Revista Colombiana de Estadística - Theoretical Statistics
January 2026, volume 49, issue 1, pp. 65–76
http://doi.org/10.15446/rce.v49n1.119457

Determination of the Effect of Measurement Error using Factor Type Estimator in two Phase Sampling

Determinación del efecto del error de medición utilizando un estimador de tipo factor en muestreo bifásico

Amita Yadav[1,a], Narendra Singh Thakur[2,b], Sarla Pareek[1,c]

[1]Department of Mathematics, Statistics and Computing, Banasthali Vidyapith, Rajasthan, India

[2]Department of Statistics, Govt. Adarsh Girls College, Sheopur, India

Abstract

In this study, factor-type estimator is used for estimating the finite population mean under two-phase sampling scheme in the presence of measurement error. The Bias and Mean Squared Error (MSE) of factor-type estimator are derived in this paper and compared with the existing estimators in literature. Then, the conditions under which the suggested estimator is better than the existing estimators in terms of efficiency are provided here. Theoretical as well as real data comparison of factor-type estimator are done to get precious result.

***Keywords*:** Bias; Estimation; Factor-type estimator; Measurement error; MSE; Two-phase sampling.

Resumen

En este estudio, se utiliza un estimador factorial para estimar la media de una población finita bajo un esquema de muestreo bifásico en presencia de error de medición. El sesgo y el error cuadrático medio (EMM) del estimador factorial se derivan en este trabajo y se comparan con los estimadores existentes en la literatura. Posteriormente, se presentan las condiciones bajo las cuales el estimador sugerido es superior a los estimadores existentes en términos de eficiencia. Se realizan comparaciones teóricas y reales del estimador factorial para obtener resultados valiosos.

***Palabras clave*:** Error de medición; Estimación; Estimador tipo factorial; MSE; Muestreo bifásico; Sesgo.

[a]Ph.D. E-mail: amitayadav73@yahoo.in
[b]Ph.D. E-mail: nst_stats@yahoo.co.in
[c]Ph.D. E-mail: psarla13@gmail.com

1. Introduction

Two-phase sampling (also known as double sampling) is a statistical technique used when information is collected in two phases. This method is often employed to improve the efficiency and cost-effectiveness of data collection, particularly when the measurement of interest is expensive or time-consuming to obtain. Suppose you are conducting a survey to estimate the average household income in a city. Measuring household income accurately is costly because it requires detailed interviews and possibly verification from financial records.

1. **Phase I**: You conduct a quick survey on a large number of households asking for a rough estimate of their income brackets, which is less accurate but much cheaper to obtain.

2. **Phase II**: From this initial sample, you select a smaller subsample of households for detailed income interviews to get precise income data.

Two-phase sampling is widely used in various fields, estimating crop yields where initial surveys might gather quick, less precise data, followed by detailed measurements on a subsample. Initial broad surveys to identify high-risk groups followed by detailed health assessments. Preliminary surveys to identify customer preferences followed by detailed studies on a targeted subsample. In summary, two-phase sampling is a powerful technique that helps balance cost and accuracy in data collection efforts, making it a valuable tool in various research and survey applications. Two-phase sampling can be particularly useful in dealing with measurement error, which occurs when the data collected deviates from the true values due to inaccuracies in the measurement process. This approach helps to improve the quality and reliability of the data by allowing for correction of measurement errors in the second phase.

Two-phase sampling is an effective strategy to address measurement errors by leveraging a combination of large-scale, error-prone initial data and smaller-scale, accurate subsample data. By employing statistical methods like regression calibration, errors-in-variables models, multiple imputation, and Bayesian methods, researchers can significantly improve the accuracy and reliability of their estimates, making the technique valuable in various fields, including health studies, social sciences, and market research.

The Hansen and Hansen & Hurwitz (1946) method addressed the issue of non-response (by sub-sampling non-respondents) and constructing an unbiased estimator for the population mean. This approach enhances the accuracy of survey estimates by mitigating the bias introduced by non-response and effectively utilizing the available data. Wu & Luan (2003) highlighted that a significant advantage of two-phase sampling is its ability to achieve high precision in estimates without a substantial increase in cost. This methodology effectively balances cost and precision, making it an attractive option for survey practitioners. This paper by Singh & Vishwakarma (2007) focused on improving estimation techniques for population means in two-phase sampling by using auxiliary information. By introducing exponential (ratio and product) estimators, the paper seemed to aim at refining the accuracy of these estimates compared to existing methods like the simple mean per unit, and the usual ratio and product estimators in two-phase sampling. Khare & Sinha (2009); Sinha & Khare (2011) made substantial contribution in estimators' development for the population mean under non-response using multi-auxiliary variables within the framework of double sampling.

Measurement errors can manifest in two ways: measurement bias and measurement variance. Measurement bias is a systematic error, where respondents consistently provide incorrect answers (e.g., forgetting to report certain income sources), leading to a consistent under- or over-estimation of the true value. Measurement variance, on the other hand, reflects random error, where a respondent might give distinct responses to the similar questions upon being asked repeatedly. The total survey error under measurement errors can be thought of as having both types of errors, fixed errors i.e. bias and variable errors i.e. variance, which contribute to the overall inaccuracy in survey estimations. Measurement error is a prevalent issue in survey research that can significantly distort the accuracy of collected data and the validity of resulting estimates. Recognizing and addressing measurement error through improved instrument design, error modeling, data validation, and robust estimation methods is crucial for obtaining reliable and accurate survey results. Buonaccorsi (2010) provides valuable insights and methodologies for dealing with measurement errors, underscoring the importance of considering these errors in statistical analysis.

Yadav et al. (2024) addressed factor-type estimator under measurement errors. These errors could introduce bias and affected the accuracy of the survey results, making it essential to address them in the design and analysis of surveys. Yadav et al. (2025) offered an important extension to the classical ratio and product estimators by incorporating measurement errors and applying them in the context of post-stratification. Their work demonstrated that by using these extended estimators, researchers could obtain more accurate estimates of population means, even under data errors. The empirical validation further strengthened the utility of these methods for real-world survey sampling applications. This paper was especially relevant for situations where data was prone to measurement errors, as it presented a practical way to improve the accuracy of estimates without resorting to full data collection (census), which might not always be feasible or cost-effective.

Many researchers have studied measurement errors like (Manisha & Singh, 2001, 2002; Singh & Karpe, 2009, 2010; Manisha & Singh, 2002; Shukla et al., 2012*a*,*b*; Singh et al., 2020; Sabir et al., 2022; Singh & Vishwakarma, 2019).

2. Notations and Set Up

Let us consider a (finite) population having variables (Y_{dsi}, X_{dsi}) of size N with means $(\underline{\mathrm{Y}}, \underline{\mathrm{X}})$. And S_y^2 and S_x^2 be the population variance for y and x, respectively. Consider a preliminary large sample S' of size n' drawn from population of size N by SRSWOR and a secondary sample S of size n (where $n < n'$) drawn as a sub-sample from sample S' as in the given Figure 1. Assume for samples S and

S', the data x_{dsi} and x'_{dsi} are known with mean $\underline{\mathrm{x}}_{ds}$ and $\underline{\mathrm{x}}'_{ds}$, respectively.

Let sample having measurement errors while doing selection, recording and process data for further task. The i^{th} observed unit of variables y and x under measurement error in the sample is (y_{dsi}, x_{dsi}), respectively. Let (u_{dsi}, v_{dsi}) be the values of measurement errors corresponding to the (y_{dsi}, x_{dsi}) such that $(y_{dsi} = Y_{dsi} + u_{dsi},\ x_{dsi} = X_{dsi} + v_{dsi})$ where Y_{dsi} and X_{dsi} be the actual values of y_{dsi} and x_{dsi}, respectively.

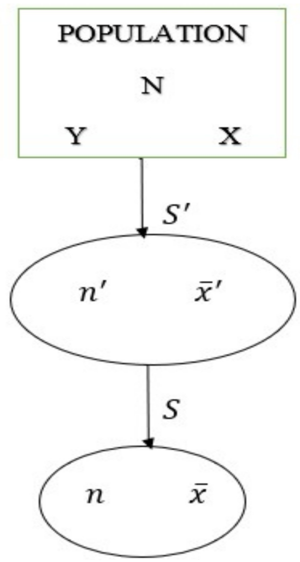

Figure 1:

Since observations taken for any sample are independent of each other. Therefore, measurement error (u_{dsi}, v_{dsi}) are also independent of each other and average values of measurement errors are zero (let). The population variances for (u_{dsi}, v_{dsi}) are S_u^2 and S_v^2, respectively.

Let

$$\underline{\mathrm{Y}} = \frac{1}{N}\sum_{i=1}^{N} Y_{dsi},$$

$$\underline{\mathrm{X}} = \frac{1}{N}\sum_{i=1}^{N} X_{dsi},$$

$$\underline{y}_{ds} = \frac{1}{n}\sum_{i=1}^{n} y_{dsi},$$

$$\underline{x}_{ds} = \frac{1}{n}\sum_{i=1}^{n} x_{dsi},$$

$$\underline{x}'_{ds} = \frac{1}{n'}\sum_{i=1}^{n'} x'_{dsi},$$

Here are some notations:

$$\begin{aligned}
\underline{y}_{ds} &= \underline{Y}(1+\in_0),\\
\underline{x}_{ds} &= \underline{X}(1+\in_1),\\
\underline{x}'_{ds} &= \underline{X}(1+\in_2)\\
E(\in_0) &= 0,\\
E(\in_1) &= 0,\\
E(\in_2) &= 0,\\
E(\in_0^2) &= \lambda_1(S_y^2+S_u^2)\underline{Y}^{-2},\\
E(\in_1^2) &= \lambda_1(S_x^2+S_v^2)\underline{X}^{-2},\\
E(\in_2^2) &= \lambda_2(S_x^2+S_v^2)\underline{X}^{-2}\\
E(\in_0\in_1) &= \lambda_1\rho S_x S_y \underline{X}^{-1}\underline{Y}^{-1}\\
E(\in_1\in_2) &= \lambda_2(S_x^2+S_v^2)\underline{X}^{-2}\\
E(\in_0\in_2) &= \lambda_2\rho S_x S_y \underline{X}^{-1}\underline{Y}^{-1}
\end{aligned}$$

where

$$\lambda_1 = \frac{1}{n} - \frac{1}{N}, \quad \lambda_2 = \frac{1}{n'} - \frac{1}{N},$$

and

$$\lambda = \lambda_1 - \lambda_2 = \frac{1}{n} - \frac{1}{n'}$$

$$\begin{aligned}
\theta_{1ds} &= \frac{fB_d}{A_d+fB_d+C_d},\\
\theta_{2ds} &= \frac{C_d}{A_d+fB_d+C_d},\\
\theta_{ds} = \theta_{1ds}-\theta_{2ds} &= \frac{fB_d-C_d}{A_d}+fB_d+C_d
\end{aligned}$$

and $\phi = \frac{a\underline{X}}{2(a\underline{X}+b)}$ (let).

3. Existing Estimator

Some existing estimators of (population mean) $\underline{Y}$ under two-phase sampling are discussed here in this section.

1. Singh & Vishwakarma (2007) modified the exponential ratio type estimator for population mean under two phase sampling and its expression as

$$\underline{y}_1 = \underline{y}_{ds} \exp\left(\frac{\underline{x}'_{ds} - \underline{x}_{ds}}{\underline{x}'_{ds} + \underline{x}_{ds}}\right). \tag{1}$$

Here, the value of mean squared error i.e. MSE of estimator $\underline{y}_1$ is expressed as-

$$MSE\left(\underline{y}_1\right) = \lambda_1 \left(S_y^2 + S_u^2\right) + 0.25\lambda R^2 \left(S_x^2 + S_v^2\right) - \lambda\rho R S_x S_y. \tag{2}$$

2. Ozgul & Cingi (2014) modified the class of exponential regression cum ratio estimator under two phase sampling and its expression is given as

$$\underline{y}_2 = \left(k_1 \underline{y}_{ds} + k_2\left(\underline{x}'_{ds} - \underline{x}_{ds}\right)\right) \exp\left(\frac{\underline{z}'_{ds} - \underline{z}_{ds}}{\underline{z}'_{ds} + \underline{z}_{ds}}\right), \tag{3}$$

where k_1 and k_2 are constants and $\underline{z}'_{ds} = a\underline{x}'_{ds} + b$ and $\underline{z}_{ds} = a\underline{x}_{ds} + b$.
Here, the value of minimum mean squared error of the estimator $\underline{y}_2$ is

$$\min MSE\left(\underline{y}_2\right) = R^4 \frac{\left(\lambda_1 - \lambda\rho^2\right)\left(S_y^2 + S_u^2\right)\left(\underline{X}^2 - \lambda\phi^2\left(S_x^2 + S_v^2\right)\right) - 0.25\lambda_1^2\phi^4\left(S_x^2 + S_v^2\right)^2}{\underline{Y}^2 + \left(\lambda_1 - \lambda\rho^2\right)\left(S_y^2 + S_u^2\right)R^2}. \tag{4}$$

3. Muhammad et al. (2021) modified an alternate ratio-regression-type in two phase sampling and its expression is given as

$$\underline{y}_3 = \left(\alpha_1 \underline{y}_{ds}\left(\frac{\underline{x}'_{ds}}{\underline{x}_{ds}}\right) + \alpha_2\left(\underline{x}'_{ds} - \underline{x}_{ds}\right)\right) \exp\left(\frac{\underline{z}'_{ds} - \underline{z}_{ds}}{\underline{z}'_{ds} + \underline{z}_{ds}}\right), \tag{5}$$

where α_1 and α_2 are real parameters and $\underline{z}'_{ds} = a\underline{x}'_{ds} + b$ and $\underline{z}_{ds} = a\underline{x}_{ds} + b$.
The value of minimum mean squared error of estimator $\underline{y}_3$ is

$$\min MSE\left(\underline{y}_3\right) = \underline{Y}^2 - \frac{W_2^2 W_3 + W_4^2 W_1 - 2W_2 W_4 W_5}{4(W_1 W_3 - W_5^2)}, \tag{6}$$

where

$$W_1 = \underline{Y}^2 + \lambda\left(S_y^2 + S_u^2\right) + \lambda R^2\left(S_x^2 + S_v^2\right)\left(\phi^2 + 2\phi + 2\right) - 2\lambda\rho R S_x S_y\left(\phi + 1\right)$$

$$W_2 = 2\underline{Y}^2 + 3\lambda R^2\left(S_x^2 + S_v^2\right)\left(1 + 0.25\phi^2\right) - \lambda\rho R S_x S_y\left(\phi + 2\right)$$

$$W_3 = \lambda\left(S_x^2 + S_v^2\right)$$

$$W_4 = \lambda\phi R\left(S_x^2 + S_v^2\right)$$

$$W_5 = \lambda\left(R\left(S_x^2 + S_v^2\right)\left(\phi + 1\right) + \rho R S_x S_y\right).$$

4. Proposed Estimator and its Properties

Under two-phase sampling scheme, the factor-type estimator (given by Singh & Shukla, 1987) under measurement error setup is proposed and its expression is as:

$$\underline{y}_{FT} = \underline{y}_{ds}\left[\frac{(A_d + C_d)\,\underline{x}'_{ds} + fB_d\underline{x}_{ds}}{(A_d + fB_d)\,\underline{x}'_{ds} + C_d\underline{x}_{ds}}\right], \tag{7}$$

where, $A_d = (k_d-1)(k_d-2)$, $B_d = (k_d-1)(k_d-4)$, $C_d = (k_d-2)(k_d-3)(k_d-4)$, $f_p = \frac{n}{N}$ and $k_d =$ constant such that $0 < k_d < \infty$.

Putting the values of $\underline{y}_{ds}$, $\underline{x}'_{ds}$ and $\underline{x}_{ds}$ from notations and equation (7) will be

$$\underline{y}_{FT} = \underline{Y}(1+\in_0)\left[\frac{(A_d + C_d)\,\underline{X}\,(1+\in_2) + \ fB_d\underline{X}(1+\in_1)}{(A_d + fB_d)\,\underline{X}\ \ (1+\in_2) + C_d\underline{X}(1+\in_1)}\right]$$

$$\underline{y}_{FT} = \underline{Y}\ \ (1+\in_0)\,(1+(1-\theta_{1ds})\in_2 + \theta_{1ds}\in_1)(1+(1-\theta_{2ds})\in_2 + \theta_{2ds}\in_1)^{-1}$$

$$\underline{y}_{FT} - \underline{Y} = \underline{Y}\big(\in_0 + \theta_{1ds}\,(\in_1 - \in_2) + \theta_{1ds}(\in_0\in_1 - \in_0\,\in_2) - \theta_{1ds}\,(\in_1\in_2 - \in_2^2) + \theta_{2ds}^2(\in_1^2 + \in_2^2 - 2\in_1\in_2)\big). \tag{8}$$

Taking expectations of Equation (8)

$$E\left(\underline{y}_{FT} - \underline{Y}\right) = E\big(\underline{Y}\ (\in_0 + \theta_{ds}\,(\in_1 - \in_2) + \theta_{ds}(\in_0\in_1 - \in_0\,\in_2) - \theta_{ds}\,(\in_1\in_2 - \in_2^2) + \theta_{2ds}^2(\in_1^2 + \in_2^2 - 2\in_1\in_2))\big)$$

$$Bias\,(\underline{y}_{FT}) = E\left(\underline{y}_{FT} - \underline{Y}\right) = \underline{Y}\left(\theta_{ds}\,(\lambda_1 - \lambda_2)\,\rho S_x S_y \underline{X}^{-1}\underline{Y}^{-1} + \theta_{ds}^2\,(\lambda_1 - \lambda_2)\,(S_x^2 + S_v^2)\ \underline{X}^{-2}\right). \tag{9}$$

Squaring both sides of Equation (8),

$$\left(\underline{y}_{FT} - \underline{Y}\right)^2 = \big(\underline{Y}\big(\in_0 + \theta_{ds}\,(\in_1 - \in_2) + \theta_{ds}(\in_0\in_1 - \in_0\,\in_2) - \theta_{ds}\,(\in_1\in_2 - \in_2^2) + \theta_{ds}^2(\in_1^2 + \in_2^2 - 2\in_1\in_2)\big)\big)^2$$

$$\left(\underline{y}_{FT} - \underline{Y}\right)^2 = \underline{Y}^2\left(\in_0^2 + \theta_{ds}^2\left(\in_1^2 + \in_2^2 - 2\in_1\in_2\right) + 2\theta_{ds}(\in_0\in_1 - \in_0\,\in_2)\right). \tag{10}$$

Taking expectations of Equation (10),

$$MSE(\underline{y}_{FT}) = \underline{Y}^2\big(\lambda_1\left(S_y^2 + S_u^2\right)\underline{Y}^{-2} + \theta_{ds}^2\,(\lambda_1 - \lambda_2)\,(S_x^2 + S_v^2)\underline{X}^{-2} + 2\theta_{ds}\,(\lambda_1 - \lambda_2)\,\rho S_x S_y \underline{X}^{-1}\underline{Y}^{-1}\big)$$

$$MSE(\underline{y}_{FT}) = \lambda_1 \left(S_y^2 + S_u^2\right) + \theta_{ds}^2 \lambda (S_x^2 + S_v^2) R^2 + 2\theta_{ds} \lambda R \rho S_x S_y. \tag{11}$$

To get minimum MSE at optimum value of θ_{ds}, we differentiate Equation (11) w.r.t. θ_{ds}

$$\frac{\partial\,(MSE)}{\partial \theta_{ds}} = 2\theta_{ds} \lambda (S_x^2 + S_v^2) R^2 + 2\lambda R \rho S_x S_y.$$

On putting $\frac{\partial (MSE)}{\partial \theta_{ds}} = 0$, we get

$$\theta_{ds} = -\frac{\rho S_x S_y}{(S_x^2 + S_v^2) R} \tag{12}$$

$$\min MSE\left(\underline{y}_{FT}\right) = \lambda_1 \left(S_y^2 + S_u^2\right) - \frac{\lambda \rho^2 S_x^2 S_y^2}{(S_x^2 + S_v^2)}. \tag{13}$$

5. Efficiency Comparison with the Proposed Estimator

Here, we exemplify that the suggested factor type estimator will have minimum MSE as compared to all other existing estimators in the literature for estimating population mean. In addition to this, our study also explains the stages under which the suggested estimator is better than other estimators taken here.

1. Comparison of factor type estimator with usual ratio estimator, we have:

$$\min MSE\left(\underline{y}_{FT}\right) < MSE\left(\underline{y}\right)$$

$$\min MSE\left(\underline{y}_{FT}\right) - MSE\left(\underline{y}\right) < 0$$

$$\lambda_1 \left(S_y^2 + S_u^2\right) - \frac{\lambda \rho^2 S_x^2 S_y^2}{(S_x^2 + S_v^2)} - \left(\lambda_1 \left(S_y^2 + S_u^2\right) + \lambda R^2 \left(S_x^2 + S_v^2\right) - 2\lambda \rho R S_x S_y\right) < 0$$

$$\frac{\lambda \rho^2 S_x^2 S_y^2}{(S_x^2 + S_v^2)} + \lambda R^2 \left(S_x^2 + S_v^2\right) - 2\lambda \rho R S_x S_y > 0$$

$$\left(\rho S_x S_y - R\left(S_x^2 + S_v^2\right)\right)^2 > 0.$$

Since above condition is true. So, factor type estimator is better in terms of efficiency than usual ratio estimator.

2. Comparison of factor type estimator with Singh & Vishwakarma (2007) exponential-ratio type estimator, we have:

$$\min MSE\left(\underline{y}_{FT}\right) < MSE\left(\underline{y}_1\right)$$

$$\min MSE\left(\underline{y}_{FT}\right) - MSE\left(\underline{y}_1\right) < 0$$

$$\lambda_1\left(S_y^2+S_u^2\right)-\frac{\lambda\rho^2S_x^2S_y^2}{\left(S_x^2+S_v^2\right)} \\ -\left(\lambda_1\left(S_y^2+S_u^2\right)+0.25\lambda R^2\left(S_x^2+S_v^2\right)+\lambda\rho RS_xS_y\right)<0$$

$$\frac{\lambda\rho^2S_x^2S_y^2}{\left(S_x^2+S_v^2\right)}+0.25\lambda R^2\left(S_x^2+S_v^2\right)+\lambda\rho RS_xS_y>0$$

$$\left(\rho S_xS_y+0.5R\left(S_x^2+S_v^2\right)\right)^2>0.$$

Since above condition is true. So, factor type estimator is better in terms of efficiency of exponential-ratio type estimator.

3. Comparison of factor type estimator with Ozgul & Cingi (2014) the class of exponential regression cum ratio estimator, we have:

$$\min MSE\left(\underline{y}_{FT}\right)<\min MSE\left(\underline{y}_2\right)$$

$$\min MSE\left(\underline{y}_{FT}\right)-\min MSE\left(\underline{y}_2\right)<0$$

$$\lambda_1\left(S_y^2+S_u^2\right)-\frac{\lambda\rho^2S_x^2S_y^2}{\left(S_x^2+S_v^2\right)} \\ -R^4\frac{\left(\lambda_1-\lambda\rho^2\right)\left(S_y^2+S_u^2\right)\left(\underline{X}^2-\lambda\phi^2\left(S_x^2+S_v^2\right)\right)-0.25\lambda_1^2\phi^4\left(S_x^2+S_v^2\right)^2}{\underline{Y}^2+\left(\lambda_1-\lambda\rho^2\right)\left(S_y^2+S_u^2\right)R^2}<0.$$

Since above condition is true. So, factor type estimator is better in terms of efficiency of class of exponential regression cum ratio estimator.

4. Comparison of factor type estimator with Muhammad et al. (2021) alternative ratio-regression-type estimator, we have:

$$\min MSE\left(\underline{y}_{FT}\right)<\min MSE\left(\underline{y}_3\right)$$

$$\min MSE\left(\underline{y}_{FT}\right)-\min MSE\left(\underline{y}_3\right)<0$$

$$\lambda_1\left(S_y^2+S_u^2\right)-\frac{\lambda\rho^2S_x^2S_y^2}{\left(S_x^2+S_v^2\right)}-\left(\underline{Y}^2-\frac{W_2^2W_3+W_4^2W_1-2W_2W_4W_5}{4\left(W_1W_3-W_5^2\right)}\right)<0.$$

Since above condition is true. So, factor type estimator is better in terms of efficiency of Muhammad et al. (2021) alternative ratio-regression-type estimator.

6. Empirical Study

To justify the performance of factor-type estimator over other estimators in terms of efficiency of factor-type estimator as compared to other previously existing estimators under two-phase sampling schemes under measurement error by using different real datasets.

1. Data for first population is taken from Ozgul & Cingi (2014). In this data set, x and y be the number of students (strength of both primary and secondary schools) for 923 districts and the number of teachers, respectively.
2. Data for second population is taken from Kadilar & Cingi (2006). In this data set, x and y be the number of apple trees and level of apple, production respectively.

Table 1: Different parameters of the populations.

Parameters	N	n'	n	ρ	$\underline{Y}$	$\underline{X}$	C_y	C_x
Population I	923	400	200	0.955	436.3	11440.50	1.72	1.86
Population II	104	40	20	0.865	625.37	13.930	1.866	1.653

Now, the bias and MSE of factor-type estimator were derived. After that MSE's of different estimators were compared as shown in the Table 2. Numerical examples by taking different datasets were carried out. The mean squared error (MSE) and its percentage relative efficiency of all estimators were shown in Table 2.

Table 2: MSE's and efficiency of different estimators for population I and II.

Population	I		II	
	MSE	Efficiency	MSE	Efficiency
$\underline{y}$	1131322.411	100.000	2834.137	100.000
$\underline{y}_1$	4071.206	27788.382	87759.002	3.229
$\underline{y}_2$	7915.871	14291.825	382757.291	0.740
$\underline{y}_3$	1943.633	58206.600	389893.830	0.727
$\underline{y}_{FT}$	921.610	**122755.053**	2521.364	**112.405**

Based on the above results from first dataset of population I, it is found that the factor-type estimator has minimum MSE and maximum efficiency values (921.610,122755.053) taken here. For dataset of population II, it is found that the factor-type estimator has minimum MSE and maximum efficiency values (2521.364,112.405) as compared to other existing estimators taken here.

7. Results and Conclusion

This paper proposes a new factor type estimator for estimation of population mean under measurement error in two-phase sampling. We derive expressions for bias, mean squared error (MSE), and optimum MSE (up to the first degree of approximation). Theoretically, we demonstrate that the suggested estimator is better in terms of efficiency than existing estimators in two-phase sampling. Numerical results, based on real datasets, show that the suggested estimator attains minimum MSE values. The evidence from this study reveals that the suggested estimator performs better in terms of efficiency of existing estimators in terms of bias, MSE, and their relative efficiencies.

[Received: February 2024 — Accepted: November 2025]

References

Buonaccorsi, J. P. (2010), *Measurement error: Models, methods and applications*, 1 edn, Chapman and Hall/CRC, New York.

Hansen, M. H. & Hurwitz, W. N. (1946), 'The problem of nonresponse in sample surveys', *Journal of the American Statistical Association* **41**, 517–529.

Kadilar, C. & Cingi, H. (2006), 'Ratio estimators for the population variance in simple and stratified random sampling', *Applied Mathematics and Computation* **173**, 1047–1059.

Khare, B. B. & Sinha, R. R. (2009), 'On class of estimators for population mean using multi-auxiliary characters in the presence of non-response', *Statistics in Transition: New Series* **10**, 3–14.

Manisha & Singh, R. K. (2001), 'An estimation of population mean in the presence of measurement errors', *Journal of the Indian Society of Agricultural Statistics* **54**, 13–18.

Manisha & Singh, R. K. (2002), 'Role of regression estimator involving measurement errors', *Brazilian Journal of Probability and Statistics* **16**, 39–46.

Muhammad, I., Zakari, Y. & Audu, A. (2021), 'An alternative class of ratio-regression-type estimator under two-phase sampling scheme', *CBN Journal of Applied Statistics* **12**, 1–26.

Ozgul, N. & Cingi, H. (2014), 'A new class of exponential regression cum ratio estimator in two-phase sampling', *Hacettepe Journal of Mathematics and Statistics* **43**, 131–140.

Sabir, S., Sanaullah, A. & Gupta, S. (2022), 'Estimation of mean considering the joint influence of measurement errors and non-response in two-phase sampling designs', *Journal of the National Science Foundation of Sri Lanka* **50**, 13.

Shukla, D., Pathak, S. & Thakur, N. S. (2012*a*), 'Classes of factor-type estimator in presence of measurement error', *Journal of Modern Applied Statistical Methods* **11**, 336–347.

Shukla, D., Pathak, S. & Thakur, N. S. (2012*b*), 'An estimator for mean estimation in presence of measurement error', *Research and Reviews: A Journal of Statistics* **2**, 1–8.

Singh, H. P. & Karpe, N. (2009), 'On the estimation of ratio and product of two population means using supplementary information in presence of measurement errors', *Statistica* **69**, 27–47.

Singh, H. P. & Karpe, N. (2010), 'Effect of measurement errors on the separate and combined ratio and product estimators in stratified random sampling', *Journal of Modern Applied Statistical Methods* **9**.

Singh, H. P. & Vishwakarma, G. K. (2007), 'Modified exponential ratio and product estimators for finite population mean in two-phase sampling', *Austrian Journal of Statistics* **36**, 217–225.

Singh, N. & Vishwakarma, G. K. (2019), 'A generalized class of estimator of population mean with the combined effect of measurement errors and non-response in sample survey', *Revista Investigación Operacional* **40**, 275–285.

Singh, R., Mishra, M. & Mishra, M. (2020), 'Estimation in stratified random sampling in the presence of errors', *Revista Investigación Operacional* **41**, 125–137.

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

Sinha, R. R. & Khare, B. B. (2011), 'Estimation of population mean using multi-auxiliary characters with subsampling the non-respondents', *Statistics in Transition: New Series* **12**, 45–56.

Wu, C. & Luan, Y. (2003), 'Optimal calibration estimators under two-phase sampling', *Journal of Official Statistics* **19**, 119–131.

Yadav, A., Thakur, N. S. & Pareek, S. (2024), 'Determining the effect of measurement errors and estimation of mean by factor-type estimator in simple random sampling', *Journal of Modern Applied Statistical Methods* **23**, 6.

Yadav, A., Thakur, N. S. & Pareek, S. (2025), 'Ratio and product type estimators in post-stratified sampling with measurement errors', *Model Assisted Statistics and Applications* .