - S_{yz}}{S_{yz}},\ e_6 = \frac{s_z - S_Z^2}{S_z}$$

$$E(e_i),\ i = (0, 1, \ldots, 6),\ E(e_j') = 0,\ j = (1, 2),\ E(e_0^2) = \theta_1 C_y^2,\ E(e_1^2) = \theta_1 C_x^2$$

$$E(e_2^2) = \theta_1 C_z^2,\ E(e_1'^2) = E(e_1 e_1') = \theta_2 C_x^2,\ E(e_0 e_1') = \theta_2 \rho_{yx} C_y C_x,\ E(e_0 e_2') = \theta_2 \rho_{yz} C_y C_z$$

$$E(e_2 e_1') = E(e_2' e_1') = E(e_1 e_2') = \theta_2 \rho_{xz} C_x C_z,\ E(e_0 e_1) = \theta_1 \rho_{yx} C_y C_x^2,\ E(e_0 e_2) = \theta_1 \rho_{yz} C_y C_z$$

$$E(e_1 e_3) = \theta_1 \lambda_{210} C_x,\ E(e_1' e_3) = \theta_2 \lambda_{210} C_x,\ E(e_1 e_4) = \theta_1 \lambda_{300} C_x,\ E(e_1' e_4) = \theta_2 \lambda_{300} C_x$$

$$E(e_2' e_3) = \theta_2 \lambda_{111} C_z,\ E(e_2' e_4) = \theta_2 \lambda_{201} C_z,\ E(e_2' e_5) = \theta_2 \lambda_{012},\ C_z,\ E(e_2' e_6) = \theta_2 \lambda_{003} C_z$$

$$E(e_1 e_2) = \theta_1 \rho_{xz} C_x C_z,$$

$$\theta_1 = (\frac{1}{n} - \frac{1}{N}), \theta_2 = (\frac{1}{n'} - \frac{1}{N}),\ \theta_3 = (\frac{1}{n} - \frac{1}{n'})$$

$$\mu_{rst} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^r (y_i - \bar{Y})^s (z_i - \bar{Z})^t,\ \lambda_{rst} = \frac{\mu_{rst}}{\mu_{200}^{\frac{r}{2}} \mu_{020}^{\frac{s}{2}} \mu_{002}^{\frac{t}{2}}}$$

$$\bar{Y}=\frac{1}{N}\sum_{i=1}^{N}y_i,\ \bar{X}=\frac{1}{N}\sum_{i=1}^{N}x_i,\ \bar{Z}=\frac{1}{N}\sum_{i=1}^{N}z_i$$

$$S_y^2=\sum_{i=1}^{N}\frac{(y_i-\bar{Y})^2}{N-1},\ S_x^2=\sum_{i=1}^{N}\frac{(x_i-\bar{X})^2}{N-1},\ S_z^2=\sum_{i=1}^{N}\frac{(z_i-\bar{Z})^2}{N-1},\ \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i,$$

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \bar{z}=\frac{1}{n}\sum_{i=1}^{n}z_i,\ \bar{x}'=\frac{1}{n'}\sum_{i=1}^{n'}x_i',\ \bar{z}'=\frac{1}{n'}\sum_{i=1}^{n'}z_i',\ s_y^2=\sum_{i=1}^{n}\frac{(y_i-\bar{y})^2}{n-1},$$

$$s_x^2=\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{n-1},\ s_z^2=\sum_{i=1}^{n}\frac{(z_i-\bar{z})^2}{n-1},\ s_x'^2=\sum_{i=1}^{n'}\frac{(x_i-\bar{x}')^2}{n'-1},s_z'^2=\sum_{i=1}^{n'}\frac{(z_i-\bar{z}')^2}{n'-1}$$

$$C_y=\frac{S_y}{\bar{Y}},\ C_x=\frac{S_x}{\bar{X}},\ C_z=\frac{S_z}{\bar{Z}},\ \rho_{yx}C_yC_x=C_{yx},\ \rho_{yz}C_yC_z=C_{yz},\ \rho_{xz}C_xC_z$$

2.3 Existing estimators

In this section we discuss the Mean Square Error (MSEs)of existing estimators under double sampling:

(i) **Usual mean estimator**

$$\hat{\bar{Y}}_0=\bar{y}, \tag{2.1}$$

The variance of $\hat{\bar{Y}}_0$, is given by:

$$Var(\hat{\bar{Y}}_0)=\theta_1 S_y^2. \tag{2.2}$$

(ii) **Classical ratio estimator**

Classical ratio estimator under double sampling, is:

$$\hat{\bar{Y}}_R=\bar{y}\left(\frac{\bar{x}'}{\bar{x}}\right). \tag{2.3}$$

The bias and MSE of $\hat{\bar{Y}}_R$, upto first order of approximation are:

$$Bias(\hat{\bar{Y}}_R) \cong \bar{Y}\theta_3 C_x^2\Big(1 - \frac{\rho_{yx}C_y}{C_x}\Big) \tag{2.4}$$

$$MSE(\hat{\bar{Y}}_R) \cong \bar{Y}^2\Big\{\theta_3(C_x^2 - 2\rho_{yx}C_xC_y) + \theta_1 C_y^2\Big\}. \tag{2.5}$$

(iii) **Bahl and Tuteja (1991) exponential ratio estimator**

Bahl and Tuteja (1991) exponential ratio estimator under double sampling, is:

$$\hat{\bar{Y}}_{BT} = \bar{y}exp\Big(\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\Big). \tag{2.6}$$

The bias and MSE of $\hat{\bar{Y}}_{BT}$, up to first oder of approximation are

$$Bias(\hat{\bar{Y}}_{BT}) \cong \bar{Y}C_x^2\theta_3(\frac{3}{8} - \frac{\rho_{yx}C_y}{2C_x}) \tag{2.7}$$

$$MSE(\hat{\bar{Y}}_{BT}) \cong \bar{Y}^2\Big\{\theta_1 C_y^2 + \theta_3 C_x^2(\frac{1}{4} - \frac{\rho_{yx}C_y}{C_x})\Big\} \tag{2.8}$$

(iv) **Traditional regression estimator**

Traditional regression estimator is given below

$$\hat{\bar{Y}}_{Reg} = \bar{y} + b_{yx}\Big(\bar{x}' - \bar{x}\Big), \tag{2.9}$$

where b_{yx} is the sample regression coefficient of y on x.

The bias and MSE of $\hat{\bar{Y}}_{Reg}$, up to first oder of approximation are

$$Bias(\hat{\bar{Y}}_{Reg}) \cong -\bar{X}B_{yx}C_x\Big(\lambda_{210} - \lambda_{300}\Big)\theta_3 \tag{2.10}$$

$$MSE(\hat{\bar{Y}}_{Reg}) \cong \bar{Y}^2C_y^2\Big\{\theta_1(1 - \rho_{yx}^2) + \theta_2\rho_{yx}^2\Big\} \tag{2.11}$$

(v) **Kiregyera (1984) ratio-regression type estimator**

Kiregyera (1984) ratio-regression type estimator, is

$$\hat{\hat{Y}}_K = \bar{y} + b_{yx}\left\{\bar{Z}\left(\frac{\bar{x}'}{\bar{z}'}\right) - \bar{x}\right\}, \tag{2.12}$$

The bias and MSE of $\hat{\hat{Y}}_K$, by using first oder of approximation are

$$Bias(\hat{\hat{Y}}_K) \cong -\Big[\Big\{(-C_z\rho_{xz}-\lambda_{210}+\lambda_{300})C_x-C_z(C_z-\lambda_{111}+\lambda_{201})\Big\}\theta_2+\theta_1C_x(\lambda_{210}-\lambda_{300})\Big]\bar{X}B_{yx} \tag{2.13}$$

$$MSE(\hat{\hat{Y}}_K) \cong \bar{Y}^2\left[\theta_1C_y^2 - \theta_3C_y^2\rho_{yx}^2 + \theta_2\left\{\left(\frac{C_yC_z\rho_{yx}}{C_x} - C_y\rho_{yz}\right)^2 - C_y^2\rho_{yz}^2\right\}\right]. \tag{2.14}$$

(vi) **Sahoo et al. (1993) regression-type estimator**

Sahoo et al. (1993) regression-type estimator, is:

$$\hat{\hat{Y}}_S = \bar{y} + b_{yx}(\bar{x}' - \bar{x}) + b_{yz}(\bar{Z} - \bar{z}'), \tag{2.15}$$

The bias and MSE of $\hat{\hat{Y}}_S$ is given below by using first order of approximation

$$Bias(\hat{\hat{Y}}_S) \cong C_xB_{yx}\theta_3\Big(-\lambda_{210} + \lambda_{300}\Big)\bar{X} - \bar{Z}\theta_2C_zB_{yz}\Big(\lambda_{012} - \lambda_{003}\Big) \tag{2.16}$$

$$MSE(\hat{\hat{Y}}_S) \cong \bar{Y}^2C_y^2\left\{\theta_1(1 - \rho_{yx}^2) + \theta_2(\rho_{yx}^2 - \rho_{yz}^2)\right\}. \tag{2.17}$$

(vii) **Roy (2003) regression type estimator**

Roy(2003) regression type estimator of $\hat{\hat{Y}}_{Reg}$, is

$$\hat{\hat{Y}}_{Roy} = \bar{y} + k_1\left[\bar{x}' + k_2(\bar{Z} - \bar{z}') - \left\{\bar{x} + k_3(\bar{Z} - \bar{z})\right\}\right]. \tag{2.18}$$

The bias and MSE of $\hat{\bar{Y}}_{Roy}$, by using first order of approximation are

$$Bias(\hat{\bar{Y}}_{Roy}) \cong \bar{Y}^2 \tag{2.19}$$

$$MSE(\hat{\bar{Y}}_{Roy}) = \bar{Y}^2 C_y^2 \left[\theta_1 (1 - \rho_{y.xz}^2) + \theta_2 (1 - \rho_{yz}^2) \rho_{yx.z}^2 \right], \tag{2.20}$$

where

$$\rho_{y.xz}^2 = \frac{\rho_{yx}^2 + \rho_{yz}^2 - 2\rho_{yz}\rho_{yx}\rho_{xz}}{1 - \rho_{xz}^2} \; and \; \rho_{yx.z}^2 = \frac{(\rho_{yx} - \rho_{yz}\rho_{xz})^2}{(1 - \rho_{yz}^2)(1 - \rho_{xz}^2)}$$

are multiple and partial correlation coefficients.

(viii) **Grover (2019) regression type estimator**

Grover (2019) regression type estimator of $\hat{\bar{Y}}_G$, is

$$\hat{\bar{Y}}_G = k_0 \bar{y} + k_1 \left[\bar{x}' + k_2 (\bar{Z} - \bar{z}') - \left\{ \bar{x} + k_3 (\bar{Z} - \bar{z}) \right\} \right], \tag{2.21}$$

where k_0, k_1, k_2 are constants.

The bias and MSE of $\hat{\bar{Y}}_G$ are given below by using first order of approximation:

$$Bias(\hat{\bar{Y}}_G) \cong \bar{Y}(K_0 - 1) \tag{2.22}$$

$$MSE(\hat{\bar{Y}}_G) \cong \bar{Y}^2 \frac{(A_1 + B_1 + C_1)}{(S_1 + A_1 + B_1 + C_1)}, \tag{2.23}$$

where

$A_1 = \{(-S_x^2 S_y^2 + S_{xy}^2)\theta_1^2 + \theta_2 (S_x^2 S_y^2 - 2S_{xy}S_{yx})\theta_1 + \theta_2^2 S_{yx}^2\} S_z^4,$

$B_1 = \{(S_x^2 S_{yz}^2 + S_{xz}^2 S_y^2 - 2S_{yx}S_{xz}S_{yz})\theta_1 + 2S_{yz}\theta_2 S_{xz} S_{yx}\}\theta_3 S_z^2,$

$C_1 = -S_{yz}^2 \theta_2 S_{xz}^2 \theta_3, \; S_1 = -(S_x^2 S_z^2 - S_{xz}^2)\theta_3 \bar{Y}^2 S_z^2.$

2.4 Proposed estimator

By using Searls (1964) technique, we have proposed the following estimator for population mean $\bar{Y}$.

$$\hat{\bar{Y}}_p = k_0\bar{y} + k_1\left[\bar{x}' + k_2 exp\left(\frac{\bar{Z}-\bar{z}'}{\bar{Z}+\bar{z}'}\right) - \left\{\bar{x} + k_3 exp\left(\frac{\bar{Z}-\bar{z}}{\bar{Z}+\bar{z}}\right)\right\}\right], \tag{2.24}$$

where $k_i (i = 0, 1, 2, 3)$ are constants.

In terms of error, we have

$$\hat{\bar{Y}}_p = k_0\bar{Y} + k_0\bar{Y}e_0 + k_1\bar{X}e_1' + k_1k_2 - \frac{k_1k_2e_2'}{2} + \frac{3k_1k_2e_2^{2'}}{8} - k_1\bar{X}e_1 - k_1k_3 + \frac{k_1k_3e_2}{2} - \frac{3k_3k_1e_2^2}{8}. \tag{2.25}$$

Bias and Mean square error (MSE) of $\hat{\bar{Y}}_p$, up to first order of approximation, are:

$$Bias(\hat{\bar{Y}}_p) \cong \bar{Y}k_0 + \frac{3}{8}\frac{k_1k_2\theta_2S_z^2}{\bar{Z}^2} - \frac{3}{8}\frac{k_1k_3\theta_1S_z^2}{\bar{Z}^2} + k_1k_2 - k_1k_3 - \bar{Y}, \tag{2.26}$$

$$\begin{aligned} MSE(\hat{\bar{Y}}_p) &= (k_0-1)^2\bar{Y}^2 + \Big\{(C_{yz}k_0\theta_1 - 2k_0 + 2)k_3 - (2 + (C_{yz}\theta_2 - 2)k_0)k_2\Big\}k_1\bar{Y} \\ &\quad - C_{xz}k_1^2k_3\theta_3\bar{X} + \frac{1}{4}\Big\{(C_z^2\theta_1 + 4)k_3^2 + (-2C_z^2\theta_2 - 8)k_2k_3 + (C_z^2\theta_2 + 4)k_2^2 \\ &\quad + 4S_x^2\theta_3\Big\}k_1^2 - 2S_{yx}k_0\theta_3 + k_0^2\theta_1S_y^2, \end{aligned} \tag{2.27}$$

Optimum values of $k_0,\ k_1,\ k_2$ and k_3, are given as:

$$k_{0(opt)} = \frac{-\theta_2\bar{Y}^2(\bar{X}C_{xz} - C_zS_x)(\bar{X}C_{xz} + C_zS_x)(C_z^2 + 2C_{yz})}{\begin{aligned}&\bar{Y}^2\Big\{\theta_2C_{xz}^2 l\bar{X}^2 + S_x^2((C_z^4 + (-C_{yz}^2\theta_1 + 4C_{yz})C_z^2 + 4C_{yz}^2)\theta_2 - 4C_{yz}^2\theta_1)\Big\} \\ &+ 2\bar{X}S_{yx}C_{xz}C_{yz}p\theta_3\bar{Y} + p\Big\{-\bar{X}^2S_y^2C_{xz}^2\theta_1 + C_z^2(S_{yx}^2\theta_2 + \theta_1(S_x^2S_y^2 - S_{yx}^2))\Big\}\end{aligned}}, \tag{2.28}$$

$$k_{1(opt)} = \frac{-(C_z^2+2C_{yz})(\bar{X}\bar{Y}C_{xz}C_{yz}-C_z^2S_{yx})\bar{Y}^2\theta_2}{\bar{Y}^2\Big\{\theta_2 C_{xz}^2 l\bar{X}^2 + S_x^2((C_z^4+(-C_{yz}^2\theta_1+4C_{yz})C_z^2+4C_{yz}^2)\theta_2 - 4C_{yz}^2\theta_1)\Big\} + 2\bar{X}S_{yx}C_{xz}C_{yz}p\theta_3\bar{Y} + p\Big\{-\bar{X}^2S_y^2C_{xz}^2\theta_1 + C_z^2(S_{yx}^2\theta_2+\theta_1(S_x^2S_y^2-S_{yx}^2))\Big\}}, \quad (2.29)$$

$$k_{2(opt)} = \frac{(2\theta_2\bar{X}^2C_{xz}^2C_{yz}+4S_x^2C_{yz}^2\theta_1)\bar{Y}^2 - 2C_{xz}\ \{(-2\theta_2+4\theta_1)C_{yz} + \theta_2C_z^2\ \}\bar{X}S_{yx}\bar{Y} + 4\bar{X}^2S_y^2C_{xz}^2\theta_1 - 4qC_z^2}{\bar{Y}\theta_2(C_z^2+2C_{yz})(\bar{X}\bar{Y}C_{xz}C_{yz}-C_z^2S_{yx})}, \quad (2.30)$$

$$k_{3(opt)} = \frac{2(-\bar{Y}S_x^2C_{yz}+C_{xz}S_{yx}\bar{X})}{(\bar{X}\bar{Y}C_{xz}C_{yz}-C_z^2S_{yx})}, \quad (2.31)$$

where,

$q = \theta_1 S_x^2 S_y^2 - \theta_3 S_{yx}^2,\ p = C_z^2\theta_2 + 4,\ l = C_{yz}^2\theta_2 - C_z^2 - 4C_{yz},$

The minimum MSE of $\hat{\bar{Y}}_p$, is given by:

$$MSE_{min}(\hat{\bar{Y}}_p) = \frac{\bar{Y}^2A^*}{B^*}, \quad (2.32)$$

where

$$A^* = \theta_2[C_{yz}^2(\theta_2\bar{X}^2C_{xz}^2 - C_z^2S_x^2\theta_1)\bar{Y}^2 + 2\bar{X}C_z^2S_{yx}C_{xz}C_{yz}\theta_3\bar{Y} + C_z^2(-\bar{X}^2S_y^2C_{xz}^2\theta_1 + C_z^2q)]$$

and

$$B^* = \bar{Y}^2\{\theta_2C_{xz}^2l\bar{X}^2 + S_x^2(\theta_2C_z^4 - \theta_2C_{yz}(C_{yz}\theta_1-4)C_z^2 - 4C_{yz}^2\theta_3)\} + 2\bar{X}S_{yx}C_{xz}C_{yz}p\theta_3\bar{Y} + p(-\bar{X}^2S_y^2C_{xz}^2\theta_1 + C_z^2q)$$

2.5 Efficiency comparisons

Comparing (2.2) and (2.32), $MSE(\hat{\bar{Y}}_0) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$\theta_1 S_y^2 > \frac{A^*}{B^*}.$$

Comparing (2.5) and (2.32), $MSE(\hat{\bar{Y}}_R) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$\left\{\theta_3(C_x^2 - 2\rho_{yx}C_xC_y) + \theta_1 C_y^2\right\} - \frac{A^*}{B^*} > 0$$

Comparing (2.8) and (2.32), $MSE(\hat{\bar{Y}}_{BT}) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$\left\{\theta_3(C_x^2 - \rho_{yx}C_xC_y) + \theta_1 C_y^2\right\} - \frac{A^*}{B^*} > 0$$

Comparing (2.11) and (2.32), $MSE(\hat{\bar{Y}}_{Reg}) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$C_y^2\left\{\theta_1(1-\rho_{yx}^2) + \theta_2\rho_{yx}^2\right\} - \frac{A^*}{B^*} > 0$$

Comparing (2.14) and (2.32), $MSE(\hat{\bar{Y}}_K) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$\left[\theta_1 C_y^2 - \theta_3 C_Y^2\rho_{yx}^2 + \theta_2\left(\left(C_y\rho_{yx}\left(\frac{C_z}{C_x}\right) - C_y p_{yz}\right)^2 - Cy^2\rho_{yz}^2\right)\right] - \frac{A^*}{B^*} > 0$$

Comparing (2.17) and (2.32), $MSE(\hat{\bar{Y}}_S) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$C_y^2\left\{\theta_1(1-\rho_{yx}^2) + \theta_2(\rho_{yx}^2 - \rho_{yz}^2)\right\} - \frac{A^*}{B^*} > 0$$

Comparing (2.20) and (2.32), $MSE(\hat{\bar{Y}}_{Roy}) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$C_y^2\left[\theta_1(1-\rho_{y.xz}^2)+\theta_2(1-\rho_{yz}^2)\rho_{yx.z}^2\right]-\frac{A^*}{B^*}>0$$

Comparing (2.23) and (2.32), $MSE(\hat{\bar{Y}}_G) - MSE_{min}(\hat{\bar{Y}}_p) > 0$ if

$$\frac{(A_1+B_1+C_1)}{(S_1+A_1+B_1+C_1)}-\frac{A^*}{B^*}>0$$

2.6 Numerical illustration

Population I [Source: Murthy et al. (1967)]

y:Output

x:Number of workers

z:Fixed capital

$$N=80, \bar{Y}=5182.637,\ \bar{X}=285.125,\ \bar{Z}=1126.463,\ S_y=1835.659,$$

$$S_x=270.4294,\ S_z=845.6097,\ \rho_{yx}=0.9149811,\ \rho_{yz}=0.9413055,$$

$$\rho_{xz}=0.9884207,\ S_{yx}=454211.4,\ S_{yz}=1461142,\ S_{xz}=226029.8$$

Population II [Source: Koyuncu and Kadilar (2009)]

y: Number of teachers

x: Number of students both primary and secondary

z: Number of classes both primary and secondary school

$$N = 127,\ \bar{Y} = 703.7402,\ \bar{X} = 20804.59,\ \bar{Z} = 498.2756,\ S_y = 883.8348,$$

$$S_x = 30486.75, S_z = 555.5816,\ \rho_{yx} = 0.9366086,\ \rho_{yz} = 0.978914,$$

$$\rho_{xz} = 0.9395892, S_{yx} = 25237154,\ S_{yz} = 480688.2,\ S_{xz} = 15914648$$

Table 2.1: MSE's and PREs of different estimators with respect to $\hat{\bar{Y}}_0$ for Population I

	n'=24 n=7	n'=28 n=9	n'=32 n=10	n'=24 n=7	n'=28 n=9	n'=32 n=10
		MSE			*PRE*	
$\hat{\bar{Y}}_0$	439256.9	332284.2	294843.7	100.0000	100.0000	100.0000
$\hat{\bar{Y}}_R$	1213384	909084.6	820794.6	36.2009	36.5515	35.9217
$\hat{\bar{Y}}_{BT}$	215070.6	165243.4	142528.9	204.2384	201.0877	206.8659
$\hat{\bar{Y}}_{Reg}$	153795.3	119587.3	100897.7	285.6114	277.8591	292.2204
$\hat{\bar{Y}}_K$	71345.94	53964.3	47894.54	615.6719	615.7481	615.6102
$\hat{\bar{Y}}_S$	66712.63	50276.56	44915.99	658.4314	660.9126	656.4337
$\hat{\bar{Y}}_{Roy}$	46527.21	35236.45	31201.77	944.0861	943.0127	944.9581
$\hat{\bar{Y}}_G$	46446.75	35190.28	31165.57	945.7215	944.2498	946.0558
$\hat{\bar{Y}}_p$	11664.19	8880.074	7787.738	3765.86	3741.908	3785.999

In table 2.1 the MSE of the proposed estimator is smaller and the efficiency of the proposed estimator is maximum than all other considered existing estimators for different combination of n' and n.

Table 2.2: MSE's and PRE's of different estimators with respect to $\hat{\bar{Y}}_0$ for Population II

	n'=38 n=12	n'=51 n=16	n'=64 n=19	n'=38 n=12	n'=51 n=16	n'=64 n=19
		MSE			PRE	
$\hat{\bar{Y}}_0$	58946.09	42671.85	34962.99	100.0000	100.0000	100.0000
$\hat{\bar{Y}}_R$	22233.89	15054.64	11135.36	265.1182	283.4465	313.9817
$\hat{\bar{Y}}_{BT}$	25430.76	17459.52	13210.25	231.7905	244.4044	264.6656
$\hat{\bar{Y}}_{Reg}$	19873.98	13279.36	9603.685	296.5994	321.3396	364.0581
$\hat{\bar{Y}}_K$	7090.293	5145.569	4230.768	831.3633	829.2931	826.3983
$\hat{\bar{Y}}_S$	6069.057	4495.793	3801.547	971.2562	949.1505	919.7043
$\hat{\bar{Y}}_{Roy}$	2351.981	1699.576	1389.022	2506.232	2510.734	2517.095
$\hat{\bar{Y}}_G$	2340.864	1693.764	1385.137	2518.134	2519.35	2524.154
$\hat{\bar{Y}}_p$	211.7601	151.5051	121.4121	27836.27	28165.28	28796.95

In table 2.2 the MSE of the proposed estimator is smaller and the efficiency of the proposed estimator is maximum than all other considered existing estimators for different combination of n' and n.

Table 2.3: Conditional values for Population I and Population II

Conditions	Population1	Population2
i	0.01635373	0.1190228
ii	0.04517484	0.04489426
iii	0.008007176	0.05134932
iv	0.005725869	0.04012917
v	0.002656242	0.01431659
vi	0.002483742	0.01225453
vii	0.001732229	0.004749076
viii	0.001729234	0.004726629
ix	0.000434263	0.000427582

All the values in Table 3.5 are greater than zero (positive) which shows that proposed estimator perform better than existing estimators.

Conclusion

From Tables 2.1-2.2, we observed that the efficiency of the suggested estimator is more as compared to the usual mean estimator, ratio estimator, traditional regression estimator, Kiregyera (1984) ratio in regression estimator, Sahoo (1993) regression type estimator, Roy(2003) regression type estimator and Grover(2019)s regression type estimator. Further, we can observe that MSEs of the proposed estimator is smaller and PREs of the suggested estimator is larger as compare to the considered estimators for different combinations of n', and n. In Table 2.3, all the conditional values are greater than zero which shows the excellency of proposed estimator.

Chapter 3

Estimation of Population Mean in the Presence of Measurement Error

3.1 Introduction

In survey sampling, generally we assume that the data we obtained for estimating different characteristics to be free of measurement errors. Such supposition is not satisfied in many situations and the measurement error do occurs in observed values due to various reasons. Measurement error is defined as the difference between observed value and actual value of the variable due to the wrong way in which the data is collected. There are many reasons of occurrence of measurement errors. Some of these reasons are: inadequate sampling frame, incorrect identification of target population, incorrect questionnaire design respondent bias and interviewer bias etc. We can avoid measurement error by using an up-to-date sampling frame, by designing carefully questionnaire, by providing complete training for processing staff and interviewers, by using correct identification of target population. Let a population $U = \{S_1, S_2, \ldots, S_N\}$ of size N. Let Y, X and Z be study and two auxiliary variables respectively. By using these two auxiliary variables, we have to estimate $\bar{Y}$. It is considered that the population mean $\bar{X}$ is not known but the population mean $\bar{Z}$ is known. As information on the auxiliary variable is not available so we estimate it by using two phase sampling. We assume that there is no measurement error at the first phase but on at second phase there is measurement error. Let (x', z') and (X_i, Z_i) be the observed and actual values of the auxiliary characteristics associated with the i^{th} $\{i = 1, 2, \dots, n'\}$ unit in the first phase sample. We have assume that there is no measurement error in 1st phase. Let (y_i^*, x_i^*, z_i^*) and (Y_i, X_i, Z_i) be the observed and true values on characteristics (Y, X, Z) respectively associated with the i^{th} $\{i = 1, 2, \dots, n\}$ unit of the second-phase sample. The measurement error is given by

$u_i = y_i - Y_i,$

$v_i = x_i - X_i,$

$w_i = z_i - Z_i,$

where (u_i, v_i, w_i) are random in nature and are uncorrelated with mean zero and variance S_u^2, S_v^2 and S_w^2 respectively.

3.2 Notations

$$E(e_i^*),\ \ i = (0, 1, 2), \qquad E(e_j') = 0,\ j = (1, 2)$$

$$E(e_0^{*2}) = \frac{\theta_1(S_u^2 + S_y^2)}{\bar{Y}^2} = C_{yu}^2\theta_1,\ E(e_1^{*2}) = \frac{\theta_1(S_v^2 + S_x^2)}{\bar{X}^2} = C_{xv}^2\theta_1,$$

$$E(e_2^{*2}) = \frac{\theta_1(S_w^2 + S_z^2)}{\bar{Z}^2} = C_{zw}^2\theta_1,\ E(e_1'^2) = E(e_1e_1') = \frac{\theta_2 S_x^2}{\bar{X}^2} = \theta_2 C_x^2$$

$$E(e_2'^2) = E(e_2e_2') = \frac{\theta_2 S_z^2}{\bar{Z}^2} = \theta_2 C_z^2,\ E(e_0^*e_1^*) = \frac{\theta_1 S_{yx}}{\bar{Y}\bar{X}} = \theta_1 C_{yx}$$

$$E(e_0^*e_2^*) = \frac{\theta_1 S_{yz}}{\bar{Y}\bar{Z}} = \theta_1 C_{yz},\ E(e_1^*e_2^*) = \frac{\theta_1 S_{xz}}{\bar{Z}\bar{X}} = \theta_1 C_{xz},$$

$$E(e_0e_1') = \frac{\theta_2 S_{yx}}{\bar{Y}\bar{X}} = \theta_2 C_{yx},\ E(e_0e_2') = \frac{\theta_2 S_{xz}}{\bar{Z}\bar{X}} = \theta_2 C_{xz}$$

$$E(e_2e_1') = E(e_1'e_2') = E(e_1e_2') = \frac{\theta_2 S_{xz}}{\bar{Z}\bar{X}} = \theta_2 C_{xz},$$

where

$$\theta_1 = \frac{1}{n} - \frac{1}{N},\ \theta_2 = \frac{1}{n'} - \frac{1}{N},\ \theta_3 = \frac{1}{n} - \frac{1}{n'}$$

$$S_{yx} = \rho_{yx} S_y S_x, \ S_{yz} = \rho_{yz} S_y S_z, \ S_{xz} = \rho_{xz} S_x S_z$$

3.3 Existing estimators

(i) Usual mean estimator

Usual mean estimator in the presence of measurement error is

$$\hat{\bar{Y}}_0^* = \bar{y}^* \tag{3.1}$$

The MSE of $\hat{\bar{Y}}_0^*$ is given by:

$$MSE(\hat{\bar{Y}}_0^*) = \theta_1 (S_y^2 + S_u^2) \tag{3.2}$$

(ii) **Classical ratio estimator**

Classical ratio estimator under double sampling is:

$$\hat{\bar{Y}}_R^* = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*} \right) \tag{3.3}$$

By using first order approximation, the MSE of $\hat{\bar{Y}}_R^*$ is:

$$MSE(\hat{\bar{Y}}_R^*) = \bar{Y}^2 \left[-\theta_2 C_x^2 - 2\theta_3 C_{yx} + \theta_1 (C_{xv}^2 + C_{yu}^2) \right] \tag{3.4}$$

(iii) **Bahl and Tuteja (1991) exponential type estimators**

Bahl and Tuteja (1991) exponential ratio type estimator is:

$$\hat{\bar{Y}}_{BT}^* = \bar{y}^* exp \left(\frac{\bar{x}' - \bar{x}^*}{\bar{x}' + \bar{x}^*} \right) \tag{3.5}$$

Compare (3.14) and (3.22) $MSE(\hat{\bar{Y}}^*_{Roy}) > MSE_{min}(\hat{\bar{Y}}^*_p)$ if

$$\frac{(A_1 + B_1 + C_1)}{D_1} - \frac{A^*}{B^*} > 0$$

3.6 Numerical illustration

Population1 [Source: Murthy et al. (1967)]

Y: Number of workers

X: Fixed capital

Z: Output

$$y = Y + rnorm(80, 0, 1),\ x = X + rnorm(80, 0, 1),\ z = Z + rnorm(80, 0, 3).$$

$N = 80$, $\bar{Y} = 5182.637$, $\bar{X} = 285.125, \bar{Z} = 1126.463$, $S_y = 1835.659$, $S_x = 270.4294$, $S_z = 845.6097, S_u = 1.129815$, $S_v = 1.079988$, $S_w = 3.093483$ $\rho_{yx} = 0.9149811$, $\rho_{yz} = 0.9413055$, $\rho_{xz} = 0.9884207$, $S_{yx} = 454211.4$, $S_{yz} = 1461142$, $S_{xz} = 226029.8$

Population 2 [Source: Koyuncu and Kadilar (2009)]

Y: Number of teachers

X: Number of students both primary and secondary

Z: Number of classes both primary and secondary school

$y = Y + rnorm(80, 0, 1)$, $x = X + rnorm(80, 0, 1)$, $z = Z + rnorm(80, 0, 3)$,

$N = 127$, $\bar{Y} = 703.7402$, $\bar{X} = 20804.59$, $\bar{Z} = 498.2756$, $S_y = 883.8348$, $S_x = 30486.75$, $S_z = 555.5816$, $S_u = 1.122774$, $S_v = 1.018055, S_w = 2.93855$ $S_{yx} = 25237154$, $S_{yz} = 480688.2$, $S_{xz} = 15914648$, $\rho_{yx} = 0.9366086$, $\rho_{yz} = 0.978914$, $\rho_{xz} = 0.9395892$

Table 3.1: MSE of the estimators with Measurement errors ($M.E$) and without measurement error ($W.M.E$)for Population 1

	n'=24		n'=28		n'=32	
	$n=7$		$n=9$		$n=10$	
	$M.E$	$W.M.E$	$M.E$	$W.M.E$	$M.E$	$W.M.E$
$\hat{\bar{Y}}_0^*$	439257.1	439256.9	332284.3	332284.2	294843.8	294843.7
$\hat{\bar{Y}}_R^*$	1213434	1213384	909122.7	909084.6	820828.4	820794.6
$\hat{\bar{Y}}_{BT}^*$	215083.4	215070.6	165253	165243.4	142537.4	142528.9
$\hat{\bar{Y}}_{Reg}^*$	153801.3	153795.3	119591.8	119587.3	100901.8	100897.7
$\hat{\bar{Y}}_K^*$	71351.97	71345.94	53968.86	53964.3	47898.59	47894.54
$\hat{\bar{Y}}_S^*$	66718.66	66712.63	50281.13	50276.56	44920.04	44915.99
$\hat{\bar{Y}}_{Roy}^*$	46545.61	46527.21	35250.36	35236.45	31214.13	31201.77
$\hat{\bar{Y}}_p^*$	11668.31	11664.19	8883.226	8880.074	7790.481	7787.738

Table 3.2: MSE of the estimators with Measurement errors ($M.E$) and without measurement error ($W.M.E$)for population 2

	n'=38		n'=51		n'=64	
	$n=12$		$n=16$		$n=19$	
	$M.E$	$W.M.E$	$M.E$	$W.M.E$	$M.E$	$W.M.E$
$\hat{\bar{Y}}_0^*$	58946.19	58946.09	42671.92	42671.85	34963.05	34962.99
$\hat{\bar{Y}}_R^*$	22233.99	22233.89	15054.71	15054.64	11135.42	11135.36
$\hat{\bar{Y}}_{BT}^*$	25430.86	25430.76	17459.59	17459.52	13210.31	13210.25
$\hat{\bar{Y}}_{Reg}^*$	19874.07	19873.98	13279.43	13279.36	9603.742	9603.685
$\hat{\bar{Y}}_K^*$	7090.388	7090.293	5145.638	5145.569	4230.824	4230.768
$\hat{\bar{Y}}_S^*$	6069.152	6069.057	4495.862	4495.793	3801.604	3801.547
$\hat{\bar{Y}}_{Roy}$	2353.338	2351.981	1700.559	1699.576	1389.827	1389.02
$\hat{\bar{Y}}_p^*$	211.8658	211.7601	151.5801	151.5051	121.4712	121.4121

We have compared the MSEs of estimators (without measurment error) with the MSEs of estimators (in the case of with measurement error). We conclude that the MSE of proposed estimator is smaller among all estimators consider here.

Table 3.3: PRE of the estimators with Measurement errors ($M.E$) and without measurement error ($W.M.E$)for population 1

	n'=24		n'=28		n'=32	
	n=7		n=9		n=10	
	$M.E$	$W.M.E$	$M.E$	$W.M.E$	$M.E$	$W.M.E$
$\hat{\bar{Y}}_0^*$	100	100	100	100	100	100
$\hat{\bar{Y}}_R^*$	36.1995	36.20099	36.54999	36.55151	35.92027	35.92174
$\hat{\bar{Y}}_{BT}^*$	204.2264	204.2384	201.0761	201.0877	206.8536	206.8659
$\hat{\bar{Y}}_{Reg}^*$	285.6003	285.6114	277.8486	277.8591	292.2088	292.2204
$\hat{\bar{Y}}_K^*$	615.6201	615.6719	615.6963	615.7481	615.5584	615.6102
$\hat{\bar{Y}}_S^*$	658.3722	658.4314	660.8529	660.9126	656.3748	656.4337
$\hat{\bar{Y}}_{Roy}^*$	943.7133	944.0861	942.6407	943.0127	944.5845	944.9581
$\hat{\bar{Y}}_p^*$	3764.531	3765.86	3740.581	3741.908	3784.668	3785.999

Table 3.4: PRE of the estimators with Measurement errors ($M.E$) and without measurement error ($W.M.E$)for population 2

	n'=24		n'=28		n'=32	
	n=7		n=9		n=10	
	$M.E$	$W.M.E$	$M.E$	$W.M.E$	$M.E$	$W.M.E$
$\hat{\bar{Y}}_0^*$	100	100	100	100	100	100
$\hat{\bar{Y}}_R^*$	265.1175	265.1182	283.4457	283.4465	313.9806	313.9817
$\hat{\bar{Y}}_{BT}^*$	231.79	231.7905	244.4039	244.4044	264.6649	264.6656
$\hat{\bar{Y}}_{Reg}^*$	296.5984	296.5994	321.3385	321.3396	364.0565	364.0581
$\hat{\bar{Y}}_K^*$	831.3535	831.3633	829.2833	829.2931	826.3886	826.3983
$\hat{\bar{Y}}_S^*$	971.2426	971.2562	949.1375	949.1505	919.6921	919.7043
$\hat{\bar{Y}}_{Roy}^*$	2504.791	2506.232	2509.288	2510.734	2515.641	2517.095
$\hat{\bar{Y}}_p^*$	27822.42	27836.27	28151.39	28165.28	28782.99	28796.95

In Table 3.3 and 3.4 we have compared the efficiency of estimators (without measurment error) with the efficiency of estimators (in the case of with measurement error). We conclude that the efficiency of proposed estimator is maximum among all estimators consider here.

Table 3.5: Conditional values for Population I and Population II

Conditions	Population1	Population2
i	0.01635374	0.119023
ii	0.0451767	0.04489446
iii	0.008007653	0.05134952
iv	0.005726092	0.04012935
v	0.002656466	0.01431678
vi	0.002483966	0.01225472
vii	0.001732914	0.004751816
viii	0.000434417	0.000427796

All the values in Table 3.5 are greater than zero (positive) which shows that proposed estimator perform better than existing estimators.

Conclusion

We have considered the problem of measurement error in the proposed estimator. In Table 3.1-3.2, we compared the MSEs of the suggested estimators (having no measurement error) with MSEs of the estimators (with measurement error). We have also compared efficiency of estimators (without measurement error) with efficiency of estimators (in case of with measurement error). From our empirical study we conclude that MSE is minimum and PRE (percentage relative efficiency) of our proposed estimator $\hat{\bar{Y}}_p^*$ is maximum among all the estimators consider here. We can observe in table that estimators has unexpectedly declined when measurement error is taken into account.

Chapter 4

Estimation of Population Mean In The Presence of Non-Response

Outline

We modified the proposed estimator using two phase sampling in case of non-response in order to get information on two auxiliary variables for estimating population mean. We investigate the adequacy of the suggested estimator and compare it with considered estimators. It is found that suggested estimator performs better than considered estimators.

4.1 Introduction

Most of the time, non-response occurs in surveys related to human behavior and trends as the respondents may refuse, not available at the location at the time of survey or/and cannot provide information due to some other reasons. Hansen and Hurwitz (1946) were first who initiated to deal on the problem of non-response. To tackle with the problem of non-response, they recommended a sub-sampling technique of non-respondents. In the presence of non-response for developing the unbiased estimator of population mean, population is divided into two groups, response group (respondents) and non-response group (non-respondents). For avoiding the bias arises because of non-response, they proposed to take a sub-sample of the non-responding units.Consider a population $U = \{S_1, S_2, \ldots, S_N\}$ of size N and a sample of size n is taken from the population applying simple random sampling without replacement (SRSWOR) scheme. We assume that y, x and z be the study and two auxiliary variables with respective population $\bar{Y}$, $\bar{X}$ and $\bar{Z}$. Let us consider the position in which the non-response exist both on the study as well as the auxiliary variables. It is observed that in the sample of n units, there are n_1 respondent and n_2 non-respondent units for the study and the auxiliary variables.By Using Hansen and Hurwitz (1946) technique of sub sampling of non-respondents, we select a sub-sample of h_2 non-respondent units from n_2 units in such a way that $h_2 = \frac{n_2}{k}, (k > 1)$and collect the information on sub-sample by personal interview method. It is find out that auxiliary information is used to increase the efficiency of estimators. Several authors like Cochran (1977), Rao (1986), Khare and Srivastava (1993, 1997), Kiregyera (1980) used Hansen and Hurwitz (1946) technique to estimate the population mean. .we use two phase sampling scheme in order to improve the efficiency of estimator when information is not available on auxiliary variable. In this scheme a large sample of size n' is selected from population of size N by using simple random sampling without replacement scheme. At first phase, auxiliary variables(x, z) are measured. Then relatively small sample of size n $(n < n')$ is selected. At the second stage y, x and z are measured. Let $\bar{x}'$ and $\bar{z}'$ be the first phase sample means of variables x and z respectively. Let us assume that the situation non-response is observed on study variable y and on both the auxiliary variables (x, z).

4.2 Notation and symbols

$$e_0^{**} = \frac{\bar{y}^{**} - \bar{Y}}{\bar{Y}},\ e_1^{**} = \frac{\bar{x}^{**} - \bar{X}}{\bar{X}}, e_2^{**} = \frac{\bar{z}^{**} - \bar{Z}}{\bar{Z}},\ e_1' = \frac{\bar{x}' - \bar{X}}{\bar{X}},\ e_2' = \frac{\bar{z}' - \bar{Z}}{\bar{Z}}$$

$$E(e_i),\ \ i = (0, 1, 2), \quad E(e_j') = 0,\ j = (1, 2)$$

$$E(e_o^{**2}) = \theta_1 C_y^{*2},\ E(e_1^{**2}) = \theta_1 C_x^{*2},\ E(e_2^{**2}) = \theta_1 C_z^{*2},\ E(e_o^{**} e_1^{**}) = \theta_1 C_{yx}^{*}$$

$$E(e_o^{**}e_2^{**})=\theta_1 C_{yz}^*,\ E(e_1^{**}e_2^{**})=\theta_1 C_{xz}^*,\ E(e_o^{**}e_1')=\frac{\theta_2 S_{yx}}{\bar{Y}\bar{X}}=\theta_2 C_{yx},$$

$$E(e_2^{**}e_1')=E(e_1^{**}e_2')=E(e_1'e_2')=\frac{\theta_2 S_{xz}}{\bar{Z}\bar{X}}=\theta_2 C_{xz},\ E(e_o^{**}e_2')=\frac{\theta_2}{\bar{Y}\bar{z}}S_{yz}=\theta_2 C_{yz},$$

$$E(e_1^{**}e_1')=E(e_1'^2)=\frac{\theta_2}{\bar{X}^2}S_x^2=\theta_2 C_x^2,\ E(e_2^{**}e_2')=E(e_2'^2)=\frac{\theta_2}{\bar{Z}^2}S_z^2=\theta_2 C_z^2,$$

where

$$C_y'^{*2}=\frac{1}{\bar{Y}^2}(S_y^2+\frac{\lambda}{\theta_1}S_{y(2)}^2),\ C_x'^{*2}=\frac{1}{\bar{X}^2}(S_x^2+\frac{\lambda}{\theta_1}S_{x(2)}^2),\ C_z'^{*2}=\frac{1}{\bar{Z}^2}(S_z^2+\frac{\lambda}{\theta_1}S_{z(2)}^2)$$

$$C_{yx}^*=\frac{1}{\bar{Y}\bar{X}}(S_{yx}+\frac{\lambda}{\theta_1}S_{yx(2)}^2),\ C_{xz}^*=\frac{1}{\bar{Z}\bar{X}}(S_{xz}+\frac{\lambda}{\theta_1}S_{xz(2)}^2),\ C_{yz}^*=\frac{1}{\bar{Y}\bar{Z}}(S_{yz}+\frac{\lambda}{\theta_1}S_{yz(2)}^2)$$

Further

$$\bar{Y}=\frac{1}{N}\sum_{i=1}^{N}y_i,\ \bar{X}=\frac{1}{N}\sum_{i=1}^{N}x_i,\ \bar{Z}=\frac{1}{N}\sum_{i=1}^{N}z_i$$

$$S_y^2=\sum_{i=1}^{N}\frac{(y_i-\bar{Y})^2}{N-1},\ S_x^2=\sum_{i=1}^{N}\frac{(x_i-\bar{X})^2}{N-1},\ S_z^2=\sum_{i=1}^{N}\frac{(z_i-\bar{Z})^2}{N},$$

$$S_{yx}=\sum_{i=1}^{N}\frac{(y_i-\bar{Y})(x_i-\bar{X})}{N-1},\ S_{yz}=\sum_{i=1}^{N}\frac{(y_i-\bar{Y})(z_i-\bar{Z})}{N-1},\ S_{xz}=\sum_{i=1}^{N}\frac{(x_i-\bar{X})(z_i-\bar{Z})}{N-1}$$

$$S_{y(2)}^2=\sum_{i=1}^{N_2}\frac{(y_i-\bar{Y}_2)^2}{N_2},\ S_{x(2)}^2=\sum_{i=1}^{N_2}\frac{(x_i-\bar{X}_2)^2}{N_2},\ S_{z(2)}^2=\sum_{i=1}^{N_2}\frac{(z_i-\bar{Z}_2)^2}{N_2}$$

$$S_{yx(2)}=\sum_{i=1}^{N_2}\frac{(y_i-\bar{Y}_2)(x_i-\bar{X}_2)}{1-N_2},\ S_{yz(2)}=\sum_{i=1}^{N_2}\frac{(y_i-\bar{Y}_2)(z_i-\bar{Z}_2)}{1-N_2},$$

$$S_{xz(2)}=\sum_{i=1}^{N_2}\frac{(x_i-\bar{X}_2)(z_i-\bar{Z}_2)}{1-N_2},\ \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i,\ \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \bar{z}=\frac{1}{n}\sum_{i=1}^{n}z_i,$$

$$\bar{Y}_2=\frac{1}{N_2}\sum_{i=1}^{N_2}y_i,\ \bar{X}_2=\frac{1}{N_2}\sum_{i=1}^{N_2}x_i,\ \bar{Z}_2=\frac{1}{N_2}\sum_{i=1}^{N_2}z_i,$$

where,

$$\theta_1=(\frac{1}{n}-\frac{1}{N}),\theta_2=(\frac{1}{n'}-\frac{1}{N}),\lambda=\frac{W(K-1)}{n}$$

4.3 Existing estimators

(i) Usual mean estimator

The simple mean for estimating population mean $\bar{Y}$ under nonresponse is

$$\hat{\bar{Y}}_0^{**} = \frac{n_1\bar{y}_{n1} + n_2\bar{y}'_{h2}}{n}. \tag{4.1}$$

The variance of $\hat{\bar{Y}}_0^{**}$, is

$$Var(\hat{\bar{Y}}_0^{**}) = \theta_1 S_y^2 + \lambda S_{y(2)} \tag{4.2}$$

(ii) Ratio estimator

Classical ratio estimator under double sampling, is;

$$\hat{\bar{Y}}_R^{**} = \bar{y}(\frac{\bar{x}'}{\bar{x}^{**}}). \tag{4.3}$$

By using first order of approximation , the MSE of $\hat{\bar{Y}}_R^{**}$,is;

$$MSE(\hat{\bar{Y}}_R^{**}) = \bar{Y}^2\left[\theta_1(C_x^{*2} + C_y^{*2} - 2C_{yx}^*) + \theta_2(-C_x^2 + 2C_{yx})\right] \tag{4.4}$$

(iii) Bahl and Tuteja (1991) ratio exponential estimator

Bahl and Tuteja (1991) ratio exponential estimator under double sampling, is:

$$\hat{\bar{Y}}_{BT}^{**} = \bar{y}exp\left(\frac{\bar{x}' - \bar{x}^{**}}{\bar{x}' + \bar{x}^{**}}\right) \tag{4.5}$$

The MSE of $\hat{\bar{Y}}_{BT}^{**}$ is given below by using first oder of approximation

$$MSE(\hat{\bar{Y}}_{BT}^{**}) = \bar{Y}^2\left[\theta_2(\frac{-C_x^2}{4} + C_{yx}) + \theta_1(\frac{C_x^{*2}}{4} + C_y^{*2} - C_{yx}^*)\right] \tag{4.6}$$

(iv) Traditional regression estimator

Traditional regression estimator is given below

$$\hat{\bar{Y}}^{**}_{Reg} = \bar{y} + b^{**}_{yx}\left(\bar{x}' - \bar{x}^{**}\right), \tag{4.7}$$

where b^{**}_{yx} is the sample regression coefficient of y on x.

By using first order of approximation, the MSE of $\hat{\bar{Y}}^{**}_{Reg}$, is;

$$MSE(\hat{\bar{Y}}^{**}_{Reg}) = \bar{Y}^2\left[\theta_1\left\{C_y^{*2} + C_y^2\rho_{yx}^2(\frac{-2\bar{X}\bar{Y}C^*_{yx}}{S_{yx}} + \frac{\bar{X}^2C_x^{*2}}{S_x^2})\right\} + \theta_2 C_y^2\rho_{yx}^2\right] \tag{4.8}$$

(v) **Kiregyera (1984) ratio-regression type estimator**

Kiregyera (1984) ratio-regression type estimator is

$$\hat{\bar{Y}}^{**}_K = \bar{y} + b^{**}_{yx}\left(\frac{\bar{x}'}{\bar{z}'}\bar{Z} - \bar{x}^{**}\right). \tag{4.9}$$

The MSE of $\hat{\bar{Y}}^{**}_K$, up to first oder of approximation is

$$MSE(\hat{\bar{Y}}^{**}_K) = \bar{Y}^2\left[\theta_1(C_y^{*2} + \frac{\rho_{yx}^2C_y^2\bar{X}^2C_x^{*2}}{S_x^2} - \frac{2C_y^2\bar{Y}\bar{X}C^*_{yx}\rho_{yx}}{S_{yx}}) + C_y^2\theta_2\left\{(\frac{C_z\rho_{yx}}{C_x} - \rho_{yz})^2 - \rho_{yz}^2 + \rho_{yx}^2\right\}\right] \tag{4.10}$$

(vi) **Sahoo et al. (1993) regression-type estimator**

Sahoo et al. (1993) regression-type estimator, is given by:

$$\bar{Y}^{**}_S = \bar{y} + b^{**}_{yx}(\bar{x}' - \bar{x}^{**}) + b^{**}_{yz}(\bar{Z} - \bar{z}'). \tag{4.11}$$

The MSE of $\hat{\bar{Y}}_S^{**}$ ny using first order of approximation, is:

$$MSE(\hat{\bar{Y}}_S^{**}) = \bar{Y}^2\left[\theta_1(C_y^{*2} - \frac{2\rho_{yx}^2 C_y^2 \bar{X}\bar{Y}C_{yx}^*}{S_{yx}} + \frac{\rho_{yx}^2 C_y^2 \bar{X}^2 C_x^{*2}}{S_x^2}) + \theta_2 C_y^2(\rho_{yx}^2 - \rho_{yz}^2)\right] \quad (4.12)$$

(vii) **Roy (2003) regression type estimator**

Roy(2003) regression type estimator of $\hat{\bar{Y}}_{Reg}^{**}$, is:

$$\hat{\bar{Y}}_{Roy}^{**} = \bar{y}^{**} + k_1\left[\bar{x}' + k_2(\bar{Z} - \bar{z}') - \left\{\bar{x}^{**} + k_3(\bar{Z} - \bar{z}^{**})\right\}\right] \quad (4.13)$$

The MSE of $\hat{\bar{Y}}_{Roy}^{**}$, up to first order of approximation, is

$$\hat{\bar{Y}}_{Roy}^{**} = \frac{\bar{Y}^2(A_1 + B_1 + C_1)}{D_1}, \quad (4.14)$$

where

$D_1 = S_z^2(\bar{Z}^2\bar{X}^2(-C_x^{*2}C_z^{*2} + C_{xz}^{*2})\theta_1^2 + \theta_2(S_x^2\bar{Z}^2C_z^{*2} + S_z^2\bar{X}^2C_x^{*2} - 2S_{xz}\bar{Z}\bar{X}C_{xz}^*)\theta_1 + (-S_x^2S_z^2 + S_{xz}^2)\theta_2^2)$

$A_1 = (\bar{Z}^2\bar{X}^2\bar{Y}^2(C_{xz}^{*2}C_y^*2 - 2C_{xz}^*C_{yz}^*C_{yx}^*) + \bar{X}^2\bar{Y}^2(-C_x^{*2}C_y^{*2} + C_{yx}^{*2})\bar{Z}^2C_z^{*2} + \bar{X}C_x^{*2}\bar{Z}^2\bar{Y}^2C_{yz}^{*2})S_z^2\theta_1^3$

$$\begin{aligned}B_1 = -\theta_2\{&\bar{X}^2\bar{Y}^2(-C_x^{*2}C_y^{*2} + C_{yx}^{*2})S_z^4 + (\bar{Y}\bar{Z}\bar{X}^2(2C_x^{*2}C_{yz}^* - 2C_{xz}^*C_{yx}^*)S_{yz} + \\ &\bar{X}\bar{Z}\bar{Y}^2(2C_{xz}^*C_y^{*2} - 2C_{yx}^*C_{yz}^*)S_{xz} - 2\bar{X}\bar{Z}C_{xz}^*S_{yx}\bar{Y}C_{yz}^* + (-S_x^2\bar{Y}^2C_y^{*2} + \\ &2S_{yx}\bar{Y}\bar{X}C_{yx}^*)\bar{Z}^2C_z^{*2} + \bar{Y}^2\bar{Z}^2C_{yz}^{*2}S_x^2)S_z^2 + \bar{Z}^2\bar{X}^2(-C_x^{*2}C_z^{*2} + C_{xz}^{*2})S_{yz}^2\}\theta_1^2\end{aligned}$$

$$\begin{aligned}C_1 = 2\theta_2^2\{&(-\tfrac{1}{2}S_x^2\bar{Y}^2C_y^{*2} + S_{yx}\bar{X}\bar{Y}C_{yx}^*)S_z^4 + ((S_x^2\bar{Y}\bar{Z}C_{yz}^* - S_{xz}\bar{Y}\bar{X}C_{yx}^* \\ &- S_{yx}\bar{X}\bar{Z}C_{xz}^*)S_{yz} - S_{yx}S_{xz}\bar{Z}\bar{Y}C_{yz}^* + \tfrac{1}{2}S_{xz}^2\bar{Y}^2C_y^{*2} + \tfrac{1}{2}S_{yx}^2\bar{Z}^2C_z^{*2})S_z^2 \\ &- \tfrac{1}{2}S_{yz}^2(S_x^2\bar{Z}^2C_z^2 - 2S_{xz}\bar{X}\bar{Z}C_{xz}^*)\}\theta_1 - \theta_2^3(S_{yx}S_z^2 - S_{xz}S_{yz})^2\end{aligned}$$

4.4 Proposed estimator

$$\hat{\bar{Y}}_p^{**} = k_0\bar{y} + k_1\left[\bar{x}' + k_2 exp\left(\frac{\bar{Z}-\bar{z}'}{\bar{Z}-\bar{z}'}\right) - \left\{\bar{x}^{**} + k_3 exp\left(\frac{\bar{Z}-\bar{z}^{**}}{\bar{Z}+\bar{z}^{**}}\right)\right\}\right] \tag{4.15}$$

where $k_i(i = 0,1,2,3)$ are constants

In terms of error, we have

$$\hat{\bar{Y}}_p^{**} = k_0\bar{Y} + k_0\bar{Y}e_0^{**} + k_1\bar{X}e_1' + k_1k_2 - \frac{k_1k_2e_2'}{2} + \frac{3k_1k_2e_2'^2}{8} - k_1\bar{X}e_1^{**} - k_1k_3 + \frac{k_1k_3e_2^{**}}{2} - \frac{3k_3k_1e_2^{**2}}{8} \tag{4.16}$$

The Bias and Mean square error of $\hat{\bar{Y}}_p^{**}$ are

$$Bias(\hat{\bar{Y}}_p^{**}) = \bar{Y}k_0 + k_1k_2 + \frac{3}{8}k_1k_2\theta_2C_z^2 - k_1k_3 - \frac{3}{8}k_1k_3\theta_1C_z^{*2} - \bar{Y} \tag{4.17}$$

$$\text{MSE}(\hat{\bar{Y}}_p^{**}) = \left[\begin{array}{l} (k_0-1)^2\bar{Y}^2 + k_1\left\{(C_{yz}^*k_0\theta_1 - 2k_0 + 2)k_3 - (2 + (C_{yz}\theta_2 - 2)k_0)k_2\right\}\bar{Y} - k_1^2k_3q\bar{X} \\ + \frac{1}{4}\left\{(\bar{Z}^2C_z^{*2}\theta_1 + 4)k_3^2 + (-2C_z^2\theta_2 - 8)k_2k_3 + (C_z^2\theta_2 + 4)k_2^2 + 4\bar{X}^2C_x^{*2}\theta_1 - 4S_x^2\theta_2)\right\}k_1^2 \\ + 2k_0(S_{yx}\theta_2 - \bar{Y}\bar{X}C_{yx}^*\theta_1)k_1 + k_0^2\theta_1\bar{Y}^2C_y^{*2} \end{array}\right] \tag{4.18}$$

The MSE of $\hat{\bar{Y}}_p^{**}$ up to first oder of approximation ,is Differentiate eq w.r.t k_0, k_1, k_2, k_3,we get

$$\text{k}_{0(opt)} = \frac{-\theta_2\left\{(C_{xz}\theta_2 - C_{xz}^*\theta_1)^2\bar{X}^2 + (-S_x^2\theta2 + \bar{X}^2C_x^{*2}\theta_1)(-C_z^2\theta_2 + C_z^{*2}\theta_1)\right\}\bar{Y}^2(C_z^2 + 2C_{yz})}{\begin{array}{l} f - 2\bar{X}(C_z^2\theta_2 + 4)(-C_{yz}\theta_2 + C_{yz}^*\theta_1)q(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)\bar{Y} + (C_z^2\theta_2 + 4)\left\{\theta_1\bar{Y}C_y^{*2}(-C_{yz}\theta_2\right. \\ \left. + C_{xz}^*\theta_1)^2\bar{X}^2 + m(-C_z^2\theta_2 + C_z^{*2}\theta_1)\right\} \end{array}} \tag{4.19}$$

$$\text{k}_{1(opt)} = \frac{-\theta_2\bar{Y}^2\left\{-\bar{X}(-C_{yz}\theta_2 + C_{yz}^*\theta_1)q\bar{Y} + (-C_z^2\theta_2 + C_z^{*2}\theta_1)(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)\right\}(C_z^2 + 2C_{yz})}{\begin{array}{l} f - 2\bar{X}(C_z^2\theta_2 + 4)(-C_{yz}\theta_2 + C_{yz}^*\theta_1)q(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)\bar{Y} + (C_z^2\theta_2 + 4)\left\{\theta_1\bar{Y}C_y^{*2}(-C_{yz}\theta_2\right. \\ \left. + C_{xz}^*\theta_1)^2\bar{X}^2 + m(-C_z^2\theta_2 + C_z^{*2}\theta_1)\right\} \end{array}} \tag{4.20}$$

$$\mathrm{k}_{2(opt)} = \frac{\begin{array}{l}\Big\{-2\theta_2 C_{yz} q^2 \bar{X}^2 + 2l\theta_1(-2C_{yz}^2\theta_1 + ((C_z^{*2} + 2C_{yz}^*)C_{yz} - \bar{Y}\bar{Z}C_{yz}^* C_z^2)\theta_2)\Big\}\bar{Y}^2 + 8(C_{yz}^*\theta_1 + \\ \frac{1}{4}\theta_2(C_z^2 - 2C_{yz}))(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)\bar{X}q\bar{Y} - 4\theta_1\bar{Y}^2 C_y^{*2} q^2 \bar{X}^2 - 4\Big\{(-\bar{X}^2 C_x^{*2}\bar{Y}^2 C_y^{*2} \\ + \bar{X}\bar{Y}C_{yx}^{*2})\theta_1^2 + (S_x^2\bar{Y}^2 C_y^{*2} - 2S_{yx}\bar{Y}\bar{X}C_{yx}^*)\theta_2\theta_1 + \theta_2^2 S_{yx}^2\Big\}(-C_z^2\theta_2 + C_z^{*2}\theta_1)\end{array}}{(C_z^2 + 2C_{yz})\theta_2\bar{Y}\Big\{-\bar{X}(-C_{yz}\theta_2 + C_{yz}^*\theta_1)q\bar{Y} + (-C_z^2\theta_2 + C_z^{*2}\theta_1)(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)\Big\}} \tag{4.21}$$

$$\mathrm{k}_{3(opt)} = \frac{(2(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1))q\bar{X} - 2\bar{Y}l(-C_{yz}\theta_2 + C_{yz}^*\theta_1)}{-\bar{X}(-C_{yz}\theta_2 + C_{yz}^*\theta_1)q\bar{Y} + (-C_z^2\theta_2 + C_z^{*2}\theta_1)(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)}, \tag{4.22}$$

where $q = -C_{xz}\theta_2 + C_{yx}^*\theta_1$,

$$\mathrm{f} = -\theta_2(C_{yz}^2\theta_2 - C_z^2 - 4C_{yz})q^2\bar{X}^2 + l\Big\{(C_{yz}(-2C_z^2 C_{yz}^* + C_{yz}C_z^{*2})\theta_1 + (C_z^2 + 2C_{yz})^2)\theta_2^2 - \theta_1(-C_{yz}^{*2}\theta_1 C_Z^2 + C_z^{*2}C_z^2 + 4C_{yz}(C_z^{*2} + 2C_{yz}^*))\theta_2 + 4C_{yz}^{*2}\theta_1^2\Big\}\bar{Y}^2,$$

$m = \bar{Y}^2\bar{X}^2(-C_x^{*2}C_y^{*2} + C_{yx}^{*2})\theta_1^2 + (S_x^2\bar{Y}^2 C_y^{*2} - 2S_{yx}\bar{Y}\bar{X}C_{yx}^*)\theta_2\theta_1 + \theta_2^2 S_{yx}^2$, $l = -S_x^2\theta_2 + \bar{X}^2 C_x^{*2}\theta_1$.

The minimum MSE of proposed estimator is

$$MSE_{min}(\hat{\bar{Y}}_p^{**}) = \frac{\bar{Y}^2 A^*}{B^*}, \tag{4.23}$$

where

$$A^* = \theta_2\{-C_{yz}^2\theta_2 q^2\bar{X}^2 + \theta_1 l(C_{yz}(-2C_z^2 C_{yz}^* + C_{yz}C_z^{*2})\theta_2 + C_{yz}^{*2}\theta_1 C_z^2)\}\bar{Y}^2 - 2\bar{X}C_z^2 (-C_{yz}\theta_2 + C_{yz}^*\theta_1)q(-S_{yx}\theta_2 + \bar{X}\bar{Y}C_{yx}^*\theta_1)\bar{Y} + C_z^2\{\theta_1\bar{Y}^2 C_y^{*2}q^2\bar{X}^2 + m(-C_z^2\theta_2 + C_z^{*2}\theta_1)\}$$

$$B^* = f - 2\bar{X}(C_z^2\theta_2 + 4)(-C_{yz}\theta_2 + C_{yz}^*\theta_1)q(-S_{yx}\theta_2 + \bar{Y}\bar{X}C_{yx}^*\theta_1)\bar{Y} + (C_z^2\theta_2 + 4)\{\theta_1\bar{Y}^2 C_y^{*2}(-C_{yz}\theta_2 + C_{xz}^*\theta_1)^2\bar{X}^2 + m(-C_z^2\theta_2 + C_z^{*2}\theta_1)\}$$

4.5 Efficiency comparisons

Comparing (4.2)and (4.23), $MSE(\hat{\hat{Y}}_0^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$, if

$$S_y^2 + S_{y(2)} - \frac{A^*}{B^*} > 0$$

Comparing (4.4)and (4.23), $MSE(\hat{\hat{Y}}_R^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$, if

$$\left[\theta_1(C_x^{*2} + C_y^{*2} - 2C_{yx}^*) + \theta_2(-C_x^2 + 2C_{yx})\right] - \frac{A^*}{B^*} > 0$$

Comparing (4.6)and (4.23), $MSE(\hat{\hat{Y}}_{BT}^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$, if

$$\left[\theta_2(\frac{-C_x^2}{4} + C_{yx}) + \theta_1(\frac{C_x^{*2}}{4} + C_y^{*2} - C_{yx}^*)\right] - \frac{A^*}{B^*} > 0$$

Comparing (4.8)and (4.23), $MSE(\hat{\hat{Y}}_{BT}^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$, if

$$\left[\theta_1(C_y^{*2} + \frac{\rho_{yx}^2 C_y^2 \bar{X}^2 C_x^{*2}}{S_x^2} - \frac{2C_y^2 C_{yx}^* \bar{X}\bar{Y}\rho_{yx}}{S_{yx}}) + C_y^2\theta_2\left\{(\frac{C_z\rho_{yx}}{C_x} - \rho_{yz})^2 - \rho_{yz}^2 + \rho_{yx}^2\right\}\right] - \frac{A^*}{B^*} > 0$$

Comparing (4.10)and (4.23), $MSE(\hat{\hat{Y}}_{Reg}^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$, if

$$\left[\theta_1(C_y^{*2} + \frac{\rho_{yx}^2 C_y^2 \bar{X}^2 C_x^{*2}}{S_x^2} - \frac{2C_y^2 \bar{Y}\bar{X} C_{yx}\rho_{yx}}{S_{yx}}) + C_y^2\theta_2\left\{(\frac{C_z\rho_{yx}}{C_x} - \rho_{yz})^2 - \rho_{yz}^2 + \rho_{yx}^2\right\}\right] - \frac{A^*}{B^*} > 0$$

Comparing (4.12)and (4.23), $MSE(\hat{\hat{Y}}_S^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$, if

$$\left[\theta_1(C_y^{*2} - \frac{2\rho_{yx}^2 C_y^2 \bar{X}\bar{Y} C_{yx}^*}{S_{yx}} + \frac{\rho_{yx}^2 C_y^2 \bar{X}^2 C_x^{*2}}{S_x^2}) + \theta_2 C_y^2(\rho_{yx}^2 - \rho_{yz}^2)\right] - \frac{A^*}{B^*} > 0$$

Comparing (4.14)and (4.23),

$$MSE(\hat{\hat{Y}}_{Roy}^{**}) - MSE_{min}(\hat{\hat{Y}}_p^{**}) > 0$$

$$\frac{A_1 + B_1 + C_1}{D_1} - \frac{A^*}{B^*} > 0$$

Population I [Source: Koyuncu and Kadilar (2009)]

y: Number of teachers

x: Number of students both primary and secondary

z: Number of classes both primary and secondary school

$$N = 127,\ n' = 70,\ n = 31,\ \bar{Y} = 703.7402,\ \bar{X} = 20804.59,\ \bar{Z} = 498.2756,\ S_y = 883.8348$$

$$S_x = 30486.75, S_z = 555.5816,\ \rho_{yx} = 0.9366086,\ \rho_{yz} = 0.978914$$

$$\rho_{xz} = 0.9395892,\ S_{yx} = 25237154,\ S_{yz} = 480688.2,\ S_{xz} = 15914648.$$

For $W = .1$

$$S_y = 510.57,\ S_x = 9446.927,\ S_z = 303.9178,\ \rho_{yx} = 0.9366086,\ \rho_{yz} = 0.978914$$

$$\rho_{zx} = 0.9395892,\ S_{yx} = 4804568,\ S_{yz} = 154106.3,\ S_{xz} = 2842719.$$

For $W = .2$

$$S_y = 392.5236, S_x = 7379.522,\ S_z = 243.6362,\ \rho_{yx} = 0.9366086,\ \rho_{yz} = 0.978914,$$

$$\rho_{zx} = 0.9395892,\ S_{yx} = 2883213,\ S_{yz} = 94593.51,\ S_{xz} = 1779393.$$

For $W = .3$

$$S_y = 500.264,\ S_x = 14017.99,\ S_z = 284.4409,\ \rho_{yx} = 0.9366086,$$

$$\rho_{yz} = 0.978914,\ \rho_{zx} = 0.9395892,\ S_{yx} = 6759674,\ S_{yz} = 138575.4, S_{xz} = 3631211$$

Population II [Source: Kadilar and Cingni 2003]

y: Apple production amount in 1999

x: The number of apple trees in 1999

z: Apple production amount in 1998

$$N = 854,\ n' = 342, n = 137, \bar{Y} = 2930.126, \bar{X} = 37600.12, \bar{Z} = 37474.92,\ S_y = 17105.73$$

$$S_x = 144793.7,\ S_z = 135589.7, \rho_{yx} = 0.9165007, \rho_{yz} = 0.9165057, \rho_{zx} = 0.9975857$$

$$S_{yx} = 2269991082,\ S_{yz} = 2125708059, S_{xz} = 19585141612.$$

$For\ W = .1$

$$S_y = 7093.424, S_x = 53011.33, S_z = 52950.28, \rho_{yx} = 0.8204176,$$

$$\rho_{yz} = 0.8204002, \rho_{zx} = 0.9992243,\ S_{yx} = 308503153,\ S_{yz} = 308141307, S_{xz} = 2804787807.$$

$For\ W = .2$

$$S_y = 10336.65,\ S_x = 56947.32,\ S_z = 54831.9,\ \rho_{yx} = 0.8417329,$$

$$\rho_{yz} = 0.8231078,\ \rho_{zx} = 0.9983789,\ S_{yx} = 495481351,\ S_{yz} = 466519393,\ S_{xz} = 3117468239$$

$For\ W = .3$

$$S_y = 13531.69,\ S_x = 92925.4,\ S_z = 91955.12,\ \rho_{yx} = 0.9232772,$$

$$\rho_{yz} = 0.9167727,\ \rho_{zx} = 0.9994398,\ S_{yx} = 1160963847, S_{yz} = 1140748040,\ S_{xz} = 8540180273$$

Table 4.1: : MSEs of estimators for Population I

W	K	$\hat{Y}_0^{**}$	$\hat{Y}_R^{**}$	$\hat{Y}_{BT}^{**}$	$\hat{Y}_{Reg}^{**}$	$\hat{Y}_K^{**}$	$\hat{Y}_S^{**}$	$\hat{Y}_{Roy}^{**}$	$\hat{Y}_p^{**}$
	2	19888.85	7597.775	8882.66	6943.614	2499.078	2144.022	781.4065	70.60782
.1	2.5	20309.3	7658.671	9082.16	7049.361	2604.825	2249.77	790.9972	71.4116
	3	20729.76	7719.567	9281.66	7155.109	2710.573	2355.517	800.2252	72.18366
	3.5	21150.21	7780.463	9481.16	7260.856	2816.32	2461.264	809.1404	72.92831
	2	20041.97	7613.59	8948.977	6975.02	2530.484	2175.428	785.4115	70.94365
.2	2.5	20538.98	7682.394	9181.637	7096.47	2651.934	2296.878	797.3002	71.93908
	3	21036	7751.198	9414.296	7217.92	2773.384	2418.328	808.996	72.91626
	3.5	21533.02	7820.002	9646.955	7339.37	2894.834	2539.779	820.5308	73.87793
	2	21469.85	7648.224	9236.762	7003.598	2559.062	2204.006	867.0096	77.73248
.3	2.5	22680.8	7734.345	9613.314	7139.338	2694.802	2339.746	915.4308	81.71357
	3	23891.76	7820.465	9989.865	7275.078	2830.542	2475.486	915.4308	85.45459
	3.5	25102.72	7906.586	10366.42	7410.817	2966.281	2611.225	1005.016	88.98769

Table 4.2: : MSEs of estimators for population II

W	K	$\hat{Y}_0^{**}$	$\hat{Y}_R^{**}$	$\hat{Y}_{BT}^{**}$	$\hat{Y}_{Reg}^{**}$	$\hat{Y}_K^{**}$	$\hat{Y}_S^{**}$	$\hat{Y}_{Roy}^{**}$	$\hat{Y}_p^{**}$
	2	1829908	816371.3	1180761	729826.5	300537.1	298963	297653.3	13691.62
.1	2.5	1848272	823415.1	1191907	735832.2	306542.8	304968.8	303656.6	13889.45
	3	1866635	830459	1203054	741837.9	312548.5	309659.7	298880	14085.07
	3.5	1884999	837502.9	1214201	747843.6	318554.3	316980.2	315662.7	14278.52
	2	1949160	874278	1265267	772660.6	343371.3	341797.2	340876.8	15068.14
.2	2.5	2027150	910275.2	1318667	800083.4	370794.1	369220	368011.7	15878.73
	3	2105140	946272.4	1372066	827506.2	398216.9	396642.8	394843.8	16643.05
	3.5	2183130	982269.6	1425466	854929	425639.7	424065.6	421385.1	17365.12
	2	2194144	921850.2	1390024	789933.6	360644.3	359070.2	355050.7	15496.44
.3	2.5	2394626	981633.5	1505803	825992.9	396703.5	395129.5	388788.8	16473.67
	3	2595108	1041417	1621581	862052.1	432762.8	431188.7	422055.1	17382.94
	3.5	2795590	1101200	1737360	898111.4	468822.1	467248	454920.5	18232.63

Tables 4.1-4.2 present MSE comparison of the suggested estimator with considered estimators and observe that their MSE increases by increasing the values of K and W.

Table 4.3: : PREs of estimators with respect to $\hat{\bar{Y}}_0^{**}$ for population I

W	K	$\hat{\bar{Y}}_0^{**}$	$\hat{\bar{Y}}_R^{**}$	$\hat{\bar{Y}}_{BT}^{**}$	$\hat{\bar{Y}}_{Reg}^{**}$	$\hat{\bar{Y}}_K^{**}$	$\hat{\bar{Y}}_S^{**}$	$\hat{\bar{Y}}_{Roy}^{**}$	$\hat{\bar{Y}}_p^{**}$
	2	100	261.772	223.9064	286.4337	795.8474	927.6419	2545.263	28168.05
.1	2.5	100	265.1805	223.6175	288.1013	779.68	902.7281	2567.557	28439.78
	3	100	268.5352	223.341	289.7197	764.774	880.0512	2590.49	28718.07
	3.5	100	271.8374	223.0762	291.2909	750.9874	859.323	2613.911	29001.37
	2	100	263.2394	223.9582	287.3393	792.0213	921.2887	2551.779	28250.55
.2	2.5	100	267.3514	223.6963	289.4254	774.4908	894.2131	2576.067	28550.52
	3	100	271.3903	223.4474	291.4413	758.4957	869.8571	2600.26	28849.53
	3.5	100	275.3582	223.2105	293.3905	743.842	847.8304	2624.279	29146.75
	2	100	280.7168	232.4391	306.5546	838.9733	974.1282	2476.31	27620.18
.3	2.5	100	293.2479	235.9312	317.6878	841.6501	969.3703	2477.61	27756.47
	3	100	305.503	239.16	328.4056	844.0703	965.1342	2477.61	27958.43
	3.5	100	317.4912	242.1542	338.7307	846.2689	961.3385	2497.744	28209.2

Table 4.4: : PREs of estimators with respect to $\hat{\bar{Y}}_0^{**}$ for population II

W	K	$\hat{\bar{Y}}_0^{**}$	$\hat{\bar{Y}}_R^{**}$	$\hat{\bar{Y}}_{BT}^{**}$	$\hat{\bar{Y}}_{Reg}^{**}$	$\hat{\bar{Y}}_K^{**}$	$\hat{\bar{Y}}_S^{**}$	$\hat{\bar{Y}}_{Roy}^{**}$	$\hat{\bar{Y}}_p^{**}$
	2	100	224.1514	154.977	250.7319	608.8792	612.085	614.7784	13365.17
.1	2.5	100	224.4641	155.0684	251.1812	602.9407	606.0528	608.6717	13307.02
	3	100	224.7715	155.1581	251.6231	597.2306	600.2536	602.8021	13252.58
	3.5	100	225.0738	155.2461	252.0579	591.7357	594.6742	597.1561	13201.65
	2	100	222.9452	154.0513	252.266	567.6538	570.2681	571.8079	12935.64
.2	2.5	100	222.6964	153.7273	253.3674	546.7051	549.0359	550.8385	12766.45
	3	100	222.4666	153.4285	254.3957	528.6416	530.7396	533.1578	12648.77
	3.5	100	222.2537	153.152	255.3581	512.9057	514.8096	518.0844	12571.92
	2	100	238.0153	157.8493	277.7631	608.3957	611.0627	617.9805	14159.02
.3	2.5	100	243.943	159.0265	289.9088	603.6311	606.0358	615.9195	14536.08
	3	100	249.1901	160.0356	301.0384	599.6605	601.8496	614.8742	14929.05
	3.5	100	253.8675	160.9102	311.2742	596.3007	598.3096	614.5227	15332.89

Table 4.3-4.4 presents PRE comparison of the suggested estimator with considered estimators for different combinations of W, and K for specified sample size. In Table 4.3, we found that the relative efficiency of proposed estimator $\bar{Y}_p^{**}$ increases as K and W are increases. From Table 4.4, we observe that relative efficiency decreases for increases K and W.

Table 4.5: Conditional values for Population I and Population II

Conditions	Population1	Population2
i	0.0482418	0.3022614
ii	0.01579094	0.1212975
iii	0.02017135	0.1888713
iv	0.0146897	0.1004063
v	0.005715378	0.05040541
vi	0.004998455	0.05022207
vii	0.001848421	0.04915825
viii	0.000172548	0.002024653

All the values in Table 4.5 are greater than zero (positive) which shows that proposed estimator perform better than existing estimators.

Conclusion

We considered the problem of non response in the proposed estimator. Tables 4.1-4.2 present MSE comparison of the suggested estimator with considered estimators and observe that their MSE increases by increasing the values of K and W. Table 4.3-4.4 presents PRE comparison of the suggested estimator with considered estimators for different combinations of W, and K for specified sample size. In Table 4.3, we found that the relative efficiency of proposed estimator $\bar{Y}_p^{**}$ increases as K and W are increases. From Table 4.4, we observe that relative efficiency decreases for increases K and W. Overall proposed estimator is more efficient than the considered estimators.

From our empirical study we conclude that MSE is minimum and PRE (percentage relative efficiency) of our proposed estimator $\hat{\bar{Y}}_p^*$ is maximum among all the estimators consider here. We can observe in table 3.3-3.4 that estimators has unexpectedly declined when measurement error is taken into account. In Chapter 4 we have discussed the issue of nonresponse for estimation of population mean in the proposed estimator by using Hansen and Hurwitz (1946) sub-sampling technique. We have used the criteria of MSE and PRE of different estimators to check the performance of suggested estimator. It is shown in Table 4.1-4.2, the MSE of proposed estimator is lower and PRE is higher as compared to other considered estimators. It is noticed that in table 4.3-4.4 for data set 1, PREs for proposed estimator increases if K increases 2.0 to 3.5 at $W=$ .1, .2 and .3. For dataset 2, propose estimator decreases if K increases 2.0 to 3.5 at W=.1, .2 and .3 respectively. The performance of exponential estimator is poor in both dataset.

Recommendation

Scrambling can also be included in the problem of measurement error and n0n-resp0nse to deal with more sensitive information. Further it can also be extended by using different sampling designs.

References

Bahl, S. and Tuteja, R. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences*, 12(1):159–164.

Cochran, W. G. (1977). *Sampling Techniques: 3rd Ed.* Wiley, New York.

Hansen, M. H. and Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, 41(236):517–529.

Khare, B. and Rehman, H. U. (2013). Improved chain type estimators for population mean using two auxiliary variables and double sampling scheme. *Journal Rankings on Statistics and Probability - SclImago*, 1(3):82–87.

Khare, B. and Sinha, R. (2007). Estimation of the ratio of the two population means using multi auxiliary characters in the presence of non-response. *Statistical Techniques in Life Testing, Reliability, Sampling Theory and Quality Control*, 1(1):63–171.

Khare, B. and Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response. *National Academy Science Letters*, 16(3):111–114.

Khare, B. and Srivastava, S. (1995). Study of conventional and alternative two phase sampling ratio, product and regression estimators in presence of nonresponse. *Proceedings-National Academy Of Sciences India Section A*, 65(1):195–204.

Khare, B. and Srivastava, S. (1997). Transformed ratio type estimators for the population mean in the presence of nonresponse. *Communications in Statistics-Theory and Methods*, 26(7):1779–1791.

Kiregyera, B. (1980). A chain ratio-type estimator in finite population double sampling using two auxiliary variables. *Metrika*, 27(1):217–223.

Kiregyera, B. (1984). Regression-type estimators using two auxiliary variables and the model of double sampling from finite populations. *Metrika*, 31(1):215–226.

Koyuncu, N. and Kadilar, C. (2009). Family of estimators of population mean using two auxiliary variables in stratified random sampling. *Communications in Statistics-Theory and Methods*, 38(14):2398–2417.

Kumar, M. (2011). Some ratio type estimators under measurement errors. *World Applied Sciences Journal*, 14(2):272–276.

Murthy, M. N. et al. (1967). *Sampling Theory and Methods.* Calcutta-35: Statistical Publishing Society,204/1, Barrackpore, Trunk Road.

Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, 33(201):101–116.

Okafor, F. C. and Lee, H. (2000). Double sampling for ratio and regression estimation with sub-sampling the non-respondents. *Survey Methodology*, 26(2):183–188.

Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika*, 45(1/2):154–165.

Pradhan, B. (2005). A chain regression estimator in two phase sampling using multi-auxiliary information. *Bulletin of the Malaysian Mathematical Sciences Society*, 28(1):81–86.

Rao, P. (1986). Ratio estimation with subsampling the nonrespondents. *Survey Methodology*, 12(2):217–230.

Rao, P. S. and Mudholkar, G. S. (1967). Generalized multivariate estimator for the mean of finite populations. *Journal of the American Statistical Association*, 62(319):1009–1012.

Roy, D. C. (2003). A regression-type estimator in two-phase sampling using two auxiliary variables. *Pakistan Journal of Statistics-All Series-*, 19(3):281–290.

Sahoo, J., Sahoo, L., and Mohanty, S. (1993). A regression approach to estimation in two-phase sampling using two auxiliary variables. *Current Science*, 65(1):73–75.

Searls, D. T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *Journal of the American Statistical Association*, 59(308):1225–1226.

Sharma, P. and Singh, R. (2013). A generalized class of estimators for finite population variance in presence of measurement errors. *Journal of Modern Applied Statistical Methods*, 12(2):13–19.

Singh, H. P. and Espejo, M. R. (2007). Double sampling ratio-product estimator of a finite population mean in sample surveys. *Journal of Applied Statistics*, 34(1):71–85.

Singh, H. P. and Karpe, N. (2009). A class of estimators using auxiliary information for estimating finite population variance in presence of measurement errors. *Communications in Statistics-Theory and Methods*, 38(5):734–741.

Singh, H. P. and Kumar, S. (2010). Estimation of mean in presence of non-response using two phase sampling scheme. *Statistical Papers*, 51(3):559–582.

ıgh, M. and Singh, R. K. (2002). Role of regression estimator involving measurement errors. *Brazilian Journal of Probability and Statistics*, 16(1):39–46.

.vastava, S. K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *Journal of the American Statistical Association*, 66(334):404–407.

.vastava, S. K. (1981). generalized two-phase sampling estimator. *Journal-Indian Society of Agricultural Statistics*, 12(1):15–22.

.basum, R. and Khan, I. (2004). Double sampling for ratio estimation with non-response. *Journal of the Indian Society of Agricultural Statistics*, 58(3):300–306.

.basum, R. and Khan, I. (2006). Double sampling ratio estimator for the population mean in presence of non-response. *Assam Statistical Review*, 20(2):73–83.

shwakarma, G. K. and Singh, H. P. (2012). A general procedure for estimating the mean using double sampling for stratification and multi-auxiliary information. *Journal of Statistical Planning and Inference*, 142(5):1252–1261.