

# A New Regression Type Estimator and Its Application in Survey Sampling

M. Zahid Hasan<sup>1\*</sup> , M. Sultana<sup>2</sup> , K. Fatema<sup>1</sup> , Md. Ali Hossain<sup>1</sup>, M. Murad Hossain<sup>1</sup>

<sup>1</sup>Department of Statistics, Faculty of Science, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh

<sup>2</sup>Department of Mathematics, Faculty of Science, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh

Email: \*zahid@bsmrstu.edu.bd

**How to cite this paper:** Hasan, M.Z., Sultana, M., Fatema, K., Hossain, Md.A. and Hossain, M.M. (2020) A New Regression Type Estimator and Its Application in Survey Sampling. *Open Journal of Statistics*, 10, 1010-1019.

<https://doi.org/10.4236/ojs.2020.106057>

**Received:** December 4, 2019

**Accepted:** December 12, 2019

**Published:** November 24, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In the present time, a large number of modified estimators have been proposed by authors to obtain efficiency. In this study, we suggested an alternative regression type estimator for estimating finite population means when there is either a positive or negative correlation between study variables and auxiliary variables. We obtained bias and mean square error equation of the proposed estimator ignoring the first-order approximation and found the theoretical conditions that make proposed estimator more efficient than simple random sampling mean estimator, product estimator and ratio estimator. In addition, these conditions are supported by a numerical example and it has been concluded that the proposed estimator performed better comparing with the usual simple random sampling mean estimator, ratio estimator and product estimator.

## Keywords

Auxiliary Information, Bias, Efficiency, Mean Square Error, Product and Ratio Estimator

---

## 1. Introduction

The use of auxiliary information has increased widely in survey sampling to improve the precision of estimates. A lot of works have been done by many authors using auxiliary information. The usage of the transformed auxiliary variable in estimating finite population mean is presented in [1]. The use of two auxiliary variables in estimation has been introduced by [2]. A modified ratio estimator for the estimation of finite population mean when median of the auxiliary varia-

ble was available is shown in [3]. The modified estimator shows the significant role using of auxiliary information. The cum-dual ratio estimator was more efficient than other existing estimator proved by [4]. A new modified product estimator is better than other existing estimator when median of the auxiliary variable is known [5].

Regression estimator is important in estimation theory for its improved precision. It has been widely used in estimating population parameter. Regression estimator is more useful in many practical situations including business, economics, time series forecasting, agriculture etc. [6]. Before applying regression estimator, the regression parameter needs to be estimated. In this study, we proposed new regression type estimator where correlation coefficient is used instead of regression parameter. In many practical situations, population correlation coefficients are known [7].

The structure of the paper is as follows: Section 2 contains the concept of proposed estimator and its bias and mean square error. Section 3 shows the efficiency comparison of proposed estimator compared to that of ratio, product and simple random sampling mean estimator. Section 4 illustrates simulation study of the proposed estimator. Section 5 presents the real-life example of proposed estimator. Finally, section 6 provides the conclusion of our study.

## 2. Proposed Estimator

In this section we proposed our new estimator. For this purpose, we derived bias and mean square error of proposed estimator.

Consider a finite population  $U = (u_1, u_2, \dots, u_N)$  of size  $N$  units. Let  $y$  and  $x$  denote the study variable and auxiliary variable respectively. For estimating population mean  $\bar{Y}$ , a sample of size  $n (n < N)$  is drawn using simple random sampling without replacement (SRSWOR) from the population  $U$ .

Where, the population mean is  $\bar{Y} = \frac{1}{N} \sum y_i$ .

Hansen, Hurwitz and Madow (1953) consider the linear regression estimator of the population mean  $\bar{Y}$  as

$$\bar{y}_{LR} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x}).$$

where,  $\hat{\beta} = \frac{S_{xy}}{S_x^2}$  an estimate of regression coefficient is  $\beta = \frac{S_{xy}}{S_x^2}$ .

The new regression type estimator for estimating population mean  $\bar{Y}$  as

$$\bar{y}_z = \bar{y} + r(\bar{X} - \bar{x}) \quad (1)$$

where,  $r = \frac{S_{xy}}{S_x S_y}$  is an estimate of correlation coefficient  $\rho = \frac{S_{xy}}{S_x S_y}$ .

To calculate bias and mean square error of new regression type estimator, we consider first degree approximation.

Let us define [6]

$$e_0 = \left(\frac{\bar{y}}{\bar{Y}} - 1\right), e_1 = \left(\frac{\bar{x}}{\bar{X}} - 1\right), e_2 = \left(\frac{s_y^2}{S_y^2} - 1\right), e_3 = \left(\frac{s_x^2}{S_x^2} - 1\right), e_4 = \left(\frac{s_{xy}}{S_{xy}} - 1\right)$$

This implies that,

$$\bar{y} = \bar{Y}(1 + e_0), \bar{x} = \bar{X}(1 + e_1), s_y = S_y(1 + e_2)^{\frac{1}{2}},$$

$$s_x = S_x(1 + e_3)^{\frac{1}{2}}, s_{xy} = (1 + e_4).$$

$$E(e_0) = E(e_1) = E(e_2) = E(e_4) = 0.$$

$$E(e_0^2) = \left(\frac{1-f}{n}\right) C_y^2, E(e_1^2) = \left(\frac{1-f}{n}\right) C_x^2, E(e_0 e_1) = \left(\frac{1-f}{n}\right) C_x C_y \rho_{xy}.$$

$$E(e_1 e_2) = \left(\frac{1-f}{n}\right) C_x \lambda_{21}, E(e_1 e_3) = \left(\frac{1-f}{n}\right) C_x \lambda_{03},$$

$$E(e_1 e_4) = \left(\frac{1-f}{n}\right) C_x \frac{\lambda_{12}}{\rho_{xy}}.$$

$$f = \frac{n}{N}, C_y^2 = \frac{S_y^2}{\bar{Y}^2}, C_x^2 = \frac{S_x^2}{\bar{X}^2}, \rho = \frac{S_{xy}}{S_x S_y}.$$

$$S_x^2 = \frac{1}{N-1} \sum (x_i - \bar{X})^2, S_y^2 = \frac{1}{N-1} \sum (y_i - \bar{Y})^2,$$

$$S_{xy} = \frac{1}{N-1} \sum (y_i - \bar{Y})(x_i - \bar{X}).$$

**Theorem 1.** The bias of the proposed estimator  $\bar{y}_z$  of population mean  $\bar{Y}$  is given by

$$Bias(\bar{y}_z) = \frac{1}{2} \left(\frac{1-f}{n}\right) S_x (\rho \lambda_{21} + \rho \lambda_{03} - 2 \lambda_{12}).$$

**Proof:**

From Equation (1) we can write,

$$\bar{y}_z = \bar{y} + r(\bar{X} - \bar{x})$$

$$\Rightarrow \bar{y}_z = \bar{Y}(1 + e_0) + r\{\bar{X} - \bar{X}(1 + e_1)\}$$

$$\Rightarrow \bar{y}_z = \bar{Y} + \bar{Y}e_0 + r(\bar{X} - \bar{X} - \bar{X}e_1)$$

$$\Rightarrow \bar{y}_z - \bar{Y} = \bar{Y}e_0 - r\bar{X}e_1$$

$$\Rightarrow \bar{y}_z - \bar{Y} = \bar{Y}e_0 - \frac{S_{xy}}{S_x S_y} \bar{X}e_1$$

The proposed estimator  $\bar{y}_z$ , in terms of  $e_0, e_1, e_2, e_3$  and  $e_4$  can be easily written as

$$\Rightarrow \bar{y}_z - \bar{Y} = \bar{Y}e_0 - \frac{S_{xy}(1 + e_4)}{S_x(1 + e_3)^{\frac{1}{2}} S_y(1 + e_2)^{\frac{1}{2}}} \bar{X}e_1$$

$$\Rightarrow \bar{y}_z - \bar{Y} = \bar{Y}e_0 - \frac{S_{xy}}{S_x S_y} \bar{X}e_1 (1 + e_4)(1 + e_3)^{-\frac{1}{2}}(1 + e_2)^{-\frac{1}{2}} \tag{2}$$

Using binomial expansion, we consider first order approximation and ignoring the higher order approximation. Hence, Equation (2) can be written as

$$\begin{aligned} \Rightarrow \bar{y}_z - \bar{Y} &= \bar{Y}e_0 - \rho\bar{X}e_1(1+e_4)\left(1 - \frac{e_2}{2} + \frac{3}{8}e_2^2 - \dots\right)\left(1 - \frac{e_3}{2} + \frac{3}{8}e_3^2 - \dots\right) \\ \Rightarrow \bar{y}_z - \bar{Y} &= \bar{Y}e_0 - \rho\bar{X}(e_1 + e_1e_4)\left(1 - \frac{e_2}{2} + \frac{3}{8}e_2^2 - \dots\right)\left(1 - \frac{e_3}{2} + \frac{3}{8}e_3^2 - \dots\right) \\ \Rightarrow \bar{y}_z - \bar{Y} &= \bar{Y}e_0 - \rho\bar{X}\left(e_1 - \frac{e_1e_2}{2} - \frac{e_1e_3}{2} + e_1e_4\right) \end{aligned} \quad (3)$$

After simplification, Equation (3) can be expressed as

$$\Rightarrow \bar{y}_z - \bar{Y} = \bar{Y}e_0 - \rho\bar{X}e_1 + \rho\bar{X}\frac{e_1e_2}{2} + \rho\bar{X}\frac{e_1e_3}{2} - \rho\bar{X}e_1e_4 \quad (4)$$

Now, taking expectation of both sides of Equation (4) we get,

$$\begin{aligned} \Rightarrow E(\bar{y}_z - \bar{Y}) &= \bar{Y}E(e_0) - \rho\bar{X}E(e_1) + \frac{1}{2}\rho\bar{X}E(e_1e_2) \\ &\quad + \frac{1}{2}\rho\bar{X}E(e_1e_3) - \rho\bar{X}E(e_1e_4) \\ \Rightarrow E(\bar{y}_z - \bar{Y}) &= 0 - 0 + \frac{1}{2}\rho\bar{X}\left(\frac{1-f}{n}\right)C_x\lambda_{21} + \frac{1}{2}\rho\bar{X}\left(\frac{1-f}{n}\right)C_x\lambda_{03} \\ &\quad - \rho\bar{X}\left(\frac{1-f}{n}\right)C_x\frac{\lambda_{12}}{\rho} \\ \Rightarrow E(\bar{y}_z - \bar{Y}) &= \left(\frac{1-f}{n}\right)\left(\frac{1}{2}\rho\bar{X}C_x\lambda_{21} + \frac{1}{2}\rho\bar{X}C_x\lambda_{03} - \bar{X}C_x\lambda_{12}\right) \\ \Rightarrow Bias(\bar{y}_z) &= \left(\frac{1-f}{n}\right)\bar{X}C_x\left(\frac{1}{2}\rho\lambda_{21} + \frac{1}{2}\rho\lambda_{03} - \lambda_{12}\right). \end{aligned}$$

Simplifying this, we get the bias of the estimator  $\bar{y}_z$ . Hence this is the theorem.

**Theorem 2.** The mean squared error of the proposed estimator  $\bar{y}_z$ , to the first order approximation is

$$MSE(\bar{y}_z) = \left(\frac{1-f}{n}\right)\left(\bar{Y}^2C_y^2 + \rho^2\bar{X}^2C_x^2 - 2\rho^2\bar{X}\bar{Y}C_xC_y\right).$$

**Proof:**

Squaring both sides of Equation (3) and then taking expectation

$$\begin{aligned} \Rightarrow (\bar{y}_z - \bar{Y})^2 &= \left\{ \bar{Y}e_0 - \rho\bar{X}\left(e_1 - \frac{e_1e_2}{2} - \frac{e_1e_3}{2} + e_1e_4\right) \right\}^2 \\ \Rightarrow (\bar{y}_z - \bar{Y})^2 &= (\bar{Y}e_0)^2 + \rho^2\bar{X}^2\left(e_1 - \frac{e_1e_2}{2} - \frac{e_1e_3}{2} + e_1e_4\right)^2 \\ &\quad - 2\bar{Y}e_0\rho\bar{X}\left(e_1 - \frac{e_1e_2}{2} - \frac{e_1e_3}{2} + e_1e_4\right) \\ \Rightarrow (\bar{y}_z - \bar{Y})^2 &= \bar{Y}^2e_0^2 + \rho^2\bar{X}^2e_1^2 - 2\bar{Y}e_0\rho\bar{X}e_1 \end{aligned}$$

$$\Rightarrow E(\bar{y}_z - \bar{Y})^2 = \bar{Y}^2 E(e_0^2) + \rho^2 \bar{X}^2 E(e_1^2) - 2\rho \bar{Y}\bar{X} E(e_0 e_1)$$

By the definition of mean squared error (MSE), we have

$$\begin{aligned} \Rightarrow MSE(\bar{y}_z) &= E(\bar{y}_z - \bar{Y})^2 = \bar{Y}^2 E(e_0^2) + \rho^2 \bar{X}^2 E(e_1^2) - 2\rho \bar{Y}\bar{X} E(e_0 e_1) \\ \Rightarrow MSE(\bar{y}_z) &= \bar{Y}^2 \left(\frac{1-f}{n}\right) C_y^2 + \rho^2 \bar{X}^2 \left(\frac{1-f}{n}\right) C_x^2 - 2\rho \bar{Y}\bar{X} \left(\frac{1-f}{n}\right) \rho C_x C_y \\ \Rightarrow MSE(\bar{y}_z) &= \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{Y}\bar{X} C_x C_y). \end{aligned}$$

Hence this is the theorem.

### 3. Efficiency Comparison

This section illustrates the conditions, for which the mean square error of newly proposed regression type estimator will be minimum compared to that of simple random sampling mean estimator, product estimator and ratio estimator.

**Theorem 3.** The proposed estimator  $\bar{y}_z$  is more efficient than the product estimator  $\bar{y}_p$  if

$$\begin{aligned} \rho < 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{Y}\bar{X} C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} - 2\bar{Y}^2 C_x C_y \right) > 0. \\ \text{or} \\ \rho > 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{Y}\bar{X} C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} - 2\bar{Y}^2 C_x C_y \right) < 0. \end{aligned}$$

**Proof:**

The proposed estimator  $\bar{y}_z$  is more efficient than product estimator  $\bar{y}_p$  if

$$\begin{aligned} MSE(\bar{y}_z) &< MSE(\bar{y}_p) \\ \Rightarrow \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{Y}\bar{X} C_x C_y) &< \left(\frac{1-f}{n}\right) \bar{Y}^2 (C_y^2 + C_x^2 + 2\rho C_x C_y) \\ \Rightarrow \bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{Y}\bar{X} C_x C_y - \bar{Y}^2 C_y^2 - \bar{Y}^2 C_x^2 - 2\bar{Y}^2 \rho C_x C_y &< 0 \\ \Rightarrow \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{Y}\bar{X} C_x C_y - \bar{Y}^2 C_x^2 - 2\bar{Y}^2 \rho C_x C_y &< 0 \\ \Rightarrow \rho \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{Y}\bar{X} C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} - 2\bar{Y}^2 C_x C_y \right) &< 0 \end{aligned}$$

Now there are two cases:

**Case 1.** The inequality will be satisfied if

$$\rho < 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{Y}\bar{X} C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} - 2\bar{Y}^2 C_x C_y \right) > 0.$$

This condition holds in practice.

**Case 2.** The inequality will be satisfied if

$$\rho > 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} - 2\bar{Y}^2 C_x C_y \right) < 0.$$

This condition holds in practice.

**Theorem 4.** The estimator  $\bar{y}_z$  is more efficient than the ratio estimator  $\bar{y}_R$  if

$$\rho < 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} + 2\bar{Y}^2 C_x C_y \right) > 0.$$

or

$$\rho > 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} + 2\bar{Y}^2 C_x C_y \right) < 0.$$

**Proof:**

The proposed estimator  $\bar{y}_z$  is more efficient than ratio estimator  $\bar{y}_R$  if

$$MSE(\bar{y}_z) < MSE(\bar{y}_R)$$

$$\Rightarrow \left( \frac{1-f}{n} \right) (\bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y}C_x C_y)$$

$$< \left( \frac{1-f}{n} \right) \bar{Y}^2 (C_y^2 + C_x^2 - 2\rho C_x C_y)$$

$$\Rightarrow \bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 C_y^2 - \bar{Y}^2 C_x^2 + 2\rho \bar{Y}^2 C_x C_y < 0$$

$$\Rightarrow \rho \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} + 2\bar{Y}^2 C_x C_y \right) < 0$$

$$\Rightarrow \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 C_x^2 + 2\rho \bar{Y}^2 C_x C_y < 0$$

Now there are two cases:

**Case 1.** The inequality will be satisfied if

$$\rho < 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} + 2\bar{Y}^2 C_x C_y \right) > 0.$$

This condition holds in practice.

**Case 2.** The inequality will be satisfied if

$$\rho > 0 \text{ and } \left( \rho \bar{X}^2 C_x^2 - 2\rho \bar{X}\bar{Y}C_x C_y - \bar{Y}^2 \frac{C_x^2}{\rho} + 2\bar{Y}^2 C_x C_y \right) < 0.$$

This condition holds in practice.

**Theorem 5.** The estimator  $\bar{y}_z$  is more efficient than simple random sample mean estimator if

$$\rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y}C_x C_y < 0.$$

**Proof:**

The proposed estimator  $\bar{y}_z$  is more efficient than simple random variance estimator  $\bar{y}$  if

$$MSE(\bar{y}_z) - MSE(\bar{y}) < 0$$

$$\begin{aligned} &\Rightarrow \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y) - \left(\frac{1-f}{n}\right) \bar{Y}^2 C_y^2 < 0 \\ &\Rightarrow \bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y - \bar{Y}^2 C_y^2 < 0 \\ &\Rightarrow \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y < 0. \end{aligned}$$

This condition holds in practice.

**Theorem 6.** The estimator  $\bar{y}_z$  is more efficient than simple linear regression estimator  $\bar{y}_{LR}$  if

$$\bar{X}^2 C_x^2 - 2\bar{X}\bar{Y} C_x C_y + S_y^2 < 0.$$

**Proof:**

The proposed estimator  $\bar{y}_z$  is more efficient than simple linear regression estimator  $\bar{y}_{LR}$  if

$$\begin{aligned} &MSE(\bar{y}_z) - MSE(\bar{y}_{LR}) < 0 \\ &\Rightarrow \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y) - \left(\frac{1-f}{n}\right) S_y^2 (1-\rho^2) < 0 \\ &\Rightarrow \bar{Y}^2 C_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y - S_y^2 (1-\rho^2) < 0 \\ &\Rightarrow \bar{Y}^2 \frac{S_y^2}{\bar{Y}^2} + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y - S_y^2 + \rho^2 S_y^2 < 0 \\ &\Rightarrow S_y^2 + \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y - S_y^2 + \rho^2 S_y^2 < 0 \\ &\Rightarrow \rho^2 \bar{X}^2 C_x^2 - 2\rho^2 \bar{X}\bar{Y} C_x C_y + \rho^2 S_y^2 < 0 \\ &\Rightarrow \bar{X}^2 C_x^2 - 2\bar{X}\bar{Y} C_x C_y + S_y^2 < 0. \end{aligned}$$

This condition does not hold in practice.

### 4. Simulation Study

In this section, we will discuss the simulation study for the proposed estimator. For this purpose, we have to generate data from standard normal distribution. Steps for generating data from normal distribution and estimating relative efficiency of different estimator with simple random sampling estimator are given below:

Step 1: Choose a random sample of size  $n$ .

Step 2: Generate normal random variable with specified mean and variance.

Step 3: Calculate mean square error and relative efficiency of considered estimator.

Step 4: Repeat step1 to step3 until desired result is obtained.

**Table 1** provides the estimated values of Mean Square Error and Relative Efficiency of proposed estimator and different estimators for simulated data.

The above information delineates that proposed estimator is more efficient than other estimator for simulating data.

### 5. Real Life Example

To illustrate the applicability of the proposed estimators over other estimators

through real data sets, we used data in [6]. The data used in this study to illustrate the application of proposed estimator is a database for the amount of real and non-real estate farm loans in different states during 1997. The result, description and necessary data statistics of the populations are given as follows:

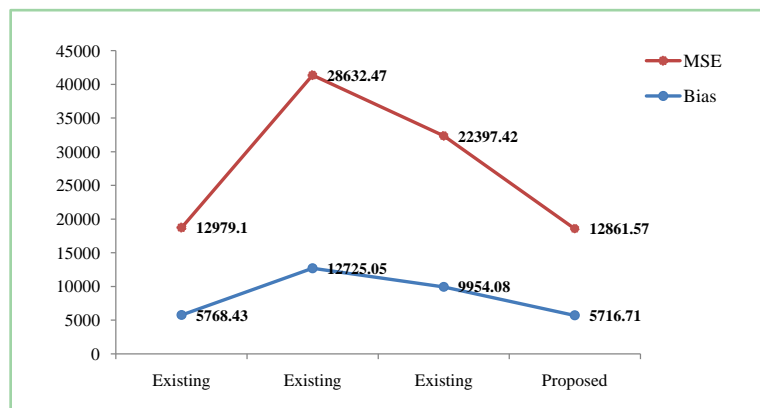
$y$  = Amount (in \$000) of real estate farm loans in different states during 1997

$x$  = Amount (in \$000) of non-real estate farm loans in different states during 1997

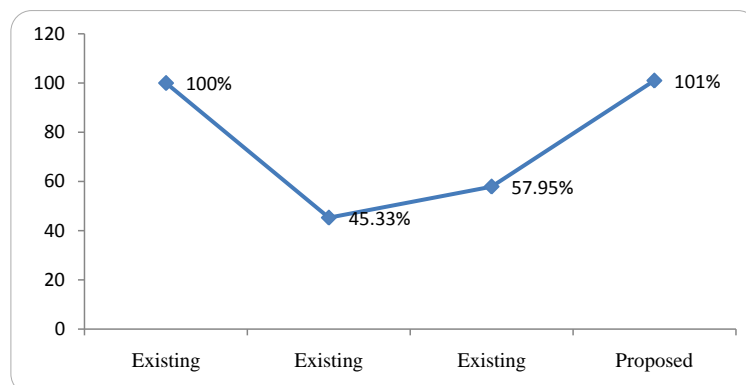
$$N = 50, \bar{x} = 964.70, \bar{y} = 582.58, n = 30, s_x^2 = 1145801, s_y^2 = 432636.7, c_x^2 = 1.2311, c_y^2 = 1.2747, \rho = -0.1222, f = 0.6.$$

**Table 2** shows the estimated values of the bias, mean square error and relative efficiency of proposed estimator and different estimator for real data.

**Figure 1** shows the bias and mean square error of the proposed estimator is less than the bias and mean square error of the existing random sample estimator, product and ratio estimators. To examine the efficiency of the proposed estimator  $\bar{y}_z$  over the estimator  $\bar{y}$ ,  $\bar{y}_R$  and  $\bar{y}_p$ , we calculated the percentage of relative efficiency of different estimator of  $\bar{Y}$  with respect to usual estimator  $\bar{y}$  provided in **Table 1** and **Table 2** and **Figure 2**. The proposed estimator  $\bar{y}_z$  performed better than other estimators.



**Figure 1.** Bias and mean square error of proposed and existing estimators.



**Figure 2.** Relative efficiency of proposed and existing estimators.



**Table 1.** Mean square error and relative efficiency comparison of proposed estimator and different estimator of  $\bar{Y}$  with respect to the usual estimator  $\bar{y}$  for simulated data.

Estimator	MSE	Relative Efficiency
$\bar{y}$	0.64	100
$\bar{y}_R$	7.47	8.53
$\bar{y}_p$	3.82	16.67
$\bar{y}_z$	0.33	193.65

**Table 2.** Bias, mean square error and relative efficiency comparison of proposed estimator and different estimator of  $\bar{Y}$  with respect to the usual estimator  $\bar{y}$ .

Estimator	Bias	MSE	Relative Efficiency
$\bar{y}$	5768.43	12,979.1	100.00
$\bar{y}_R$	12,725.05	28,632.47	45.33
$\bar{y}_p$	9954.08	22,397.42	57.95
$\bar{y}_z$	5716.71**	12,861.57**	101.00**

## 6. Conclusion

In this demonstration, we developed a new regression type estimator using the correlation coefficient Equation (1) and theoretically showed that the proposed estimator is more efficient than other usual ratio estimators, product estimator and simple random sampling mean estimator under some certain conditions. These theoretical conditions are also satisfied by the results of a numerical example.

## Acknowledgements

The authors gratefully acknowledge the excellent research facility provided by Bangabandhu Sheikh Mujibur Rahman Science and Technology University. M. Zahid Hasan would like to specifically thank Ministry of Science and Technology Government of the People's Republic of Bangladesh for research support through R & D.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Upadhyaya, L.N., Singh, H.P., Chatterjee, S. and Yadav, R. (2011) Improved Ratio and Product Exponential Type Estimators. *Journal of Statistical Theory and Practice*, **5**, 285-302. <https://doi.org/10.1080/15598608.2011.10412029>
- [2] Kadilar, C. and Cingi, H. (2005) A New Estimator Using Two Auxiliary Variables. *Applied Mathematics and Computation*, **162**, 901-908. <https://doi.org/10.1016/j.amc.2003.12.130>

- 
- [3] Subramani, J. and Kumarapandiyan, G. (2013) New Modified Ratio Estimator for Estimation of Population Mean When Median of the Auxiliary Variable Is Known. *Pakistan Journal of Statistics and Operation Research*, **9**, 137-145.  
<https://doi.org/10.18187/pjsor.v9i2.486>
- [4] Adebola, F.B., Adegoke, N.A. and Sanusi, R.A. (2015) A Class of Regression Estimator with Cum-Dual Ratio Estimator as Intercept. *International Journal of Statistics and Probability*, **4**, 42-50.
- [5] Hasan, M.Z., Hossian, M.A., Sultana, M., Fatema, K. and Hossain, M.M. (2019) A New Modified Product Estimator for Estimation of Population Mean When Median of the Auxiliary Variable Is Known. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, **6**, 108-113.
- [6] Sing, S. (2003) *Advanced Sampling Theory with Applications*. Springer Science & Business, Media Dordrecht.
- [7] Cochran, W.G. (1977) *Sampling Techniques*. 3rd Edition, John Wiley & Sons, Hoboken.