OPEN ACCESS

KJMR

Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 2 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

# Regression-Based Mean Estimator Utilizing Rank and Empirical Distribution Function Both as Dual of a Single Auxiliary Variable

***Sarhad Ullah Khan****
*National College of Business Administration & Economics, Lahore, Pakistan.*

***Muhammad Hanif***
*National College of Business Administration & Economics, Lahore, Pakistan.*

***Kalim Ullah***
*Department of Anesthesiology Aga Khan University Karachi.*

*[*]Corresponding author: Sarhad Ullah Khan (sarhadktk@gmail.com)*
***DOI: https://doi.org/10.71146/kjmr257***

**Article Info**



**Abstract**

This study proposes a novel regression-type estimator for estimating the finite population mean under simple random sampling by incorporating both the rank and the empirical distribution function (EDF) as dual of a single auxiliary variable. The pro- posed estimator utilizes the distributional properties of the auxiliary variable to enhance estimation efficiency. Theoretical properties, including the mean square error (MSE) and bias of the estimator, are derived, and its efficiency is evaluated using real-life data. The results demonstrate the practical applicability and improved accuracy of the proposed estimator in survey sampling.

**Keywords:** *Regression estimator, finite population mean, auxiliary variable, rank function, empirical distribution function, dual variable, simple random sampling, mean square error.*

## Introduction

Sampling theory plays a pivotal role in statistical research, providing tools to estimate population parameters from finite population samples. The use of auxiliary variables has long been recognized as an effective means of improving estimator precision by minimizing sampling errors. Traditionally, auxiliary variables have enhanced the efficiency of estimators through straightforward applications. However, innovative approaches have recently emerged that utilize rank and EDF as dual of a single auxiliary variable to achieve even greater precision.

The EDF serves as a non-parametric approach to estimating the cumulative distribution function based on observed values of an auxiliary variable. It proves to be effective in situations where the underlying distribution of the auxiliary variable remains unknown or complex to model. Extending its application to the estimation of population parameters introduces a robust approach for enhancing precision, particularly when conventional auxiliary variables fail to capture intricate relationships between study and auxiliary variables.

Rank, as another aspect of auxiliary variables, provides an alternative perspective by utilizing the order of observations. This method has demonstrated higher efficiency in certain scenarios, especially when traditional auxiliary variables do not yield satisfactory results.

Previous studies have explored these methods in survey sampling and finite population estimation. Mak and Kuk (1993) demonstrated the efficiency of auxiliary variable- based estimators in reducing bias. Pandey et al. (2021) addressed non-response issues using auxiliary information, while Zaman and Kadilar (2021) introduced regression-type estimators incorporating EDF to improve precision. Singh and Solanki (2013) pioneered rank-based ratio estimators, highlighting their superiority over traditional ratio estima- tors. Similarly, Kadilar and Cingi (2006) proposed a hybrid estimator combining rank and regression methods. Haq et al. (2017) developed a rank-based calibration estimator, and Hussain et al. (2022) achieved enhanced efficiency with a ratio-type estimator leveraging EDF.

This study builds on these advancements by developing a regression-type estimator for the mean of finite population under simple random sampling, that utilized rank and EDF as dual of a single auxiliary variable. The proposed methodology aims to improve estimation accuracy and efficiency, addressing gaps in traditional and contemporary approaches.

## 2. Notations and Methods

Consider the finite population of

$U = \{U_1, U_2, \ldots\ldots\ldots, U_{N}\}$ of size N , where

$Y = \{Y1, Y_2, \ldots\ldots\ldots\ldots, Y_N\}$ represents the study variable, and

$Z = \{ Z_1, Z_2, \ldots\ldots, Z_N\}$ is the corresponding auxiliary variable.

Let $y_i$ and $z_i$ denote the $i^{th}$ observations of the study variable Y and the supplementary variable Z,

respectively. These variables exhibit a certain degree of correlation $(\rho_{yz})$ within

the population U. A random subset of size n is selected from the population using SRSWOR.

Let

The population mean and variance of the study characteristic Y are given by:

$$\bar{Y} = \sum_{i=1}^{N} Y_i/N \quad \text{and} \quad S_y^2 = \sum_{i=1}^{N} (Y_i - \bar{Y})^2/N - 1,$$

respectively. In case of subset, the corresponding quantities for the study variable y are:

$$\bar{y} = \sum_{i=1}^{n} y_i/n \quad \text{and} \quad s_y^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2/(n-1).$$

Similarly, the mean and variance of the auxiliary variable Z for the population and subset are defined as follows:

$$\bar{Z} = \sum_{i=1}^{N} Z_i/N, \quad S_x^2 = \sum_{i=1}^{N} (Z_i - \bar{Z})^2/(N-1),$$

and for the sample:

$$\bar{z} = \sum_{i=1}^{n} z_i/n, \quad s_x^2 = \sum_{i=1}^{n} (z_i - \bar{z})^2/(n-1).$$

Next, we define the rank of the auxiliary variable as:

$R_z = \{r_{z1}, r_{z2}, \ldots, r_{zN}\}$, where $r_i$ represents the rank of $Z_i$, and the

Empirical Distribution Function(EDF) of the auxiliary variable $Z = \{z_1, z_2, \ldots, z_N\}$ as:

F_z = $\{f_{z1}, f_{z2}, \ldots, f_{zN}\}$,

where $f_{xi}$ is the EDF value for each $z_i$. The mean and variance of $R_z$ and $F_z$ in both the population and subset cases are defined as follows:

$$\bar{R}_z = \sum_{i=1}^{N} r_{zi}/N\,,\ S_{r_z}^2 = \sum_{i=1}^{N} (r_{zi} - \bar{R}_z)/(N-1), \quad \bar{r}_z = \sum_{i=1}^{n} r_{zi}/n \quad \text{and} \quad s_{r_z}^2 = \sum_{i=1}^{n} (r_{xi} - \bar{r}_z)/(n-1).$$

Similarly, the mean and variance of \( F_z \) for the population and subset are defined as:

$$\bar{F}_z = \sum_{i=1}^{N} f_{zi}/N, \quad S_{F_z}^2 = \sum_{i=1}^{N} (F_{zi} - \bar{F}_z)^2/(N-1),$$

for the population, and for the sample:

$$\bar{f}_z = \sum_{i=1}^{n} f_{zi}/n, \quad s_{f_z}^2 = \sum_{i=1}^{n} (f_{zi} - \bar{f}_z)^2/(n-1).$$

The covariance terms between the study variable Y and the auxiliary variables Z, $R_z$ and Fz are defined as follows:

$$S_{yz} = \sum_{i=1}^{N}(Y_i - \bar{Y})(Z_i - \bar{Z})/(N-1), \quad S_{yr_z} = \sum_{i=1}^{N}(Y_i - \bar{Y})(R_{zi} - \bar{R}_z)/(N-1),$$

$$S_{zr_z} = \sum_{i=1}^{N}(Z_i - \bar{Z})(R_{zi} - \bar{R}_z)/(N-1), \quad S_{yf_z} = \sum_{i=1}^{N}(Y_i - \bar{Y})(F_{zi} - \bar{F}_z)/(N-1),$$

and

$$S_{zf_z} = \sum_{i=1}^{N}(Z_i - \bar{Z})(F_{zi} - \bar{F}_z)/(N-1).$$

Next, we define the correlation coefficients between $Z_i$ and $R_{zi}$ , and between $Z_i$ and $F_{zi}$ as:

$$\rho_{zr_z} = \frac{S_{zr_z}}{S_z S_{r_z}} \quad \text{and} \quad \rho_{zf_z} = \frac{S_{zf_z}}{S_z S_{f_z}},$$

where $S_{zrz}$ and $S_{zfz}$ represent the population covariances between $Z_i$ and $R_{zi}$, and $Z_i$ and $F_{zi}$ , respectively.

For the derivation of the biases and MSEs of the suggested estimators, we defined the error terms as:

Let

$$\xi_0 = (\bar{y} - \bar{Y})/\bar{Y}, \quad \xi_1 = (\bar{z} - \bar{Z})/\bar{Z}, \quad \xi_2 = (\bar{r}_z - \bar{R}_z)/\bar{R}_z, \quad \xi_3 = (\bar{f}_z - \bar{F}_z)/\bar{F}_z,$$

ensuring that

$$E(\xi_0) = E(\xi_1) = E(\xi_2) = E(\xi_3) = 0,$$

$$E(\xi_0^2) = \theta C_y^2, \quad E(\xi_1^2) = \theta C_z^2, \quad E(\xi_2^2) = \theta C_{r_z}^2, \quad E(\xi_3^2) = \theta C_{f_z}^2,$$

$$E(\xi_0\xi_1) = \theta\rho_{yz}C_yC_z, \quad E(\xi_0\xi_2) = \theta\rho_{yr_z}C_yC_{r_z},$$

$$E(\xi_0\xi_3) = \theta\rho_{yf_z}C_yC_{f_z}, \quad E(\xi_1\xi_2) = \theta\rho_{zr_z}C_zC_{r_z},$$

$$E(\xi_1\xi_3) = \theta\rho_{zf_z}C_zC_{f_z}.$$

and

$\theta = (\frac{1}{n} - \frac{1}{N})$ is the correction factor of finite population.

Some considered estimators that are currently available in the literature, along with their MSEs, are thoroughly explained in Section 2. In Section 3, we explored and derived the proposed efficient regression and ratio types estimators using rank and EDF as dual of supplementary variable and compared the estimators' outcomes under various scenarios. The practical usefulness of the suggested estimators is assessed by applying them to real-world data sets in Section 4,

and the work is concluded and some future research topics are discussed in Section 5.

## 3. ESTIMATORS AVAILABLE IN LITERATURE

In this section, we analyzed a range of existing estimators employed for estimating the mean of the finite population. The MSEs of all the considered estimators were derived and presented using first-order approximations.

### 3.1 The Usual Mean Estimator

$\bar{y}$ is the mean of this unbiased estimator $\hat{y}_1$ and the variance is:

$$Var(\bar{y_1}) = \theta \bar{Y}^2 C_y^2 \qquad (1)$$

### 3.2 The Foundational Ratio and Product Estimators

The foundational ratio proposed by Cochran(1940) and product suggested by Murthy (1964)

estimators are given as:

$$\hat{\bar{Y}}_2 = \bar{y}\left(\frac{\bar{Z}}{\bar{z}}\right) \qquad (2)$$

and

$$\hat{\bar{Y}}_3 = \bar{y}\left(\frac{\bar{z}}{\bar{Z}}\right) \qquad (3)$$

The MSEs of the aforementioned ratio and product estimators are given as follows:

$$MSE(\hat{\bar{Y}}_2) \cong \theta \bar{Y}^2 (C_y^2 + C_z^2 - 2\rho_{yz} C_y C_z),$$

and

$$MSE(\hat{\bar{Y}}_3) \cong \theta \bar{Y}^2 (C_y^2 + C_z^2 + 2\rho_{yz} C_y C_z)\,.$$

### 3.3 Conversion of Power Estimator

Respectively, Srivastava (1967) proposed the power transformation estimator as:

$$\hat{\bar{Y}}_4 = \bar{y}\left(\frac{\bar{z}}{\bar{Z}}\right)^{\pi} \qquad (4)$$

The MSE of estimator $\hat{\bar{Y}}_4$ is

$$MSE(\hat{\bar{Y}}_4) \cong \theta \bar{Y}^2 [C_y^2 + \pi C_z^2 (\pi - 2\beta_{yz})]\,,$$

where

$$\beta_{yz} = \frac{C_{yz}}{{C_z}^2}\ .$$

The minimum MSE of the $\hat{\bar{Y}}_4$ for the optimum level of $\upsilon = \beta_{yz}$ is

$$MSE_{(min.)}(\hat{\bar{Y}}_4) = \theta \bar{Y}^2 C_y^2 [1 - \rho_{yz}^2]$$

where

$$\rho_{yz}^2 = \frac{C_{yz}^2}{C_y^2 C_z^2}\ .$$

\par (MSE) of $\hat{\bar{Y}}_4$ can be conveniently determined by substituting the value of $\upsilon = (1, -1)$.

### 3.4 Standard Regression Estimator

The standard regression estimator proposed by Cochran (1963) is described as:

$$\hat{\bar{Y}}_5 = \bar{y} + \hat{\beta}_{yz}(\bar{Z} - \bar{z}) \qquad (5)$$

where

$$\hat{\beta}_{yz} = \frac{s_{yz}}{s_z^2}\ .$$

The MSE of the $\hat{\bar{Y}}_5$, up to the first approximation, is given as:

$$MSE(\hat{\bar{Y}}_5) = \theta \bar{Y}^2 [C_y^2 + P_{yz} C_z^2 (P_{yz} - 2\beta_{yz})]\ ,$$

where

$$P_{yz} = R\beta_{yz} \quad and \quad R = \frac{\bar{Z}}{\bar{Y}}\ .$$

so that

$$MSE_{(min.)}(\hat{\bar{Y}}_5) = min.\, MSE(\hat{\bar{Y}}_4)$$

### 3.5 Traditional Difference-Type Estimator

The Traditional difference-type estimator presented by Riaz et al. (2014) is

$$\hat{\bar{Y}}_6 = \bar{y} + w(\bar{Z} - \bar{z}) \qquad (6)$$

The optimal value of $w$ is given by

$$w_{opt} = \frac{1}{R}\beta_{yz}.$$

An enhanced version of the estimator $\hat{\bar{Y}}_6$ as proposed by Riaz et al. (2014), is defined as:

$$\hat{\bar{Y}}_7 = w_1\bar{y} + w_2(\bar{Z} - \bar{z}) \qquad (7)$$

where the constants $w_1$ and $w_2$ are chosen at random.

Up to the first order approximation, the minimal MSE of estimators $\hat{\bar{Y}}_6$ and $\hat{\bar{Y}}_7$ is

$$MSE(\hat{\bar{Y}}_6) \cong \theta\bar{Y}^2[C_y^2 + \pi C_z^2(\pi - 2\beta_{yz})]\,,$$

the optimal level of the constants $w_1$ and $w_2$ are

$$w_{1opt} = \frac{1}{C_y^2(1-\rho_{yz}^2)} \quad \text{and} \quad w_{2opt} = (\frac{1}{R}\beta_{yz})\left(\frac{1}{C_y^2(1-\rho_{yz}^2)}\right).$$

Following the entry of these values, the estimator $\hat{\bar{Y}}_6$ and $\hat{\bar{Y}}_7$'s minimal MSE is

$$MSE(\hat{\bar{Y}}_6) = MSE_{(min.)}(\hat{\bar{Y}}_4)\,,$$

or

$$MSE_{(min.)}(\hat{\bar{Y}}_6) = MSE_{(min.)}(\hat{\bar{Y}}_4)\,,$$

After simplification

$$MSE_{(min.)}(\hat{\bar{Y}}_7) = \frac{\bar{Y}^2 MSE_{(min.)}(\hat{\bar{Y}}_6)}{\bar{Y}^2 + MSE_{(min.)}(\hat{\bar{Y}}_6)}\,.$$

**3.6 Regression-Cum Exponential Type Estimator Presented by Hussain et al. (2022)**

They suggested a regression cum exponential-type estimator using the EDF as a dual use of the auxiliary variable, defined as:

$$\hat{\bar{Y}}_6 = \left(\pi_1\bar{y} + \pi_2(\bar{Z} - \bar{z}) + \pi_3(\bar{F}_z - \bar{F}_z)\right)\exp\left(\frac{a(\bar{Z}-\bar{z})}{a(\bar{Z}+\bar{z})+2b}\right).$$

Here $\pi_1, \pi_2,$ and $\pi_3$ are the constants determined to minimize the MSE of the presented class, parameters a and b serve as key determinants, enabling the generation of various members within the suggested family.

The optimum values of constants $\pi_1, \pi_2,$ and $\pi_3$ are given as:

$$\pi_{1(opt)} = \frac{8-\theta^2 V_z^2}{8(1+V_y^2(1-R^2_{yzf_z}))},$$

$$\pi_{2(opt)} = \frac{\bar{Y}\left(\theta^3 V_z^3(\rho^2_{zf_z}-1)+V_z(\theta^2 V_z^2-8)(\rho_{yx}-\rho_{yx}\rho_{xf_x})+4\theta V_x(\rho_{xf_x}-1)(1+V_y^2(1-R^2_{yzf_z}))\right)}{8\bar{Z}V_z(\rho^2_{zf_z}-1)(1+V_y^2(1-R^2_{yzf_z}))},$$

and

$$\pi_{3(opt)} = \frac{\bar{Y}V_y(8-\theta^2V_z^2)(\rho_{yf_z}\rho_{zf_z}-\rho_{yf_z})}{8\bar{F}_z V_{f_z}(\rho^2_{zf_z}-1)(1+V_y^2(1-R^2_{yzf_z}))}.$$

The minimum MSE of the estimator $\hat{\bar{Y}}_6$ is given as:

$$MSE_{(min.)}(\hat{\bar{Y}}_5) \cong Var(\hat{\bar{Y}}_4) - L - M$$

where

$$Var(\hat{\bar{Y}}_4) = \bar{Y}^2 V_y^2(1-\rho^2_{yz}),$$

$$L = \frac{\bar{Y}^2(\theta^2V_z^2+8V_y^2(1-\rho^2_{yz}))^2}{64(1+V_y^2(1-\rho^2_{yz}))},$$

And

$$M = \frac{\bar{Y}^2V_y^2(\rho^2_{yf_z}-\rho_{yz}\rho_{zf_z})^2(-8+\theta^2V_z^2)^2}{64(1-\rho^2_{zf_z})(1+V_y^2(1-\rho^2_{yz}))(1+V_y^2(1-R^2_{yzf_z}))}.$$

## 4. PROPOSED ESTIMATOR

Building on the work of Hussain et al. (2022), we suggested a new ratio in regression cum exponential-type estimator for estimating the mean of the finite population $\bar{Y}$ of the research character Y. This estimator utilizes the auxiliary character (z) along with its rank $R_z$ and Empirical Distribution Function ($F_z$) as dual of a single auxiliary variable to improve the efficiency of estimators under simple random sampling. The proposed ratio in regression cum exponential-type estimator is formulated as follows:

$$\hat{\bar{Y}}_{Prop} = \left[\gamma\,\bar{y} + \pi_1(\bar{Z}-\bar{z}) + \pi_2(\bar{R}_z-\bar{r}_z) + \pi_3(\bar{F}_z-\bar{f}_z)\right]\exp\left(\frac{(\bar{Z}-\bar{z})}{(\bar{Z}+\bar{z})}\right). \qquad (4.1)$$

where the values of the Constants $\gamma, \pi_1, \pi_2$ and $\pi_3$ are to be selected accordingly.

The bias and MSE of the proposed Estimator are extended up to the 1st approximation.

After supposing the values of $\bar{y}, \bar{z}, \bar{r}_z$ and $\bar{f}_z$ we have

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) + \pi_1\{\bar{Z} - \bar{Z}(1+\xi_1)\} + \pi_2\{\bar{R}_z - \bar{R}_z(1+\xi_2)\} + \{\pi_3\{\bar{F}_z - \bar{F}_z(1+\xi_3)\}\right) \exp\left(\frac{(\bar{Z} - \bar{Z}(1+\xi_1))}{(\bar{Z} + \bar{Z}(1+\xi_1))}\right),$$

or

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) + \pi_1\{\bar{Z} - \bar{Z} + \bar{Z}\xi_1\} + \pi_2\{\bar{R}_z - \bar{R}_z - \bar{R}_z\xi_2\} + \{\pi_3\{\bar{F}_z - \bar{F}_z - \bar{F}_z\xi_3\}\right) \exp\left(\frac{-\bar{Z}\xi_1}{2\bar{Z} + \bar{Z}\xi_1}\right),$$

or

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3\right) \exp\left(\frac{-\bar{Z}\xi_1}{2\bar{Z}(1+\frac{\xi_1}{2})}\right),$$

Or

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3\right) \exp\left(\frac{-\xi_1}{2}(1+\frac{\xi_1}{2})^{-1}\right),$$

or

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3\right) \exp\left(\frac{-\xi_1}{2}(1-\frac{\xi_1}{2}+\frac{\xi_1^2}{4})\right),$$

Extended up-to first order

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3\right) \exp\left(-\frac{\xi_1}{2}+\frac{\xi_1^2}{4}\ldots\right),$$

or

$$\hat{\hat{Y}}_{Prop} = \left(\gamma\,\bar{Y}(1+\xi_0) - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3\right) \left(1-\frac{\xi_1}{2}+\frac{\xi_1^2}{4}\ldots\right),$$

or

$$\hat{\hat{Y}}_{Prop} = \gamma\,\bar{Y}(1+\xi_0) - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3 - (\gamma\,\bar{Y}(1+\xi_0))\frac{\xi_1}{2}$$
$$+\pi_1\bar{Z}\frac{\xi_1^2}{2} + \pi_2\bar{R}_z\frac{1}{2}\xi_1\xi_2 + \pi_3\bar{F}_z\frac{1}{2}\xi_1\xi_3 + \gamma\,\bar{Y}(1+\xi_0)\frac{1}{4}\xi_1^2$$

or

$$\hat{\bar{Y}}_{Prop} = \gamma\,\bar{Y} + \gamma\,\bar{Y}\xi_0 - \pi_1\bar{Z}\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3 - \gamma\,\bar{Y}\frac{1}{2}\xi_1 + \gamma\,\bar{Y}\frac{1}{2}\xi_0\xi_1 + \pi_1\bar{Z}\frac{1}{2}\xi_1^2 + \pi_2\bar{R}_z\frac{1}{2}\xi_1\xi_2 + \pi_3\bar{F}_z\frac{1}{2}\xi_1\xi_3 + \gamma\,\bar{Y}\frac{1}{4}\xi_1^2$$

Subtracting $\bar{Y}$ from both sides for bias

$$\hat{\bar{Y}}_{Prop} - \bar{Y} = \gamma\,\bar{Y} - \bar{Y} + \gamma\,\bar{Y}\xi_0 - (\gamma\,\bar{Y}\frac{1}{2} + \pi_1\bar{Z})\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3 - \gamma\,\bar{Y}\frac{1}{2}\xi_0\xi_1 + \pi_2\bar{R}_z\frac{1}{2}\xi_1\xi_2 + \pi_3\bar{F}_z\frac{1}{2}\xi_1\xi_3 + (\gamma\,\bar{Y}\frac{1}{4} + \pi_1\bar{Z})\frac{1}{2}\xi_1^2$$

By applying the expectation operator to both sides, we obtain.

$$E(\hat{\bar{Y}}_{Prop} - \bar{Y}) = \gamma\,\bar{Y} - \bar{Y} + \gamma\,\bar{Y}E(\xi_0) - (\gamma\,\bar{Y}\frac{1}{2} + \pi_1\bar{Z})E(\xi_1) - \pi_2\bar{R}_zE(\xi_2) - \pi_3\bar{F}_zE(\xi_3) + \gamma\,\bar{Y}\frac{1}{2}E(\xi_0\xi_1) + \pi_2\bar{R}_z\frac{1}{2}E(\xi_1\xi_2) + \pi_3\bar{F}_z\frac{1}{2}(\xi_1\xi_3) + (\gamma\,\bar{Y}\frac{1}{4} + \pi_1\bar{Z})\frac{1}{2}E(\xi_1^2)$$

Bias of the proposed estimator is

$$Bias(\hat{\bar{Y}}_{Prop}) = \gamma\,\bar{Y} - \bar{Y} + \pi_1\bar{Z}\frac{1}{2}V_z^2 + \pi_2\bar{R}_z\frac{1}{2}V_{zr_z} + \pi_3\bar{F}_z\frac{1}{2}V_{zf_z} + \gamma\,\bar{Y}\frac{1}{4}V_z^2 + \gamma\,\bar{Y}\frac{1}{2}V_{yz}\,,$$

or

$$Bias(\hat{\bar{Y}}_{Prop}) = \bar{Y}(\gamma\ - 1) + \frac{1}{4}[2(\pi_1\bar{Z}V_z^2 + \pi_2\bar{R}_zV_{zr_z} + \pi_3\bar{F}_zV_{zf_z}) + \gamma\,\bar{Y}(V_z^2 + 2V_{yz})].$$

For MSE  we square on both sides and have

$$(\hat{\bar{Y}}_{Prop} - \bar{Y})^2 = (\bar{Y}(\gamma\ - 1) + \gamma\,\bar{Y}\xi_0 - (\gamma\,\bar{Y}\frac{1}{2} + \pi_1\bar{Z})\xi_1 - \pi_2\bar{R}_z\xi_2 - \pi_3\bar{F}_z\xi_3 - \gamma\,\bar{Y}\frac{1}{2}\xi_0\xi_1 + \pi_2\bar{R}_z\frac{1}{2}\xi_1\xi_2 + \pi_3\bar{F}_z\frac{1}{2}\xi_1\xi_3 + (\gamma\,\bar{Y}\frac{1}{4} + \pi_1\bar{Z})\frac{1}{2}\xi_1^2)^2$$

Open the square up-to second ordered

$$(\hat{\bar{Y}}_{Prop} - \bar{Y})^2 = \bar{Y}^2(\gamma - 1)^2 + \gamma^2\bar{Y}^2\xi_0^2 + (\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})^2\xi_1^2 + \pi_2^2\bar{R}_z^2\xi_2^2 + \pi_3^2\bar{F}_z^2\xi_3^2$$

$$+2\bar{Y}(\gamma - 1)\gamma\bar{Y}\xi_0 - 2\bar{Y}(\gamma - 1)(\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})\xi_1 - 2\bar{Y}(\gamma - 1)\pi_2\bar{R}_z\xi_2 - 2\bar{Y}(\gamma - 1)\pi_3\bar{F}_z\xi_3 - 2\bar{Y}(\gamma - 1)\gamma\bar{Y}\frac{1}{2}\xi_0\xi_1 + 2\bar{Y}(\gamma - 1)\pi_2\bar{R}_z\frac{1}{2}\xi_1\xi_2$$

$$-2\bar{Y}(\gamma - 1)\pi_3\bar{F}_z\xi_3 - 2\bar{Y}(\gamma - 1)\gamma\bar{Y}\frac{1}{2}\xi_0\xi_1 + 2\bar{Y}(\gamma - 1)\pi_2\bar{R}_z\frac{1}{2}\xi_1\xi_2 + 2\bar{Y}(\gamma - 1)\pi_3\bar{F}_z\frac{1}{2}\xi_1\xi_3 - 2\bar{Y}(\gamma - 1)(\gamma\bar{Y}\frac{1}{4} + \pi_1\bar{Z})\frac{1}{2}\xi_1^2$$

$$-2\gamma\bar{Y}(\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})\xi_0\xi_1 - 2\gamma\bar{Y}\pi_2\bar{R}_z\xi_0\xi_2 - 2\gamma\bar{Y}\pi_3\bar{F}_z\xi_0\xi_3 + 2(\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})\pi_2\bar{R}_z\xi_1\xi_2 + 2(\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})\pi_3\bar{F}_z\xi_1\xi_3 + 2\pi_2\bar{R}_z\pi_3\bar{F}_z\xi_2\xi_3.$$

By applying the expectation operator to both sides, we obtain

$$E(\hat{\bar{Y}}_{Prop} - \bar{Y})^2 = \bar{Y}^2(\gamma - 1)^2 + \gamma^2\bar{Y}^2E(\xi_0^2) + (\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})^2E(\xi_1^2) + \pi_2^2\bar{R}_z^2E(\xi_2^2) + \pi_3^2\bar{F}_z^2E(\xi_3^2)$$

$$+2\bar{Y}(\gamma - 1)\gamma\bar{Y}E(\xi_0) - 2\bar{Y}(\gamma - 1)(\gamma\bar{Y}\frac{1}{2} + \pi_1\bar{Z})E(\xi_1) - 2\bar{Y}(\gamma - 1)\pi_2\bar{R}_zE(\xi_2) - 2\bar{Y}(\gamma - 1)\pi_3\bar{F}_zE(\xi_3)$$

$$-2\bar{Y}(\gamma-1)\gamma\bar{Y}\frac{1}{2}E(\xi_0\xi_1)+2\bar{Y}(\gamma-1)\pi_2\bar{R}_z\frac{1}{2}E(\xi_1\xi_2)+2\bar{Y}(\gamma-1)\pi_3\bar{F}_z\frac{1}{2}E(\xi_1\xi_3)-2\bar{Y}(\gamma-1)(\gamma\bar{Y}\frac{1}{4}+\pi_1\bar{Z})\frac{1}{2}E(\xi_1^2)$$
$$-2\gamma\bar{Y}(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})E(\xi_0\xi_1)-2\gamma\bar{Y}\pi_2\bar{R}_zE(\xi_0\xi_2)-2\gamma\bar{Y}\pi_3\bar{F}_zE(\xi_0\xi_3)+2(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})\pi_2\bar{R}_zE(\xi_1\xi_2)$$
$$+2(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})\pi_3\bar{F}_zE(\xi_1\xi_3)+2\pi_2\bar{R}_z\pi_3\bar{F}_zE(\xi_2\xi_3).$$

The MSE is

$$MSE(\hat{\hat{Y}}_{Prop})=\ \bar{Y}^2(\gamma-1)^2+\gamma^2\bar{Y}^2V_y^2+(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})^2V_z^2+\pi_2^2\bar{R}_z^2V_{r_z}^2+\pi_3^2\bar{F}_z^2V_{f_z}^2$$
$$+2\bar{Y}(\gamma-1)\gamma\bar{Y}(0)-2\bar{Y}(\gamma-1)(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})(0)-2\bar{Y}(\gamma-1)\pi_2\bar{R}_z(0)$$
$$-2\bar{Y}(\gamma-1)\pi_3\bar{F}_z(0)-2\bar{Y}(\gamma-1)\gamma\bar{Y}\frac{1}{2}V_{yz}+2\bar{Y}(\gamma-1)\pi_2\bar{R}_z\frac{1}{2}V_{zr_z}$$
$$+2\bar{Y}(\gamma-1)\pi_3\bar{F}_z\frac{1}{2}V_{zf_z}-2\bar{Y}(\gamma-1)(\gamma\bar{Y}\frac{1}{4}+\pi_1\bar{Z})\frac{1}{2}V_z^2-2\gamma\bar{Y}(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})V_{yz}-2\gamma\bar{Y}\pi_2\bar{R}_zV_{yr_z}$$
$$-2\gamma\bar{Y}\pi_3\bar{F}_zV_{yf_z}+2(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})\pi_2\bar{R}_zV_{zr_z}+2(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z})\pi_3\bar{F}_zV_{zf_z}+2\pi_2\bar{R}_z\pi_3\bar{F}_zV_{r_zf_z}.$$

or

$$MSE(\hat{\hat{Y}}_{Prop})=\ \bar{Y}^2(\gamma-1)^2+\gamma^2\bar{Y}^2V_y^2+\left(\gamma^2\bar{Y}^2\frac{1}{4}+\pi_1^2\bar{Z}^2+\gamma\bar{Y}\pi_1\bar{Z}\right)V_z^2$$
$$+\pi_2^2\bar{R}_z^2V_{r_z}^2+\pi_3^2\bar{F}_z^2V_{f_z}^2-\bar{Y}(\gamma-1)\gamma\bar{Y}V_{yz}+\bar{Y}(\gamma-1)\pi_2\bar{R}_zV_{zr_z}+\bar{Y}(\gamma-1)\pi_3\bar{F}_zV_{zf_z}$$
$$-2\bar{Y}(\gamma-1)\left(\gamma\bar{Y}\frac{1}{4}+\pi_1\bar{Z}\right)V_z^2-2\gamma\bar{Y}\left(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z}\right)V_{yz}-2\gamma\bar{Y}\pi_2\bar{R}_zV_{yr_z}-2\gamma\bar{Y}\pi_3\bar{F}_zV_{yf_z}$$
$$+2\left(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z}\right)\pi_2\bar{R}_zV_{zr_z}+2\left(\gamma\bar{Y}\frac{1}{2}+\pi_1\bar{Z}\right)\pi_3\bar{F}_zV_{zf_z}+2\pi_2\bar{R}_z\pi_3\bar{F}_zV_{r_zf_z}.$$

or

$$MSE(\hat{\hat{Y}}_{Prop})=\ \bar{Y}^2(\gamma-1)^2+\gamma^2\bar{Y}^2V_y^2+\left(\gamma^2\bar{Y}^2\frac{1}{4}+\pi_1^2\bar{Z}^2+\gamma\bar{Y}\pi_1\bar{Z}+\gamma^2\bar{Y}^2\frac{1}{4}+2\gamma\bar{Y}\pi_1\bar{Z}\right)V_z^2$$
$$+\pi_2^2\bar{R}_z^2V_{r_z}^2+\pi_3^2\bar{F}_z^2V_{f_z}^2-\left(2\gamma^2\bar{Y}^2-\gamma\bar{Y}^2+2\gamma\bar{Y}\pi_1\bar{Z}\right)V_{yz}-2\gamma\bar{Y}\pi_2\bar{R}_zV_{yr_z}$$
$$-2\gamma\bar{Y}\pi_3\bar{F}_zV_{yf_z}+\left(2\bar{Y}\gamma\pi_2\bar{R}_z-\bar{Y}\pi_2\bar{R}_z+2\pi_1\bar{Z}\pi_2\bar{R}_z\right)V_{zr_z}+\left(2\bar{Y}\gamma\pi_3\bar{F}_z-\bar{Y}\pi_3\bar{F}_z+2\pi_1\bar{Z}\pi_3\bar{F}_z\right)V_{zf_z}+2\pi_2\bar{R}_z\pi_3\bar{F}_zV_{r_zf_z}.$$

where

$$V_y^2=\tau C_y^2,\quad V_z^2=\tau C_z^2,\quad V_{r_z}^2=\tau C_{r_z}^2,\quad V_{f_z}^2=\tau C_{f_z}^2,$$
$$V_{yz}=\tau\rho_{yz}C_yC_z,\quad V_{yr_z}=\tau\rho_{yr_z}C_yC_{r_z},\quad V_{yf_z}=\tau\rho_{yf_z}C_yC_{f_z},$$
$$V_{zr_z}=\tau\rho_{zr_z}C_zC_{r_z},\quad V_{zf_z}=\tau\rho_{zf_z}C_zC_{f_z},\quad V_{r_zf_z}=\tau\rho_{r_zf_z}C_{r_z}C_{f_z}.$$

To get the optimum values of the constants $\gamma,\pi_1,\pi_2$ and $\pi_3$ we differentiate Equ.(4.14) with respect to $\gamma, \pi_1, \pi_2$ and $\pi_3$ respectively and after more clarifications we have

The constants $\gamma, \pi_1, \pi_2$ and $\pi_3$ at their optimal levels are:

$$\gamma_{opt} = -\left(V_z^2(\bar{Y}-1)+4\bar{Y}\right)\left(-2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}+\rho_{zr_z}^2+\rho_{zf_z}^2+\rho_{r_zf_z}^2-1\right)\Big/\Bigg(4V_y^2(\rho_{yz}^2-1)\rho_{r_zf_z}^2+\Big((-2\rho_{yz}\rho_{yr_z}+2\rho_{zr_z})\rho_{zf_z}-2\rho_{yf_z}(\rho_{yz}\rho_{zr_z}-\rho_{yr_z})\Big)\rho_{r_zf_z}$$

$$+\left(\rho_{yr_z}^2-1\right)\rho_{zf_z}^2+2\rho_{yf_z}(-\rho_{yr_z}\rho_{zr_z}+\rho_{yz})\rho_{zf_z}+\left(\rho_{yf_z}^2-1\right)\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}+2\rho_{r_zf_z}\rho_{zf_z}\rho_{zr_z}-\rho_{r_zf_z}^2-\rho_{yf_z}^2-\rho_{zr_z}^2+1\Bigg)\bar{Y},$$

$$\pi_{1opt} = \Big(2V_z((\rho_{yz}^2-1)\rho_{r_zf_z}^2+((-2\rho_{yz}\rho_{yr_z}+2\rho_{zr_z})\rho_{zf_z}+2\rho_{yf_z}(-\rho_{zr_z}\rho_{yz}+\rho_{yr_z}))\rho_{r_zf_z}+(\rho_{yr_z}-1)\rho_{zf_z}^2)-2\rho_{yf_z}(\rho_{yr_z}\rho_{zr_z}-\rho_{yz})\rho_{zf_z}+(\rho_{yf_z}^2-1)$$

$$\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}-\rho_{yz}^2-\rho_{yr_z}^2-\rho_{yf_z}^2+1)(\bar{Y}+1)V_y^2+\frac{1}{2}((\bar{Y}-1)(-\rho_{yz}\rho_{r_zf_z}^2$$

$$+(\rho_{yr_z}\rho_{zf_z}+\rho_{yf_z}\rho_{zr_z})\rho_{r_zf_z}-\rho_{yr_z}\rho_{zr_z}-\rho_{yf_z}\rho_{yr_z}+\rho_{yz})V_{yz}+V_z^2(-2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}+\rho_{zr_z}^2+\rho_{zf_z}^2+\rho_{r_zf_z}^2-1)(\bar{Y}-1))+4\bar{Y}$$

$$(-\rho_{yz}\rho_{r_zf_z}^2+(\rho_{yr_z}\rho_{zf_z}+\rho_{yf_z}\rho_{zr_z})\rho_{r_zf_z}-\rho_{yr_z}\rho_{zr_z}-\rho_{yf_z}\rho_{zf_z}+\rho_{yz})V_Y+2V_z(-2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}+\rho_{zr_z}^2+\rho_{zf_z}^2+\rho_{r_zf_z}^2-1)(\bar{Y}-1)\Big)\Big/$$

$$\Big(4V_z\bar{Z}((\rho_{yz}^2-1)\rho_{r_zf_z}^2+((-2\rho_{yz}\rho_{yr_z}+2\rho_{zr_z})\rho_{zf_z}+2\rho_{yf_z}(-\rho_{yz}\rho_{zr_z}+\rho_{yr_z}))\rho_{r_zf_z}+(\rho_{yr_z}^2-1)\rho_{zf_z}^2-2\rho_{yf_z}(\rho_{yr_z}\rho_{zr_z}$$

$$-\rho_{yz})\rho_{zf_z}+(\rho_{yf_z}^2-1)\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}-\rho_{yz}^2-\rho_{yr_z}^2-\rho_{yf_z}^2+1)V_y^2+2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}-\rho_{zr_z}^2-\rho_{zf_z}^2-\rho_{r_zf_z}^2+1)\Big),$$

$$\pi_{2opt} = -V_y\left(V_z^2(\bar{Y}-1)+4\bar{Y}\right)\left(\rho_{zf_z}^2\rho_{yr_z}+(-\rho_{r_zf_z}\rho_{yz}-\rho_{zr_z}\rho_{yf_z})\rho_{zf_z}+\rho_{yf_z}\rho_{r_zf_z}+\rho_{yz}\rho_{zr_z}-\rho_{yr_z}\right)\Big/$$

$$\Bigg(4\bar{R}_zV_{r_z}\Bigg(\Big((\rho_{yr_z}^2-1)\rho_{zf_z}^2+\Big((-2\rho_{yr_z}\rho_{yz}+2\rho_{zr_z})\rho_{r_zf_z}-2\rho_{yf_z}(\rho_{yr_z}\rho_{zr_z}-\rho_{yz})\Big)\rho_{zf_z}+(\rho_{yz}^2-1)\rho_{r_zf_z}^2\Big)$$

$$+2\rho_{yf_z}(-\rho_{yz}\rho_{zr_z}+\rho_{yr_z})\rho_{r_zf_z}+(\rho_{yf_z}^2-1)\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}-\rho_{yz}^2-\rho_{yr_z}^2-\rho_{yf_z}^2+1\Bigg)V_y^2+\left(2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}-\rho_{zr_z}^2-\rho_{zf_z}^2-\rho_{r_zf_z}^2+1\right)\Bigg)$$

and

$$\pi_{3opt} = -V_y\left(\rho_{yf_z}\rho_{zr_z}^2+(-\rho_{yz}\rho_{r_zf_z}-\rho_{yr_z}\rho_{zf_z})\rho_{zr_z}+\rho_{yr_z}\rho_{r_zf_z}+\rho_{yz}\rho_{zf_z}-\rho_{yf_z}\right)\left(V_z^2(\bar{Y}-1)+4\bar{Y}\right)\Big/$$

$$\Bigg(4\bar{F}_zV_{f_z}\Bigg((\rho_{yf_z}^2-1)\rho_{zr_z}^2+\Big((-2\rho_{yf_z}\rho_{yz}+2\rho_{zf_z})\rho_{r_zf_z}-2\rho_{yr_z}(\rho_{yf_z}\rho_{zf_z}-\rho_{yz})\Big)\rho_{zr_z}+(\rho_{yz}^2-1)\rho_{r_zf_z}^2\Bigg)$$

$$+2\rho_{yr_z}(-\rho_{yz}\rho_{zf_z}+\rho_{yf_z})\rho_{r_zf_z}+(\rho_{yr_z}^2-1)\rho_{zf_z}^2+2\rho_{yz}\rho_{yf_z}\rho_{zf_z}-\rho_{yz}^2-\rho_{yr_z}^2-\rho_{yf_z}^2+1\Bigg)+V_y^2\left(2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}-\rho_{zr_z}^2-\rho_{zf_z}^2-\rho_{r_zf_z}^2+1\right)$$

where

$$\rho_{yz}^2=\frac{V_{yz}^2}{V_y^2V_z^2},\quad \rho_{yr_z}^2=\frac{V_{yr_z}^2}{V_y^2V_{r_z}^2}\quad \rho_{zr_z}^2=\frac{V_{zr_z}^2}{V_z^2V_{r_z}^2}\rho_{yf_z}^2=\frac{V_{yf_z}^2}{V_y^2V_{f_z}^2},\rho_{zf_z}^2=\frac{V_{zf_z}^2}{V_z^2V_{f_z}^2}\ \ and\ \ \rho_{r_zf_z}^2=\frac{V_{r_zr_z}^2}{V_{r_z}^2V_{f_z}^2}$$

\par By substituting the values of $\gamma, \pi_1, \pi_2$ and $\pi_3$ in (4.14) we get:

$$MSE(\hat{\hat{\bar{Y}}}_{Prop}) = \Bigg(\bigg(((\rho_{yz}^2-1)\rho_{r_zf_z}^2+(-2\rho_{yz}\rho_{yr_z}+2\rho_{zr_z})\rho_{zf_z}-2\rho_{yf_z}(\rho_{yz}\rho_{zr_z}-\rho_{yr_z}))\rho_{r_zf_z}$$

$$+(\rho_{yr_z}^2-1)\rho_{zf_z}^2+2\rho_{yf_z}(-\rho_{yr_z}\rho_{zr_z}+\rho_{yz})\rho_{zf_z}+(\rho_{yf_z}^2-1)\rho_{zr_z}^2$$

$$+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}-\rho_{yz}^2-\rho_{yf_z}^2-\rho_{yr_z}^2+1)(V_z^2(\bar{Y}-1)+2\bar{Y})\bar{Y}V_y^2+\frac{1}{8}V_y^2$$

$$(-2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}+\rho_{zr_z}^2+\rho_{zf_z}^2+\rho_{r_zf_z}^2-1)(\bar{Y}-1)\bigg)\Big/\bigg(2\Big(V_y^2((\rho_{yz}^2-1)\rho_{r_zf_z}^2$$

$$+((-2\rho_{yz}\rho_{yr_z}+2\rho_{zf_z})\rho_{zf_z}-2\rho_{yf_z}(\rho_{yz}\rho_{zr_z}-\rho_{yr_z}))\rho_{r_zf_z}+(\rho_{yr_z}^2-1)\rho_{zf_z}^2$$

$$+2\rho_{yf_z}(-\rho_{yr_z}\rho_{zr_z}+\rho_{yz})\rho_{zf_z}+(\rho_{yf_z}^2-1)\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}$$

$$-\rho_{yz}^2-\rho_{yr_z}^2-\rho_{yf_z}^2+1)+2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}-\rho_{zr_z}^2-\rho_{zf_z}^2-\rho_{r_zf_z}^2+1\Big)\bigg)$$

or

$$MSE(\hat{\hat{\bar{Y}}}_{Prop}) = \frac{(A+B+C\ \bar{Y}V_y^2(V_z^2(\bar{Y}-1)+2\bar{Y})+\frac{1}{8}V_y^2D(\bar{Y}-1)}{2V_y^2(E+F+G+H)}$$

where

$$A=(\rho_{yz}^2-1)\rho_{r_zf_z}^2+(-2\rho_{yz}\rho_{yr_z}+2\rho_{zr_z})\rho_{zf_z}-2\rho_{yf_zz}(\rho_{yz}\rho_{zr_z}-\rho_{yr_z})\rho_{r_zf_z},$$

$$B=(\rho_{yr_z}^2-1)\rho_{zf_z}^2+2\rho_{yf_z}(-\rho_{yr_z}\rho_{zr_z}+\rho_{yz})\rho_{zf_z},$$

$$C=(\rho_{yf_z}^2-1)\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}-\rho_{yz}^2-\rho_{yf_z}^2-\rho_{yr_z}^2+1,$$

$$D=(-2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}+\rho_{zr_z}^2+\rho_{zf_z}^2+\rho_{zr_z}^2-1),$$

$$E=(\rho_{yz}^2-1)\rho_{r_zf_z}^2+((-2\rho_{yz}\rho_{yr_z}+2\rho_{zf_z})\rho_{zf_z}-2\rho_{yf_z}(\rho_{yz}\rho_{zr_z}-\rho_{yr_z}))\rho_{r_zf_z},$$

$$F=(\rho_{yr_z}^2-1)\rho_{zf_z}^2+2\rho_{yf_z}(-\rho_{yr_z}\rho_{zr_z}+\rho_{yz})\rho_{zf_z},$$

$$G=(\rho_{yf_z}^2-1)\rho_{zr_z}^2+2\rho_{yz}\rho_{yr_z}\rho_{zr_z}-\rho_{yz}^2-\rho_{yr_z}^2-\rho_{yf_z}^2+1,$$

And

$$H=2\rho_{zr_z}\rho_{zf_z}\rho_{r_zf_z}-\rho_{zr_z}^2-\rho_{zf_z}^2-\rho_{r_zf_z}^2+1.$$

**5. COMPARATIVE STUDY**

In this section, multiple population data sets were utilized to conduct a numerical comparison of the proposed estimators with existing counterparts. The assessment focused on examining the performance of the estimators by analyzing their mean squared errors (MSEs) and percentage relative efficiencies (PREs).

**1. Population-I** (Source of data: Gujarati, 2009),

Y is Eggs produced in Millions in 1991 and Z is Per dozen price of eggs in 1991.

**2. Population-II** (Source of data: Koyuncu and Kadilar, 2009),

Y is Number of teachers and Z is Number of students.

3. **Population-III** (Source of data: Punjab Bureau of Statistics, 2021-2022).
4. Book Published by Bureau of Statistics Pakistan,
Y is In 2021, COVID-19 test was carried out in Punjab and Z is In 2021, COVID-19

established cases. and

**4. Population-IV** (Source of data: Murthy, 1967),

Y is Output of the Factory and Z is Number of Workers

**Table 1: Statistic of the comparative study datasets**

| Statistic↓ | Popul-I | Popul-II | Popul-III | Popul-IV |
|---|---|---|---|---|
| $N$ | 50 | 923 | 228 | 80 |
| $n$ | 5 | 180 | 40 | 10 |
| $\tau$ | 0.18 | 0.004472132 | 0.02061404 | 0.0875 |
| $\bar{Y}$ | 1357.622 | 436.4345 | 14179.06 | 5182.637 |
| $\bar{Z}$ | 78.29 | 11440.5 | 882.9342 | 285.125 |
| $\bar{R}_z$ | 25.5 | 462 | 114.5 | 40.5 |
| $\bar{F}_z$ | 0.512 | 0.5005804 | 0.5027124 | 0.5067187 |
| $C_y$ | 1.223641 | 1.718333 | 2.091506 | 0.3541939 |
| $C_z$ | 0.2722885 | 1.864528 | 3.482949 | 0.9484593 |
| $C_{r_z}$ | 0.5715933 | 0.5770377 | 0.5760841 | 0.5721408 |
| $C_{f_z}$ | 0.5717075 | 0.5769416 | 0.5748503 | 0.5737652 |
| $\rho_{yz}$ | -0.2888328 | 0.9543029 | 0.8473361 | 0.9149811 |
| $\rho_{yr_z}$ | -0.2460575 | 0.6444158 | 0.454595 | 0.9832342 |
| $\rho_{yf_z}$ | -0.2469467 | 0.6444185 | 0.4545811 | 0.9836087 |
| $\rho_{zr_z}$ | 0.9460146 | 0.6306615 | 0.3807422 | 0.8902191 |
| $\rho_{zf_z}$ | 0.9467713 | 0.6306635 | 0.3807436 | 0.8906586 |
| $V_y^2$ | 0.9687591 | 0.9999999 | 0.9999939 | 0.9999855 |
| $V_z^2$ | 0.2695136 | 0.01320472 | 0.09017398 | 0.01097716 |
| $V_{r_z}^2$ | 0.01334538 | 0.01554721 | 0.2500674 | 0.07871281 |
| $V_{f_z}^2$ | 0.0164878 | 0.001488601 | 0.006811967 | 0.0286427 |
| $V_{yz}$ | -0.01732218 | 0.01367342 | 0.1272404 | 0.02689552 |
| $V_{yr_z}$ | -0.03108973 | 0.002857539 | 0.01129101 | 0.01749064 |
| $V_{yf_z}$ | -0.01837741 | 0.002857075 | 0.01126648 | 0.01743448 |
| $V_{zr_z}$ | 0.0265237 | 0.003034478 | 0.01574806 | 0.04238946 |
| $V_{zf_z}$ | 0.01462072 | 0.003033983 | 0.01571439 | 0.04229032 |
| $V_{r_zf_z}$ | 0.03016619 | 0.001488849 | 0.006826545 | 0.0287236 |

**Table 2: Variance & MSEs of the Proposed and Considered Estimators**

| Estimators | Popul-I | Popul-III | Popul-IV | Popul-IV |
|---|---|---|---|---|
| $\bar{y}$ | 496750.7 | 2515.169 | 18129088 | 294843.69 |
| $\hat{Y}_1$ | 585202.3 | 267.6354 | 17241817 | 964235.75 |
| $\hat{Y}_2$ | 457493.7 | 10685.41 | 119566304 | 3853861.49 |
| $\hat{Y}_3$ | 455309.5 | 224.6194 | 5112794 | 48003.36 |
| $\hat{Y}_4$ | 365115.3 | 224.3549 | 4985995 | 47917.72 |
| $\hat{Y}_5$ | 362698.6 | 216.7154 | 4626187 | 7205.118 |
| $\hat{Y}_{Prop}$ | **124649.4** | **33.20287** | **1596402** | **3177.738** |

**Figure 1: Variance & MSEs of the Proposed and Considered Estimators**

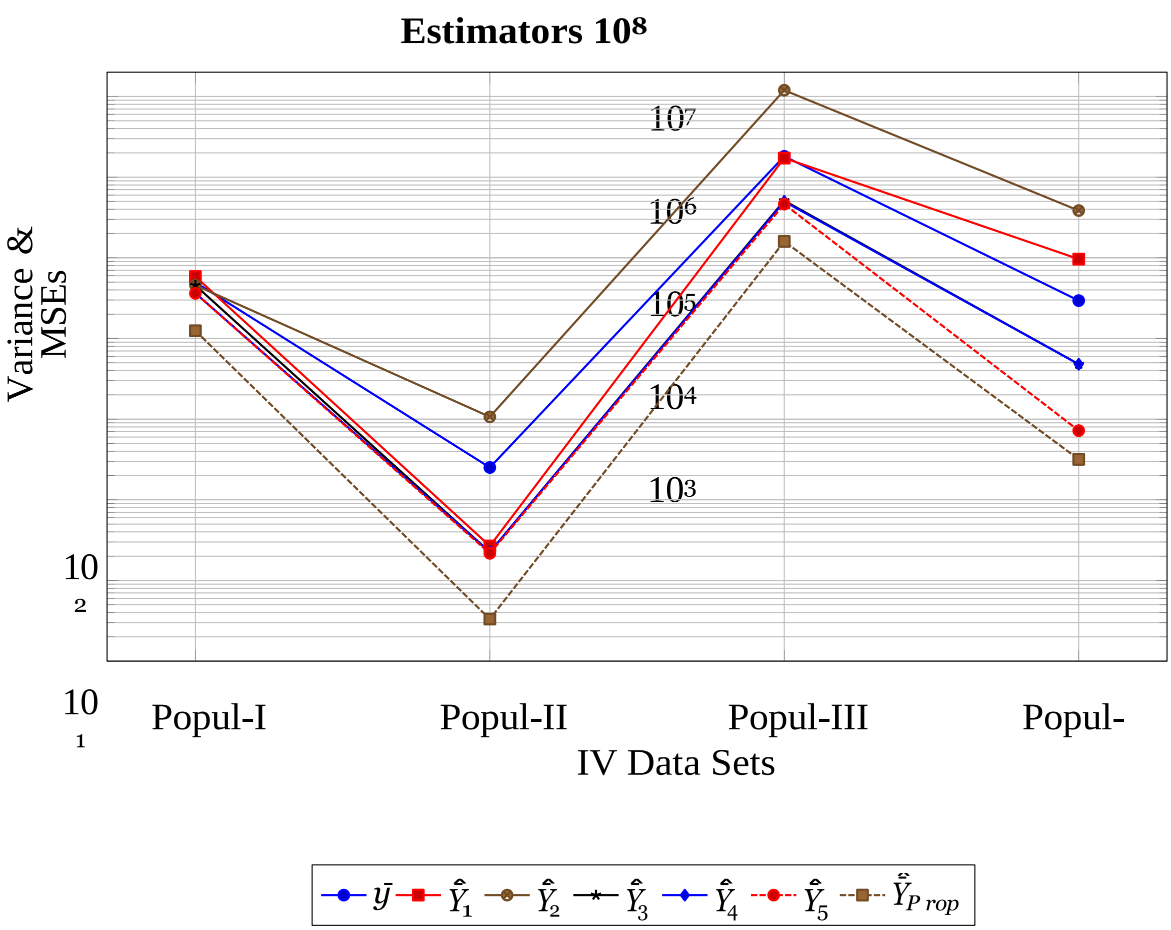
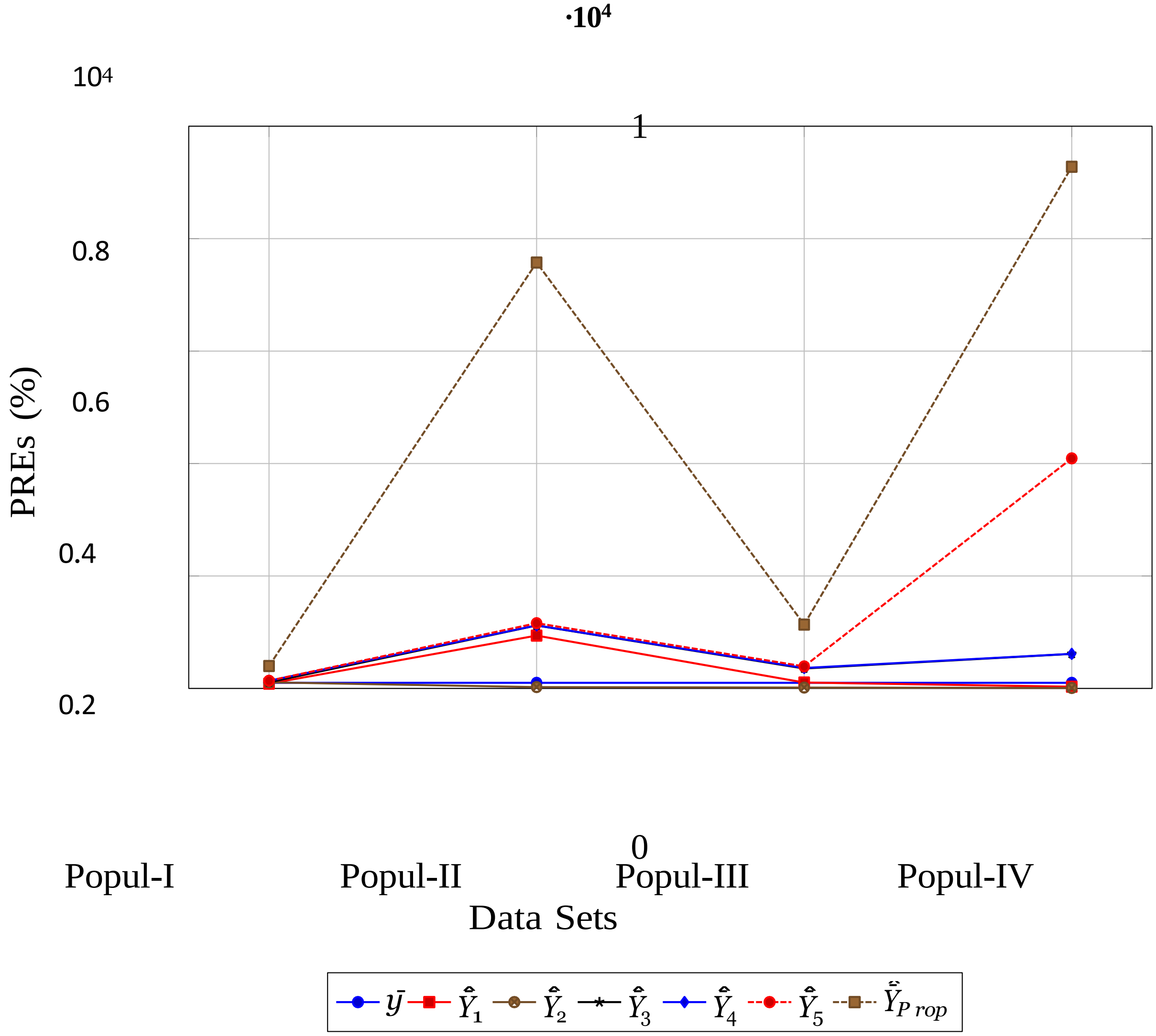


**Table 3: PREs of the Proposed and Considered Estimators**

| Estimators | Popul-I | Popul-II | Popul-III | Popul-IV |
|---|---|---|---|---|
| $\bar{y}$ | 100 | 100 | 100 | 100 |
| $\hat{\bar{Y}}_1$ | 84.88528 | 939.7742 | 105.146 | 30.57797 |
| $\hat{\bar{Y}}_2$ | 108.5809 | 23.53834 | 15.16237 | 7.650604 |
| $\hat{\bar{Y}}_3$ | 109.1017 | 1119.747 | 354.5828 | 614.2148 |
| $\hat{\bar{Y}}_4$ | 136.0531 | 1121.067 | 363.6002 | 615.3125 |
| $\hat{\bar{Y}}_5$ | 136.9596 | 1160.586 | 391.8797 | 4092.143 |
| $\hat{\bar{Y}}_{Prop}$ | **398.5183** | **7575.154** | **1135.622** | **9278.414** |

**Figure 2: PREs of the Proposed and Considered Estimators**



## 5. CONCLUSION AND DISCUSSION

### 5.1 DISCUSSION

The proposed study introduces a regression-type estimator for estimating the finite population mean by utilizing the rank and empirical distribution function (EDF) as double duals of an auxiliary variable. This approach is designed to maximize the efficiency of estimation, by incorporating both the auxiliary variable and its derived functions to reduce the mean squared error (MSE).

The results of the numerical study, as presented in Tables 3 and 2, demonstrate the superiority of the proposed estimator compared to existing estimators. The proposed estimator achieves the highest percentage relative efficiencies (PREs) and exhibits the lowest MSEs across all the examined data sets. Specifically, the PRE values for the proposed estimator outperform all counterparts, with significant improvements observed across all scenarios. These findings validate the theoretical properties of the estimator and its practical applicability in real-world scenarios where auxiliary variables have a strong correlation with the study variable.

The inclusion of the rank and EDF as double duals of the auxiliary variable intro- duces a novel dimension in survey sampling techniques. The rank function captures the relative position of the observations, while the EDF represents the cumulative distribu- tion of the auxiliary variable. By combining these aspects with the auxiliary variable in a regression-type framework, the proposed estimator leverages additional information, leading to enhanced precision.

The empirical analysis further highlights the robustness of the proposed estimator in various population settings. The variance and MSE comparisons confirm that the estimator effectively minimizes error while maintaining stability across diverse data sets. This indicates the broader applicability of the method in different domains, ranging from agriculture to socio-economic surveys.

The comparative analysis of variances, MSEs, and PREs across different population datasets highlights the superior performance of the newly proposed estimator ($\hat{Y}_{\mathrm{Prop}}$) compared to the existing estimators. From Table 4.2, it is evident that $\hat{Y}_{\mathrm{Prop}}$ consistently achieves significantly lower MSE values across all population datasets. For instance, in Population-IV, the MSE of $\hat{Y}_{\mathrm{Prop}}$ is 3177.738, which is notably lower than the MSEs of the existing estimators, such as $\hat{Y}_1$ (496750.7) and $\hat{Y}_4$ (36511.5). This reduction in MSE demonstrates the enhanced precision of the proposed estimator.

Similarly, Table 4.3 provides insight into the PREs of the estimators. The PRE values of $\hat{Y}_{\mathrm{Prop}}$ exceed those of all other estimators across every dataset. For Population- II, $\hat{Y}_{\mathrm{Prop}}$ achieves a PRE of 7575.154, significantly surpassing that of $\hat{Y}_4$ (1121.067) and $\hat{Y}_6$ (93.8577). This substantial improvement in efficiency confirms the robustness of $\hat{Y}_{\mathrm{Prop}}$ in utilizing auxiliary information more effectively than the competing estimators.

The trends observed in the MSE and PRE values suggest that the incorporation of dual auxiliary variables, such as rank and empirical distribution functions, in the proposed estimator plays a pivotal role in achieving higher accuracy and efficiency.

These findings emphasize the practical advantages of $\hat{Y}_{\mathrm{Prop}}$, particularly for applications requiring precise population mean estimation.

### 5.2 CONCLUSION

The study demonstrates that the proposed estimator, $\hat{Y}_{\mathrm{Prop}}$, outperforms existing esti- mators in terms of both mean squared error (MSE) and percentage relative efficiency (PRE). The results across various population datasets consistently reveal lower MSEs and higher PREs for $\hat{Y}_{\mathrm{Prop}}$, showcasing its superior accuracy and efficiency. The find- ings validate the effectiveness of incorporating dual auxiliary variables, such as rank and empirical distribution functions, in population mean estimation under simple random sampling. This advancement provides a valuable contribution to statistical estimation methodologies, offering practitioners a more reliable and efficient tool for finite population mean estimation.

**Recommendation**

Future research could explore extending this approach to other sampling schemes, other population parameters, and auxiliary variable structures.

**References**

[1] Bhushan, S., Pandey, A. P. (2018). Optimality of ratio type estimation methods for population mean in the presence of missing data. Communications in Statistics - Theory and Methods, **47**(11), 2576–2589.

[2] Cochran, W.G. (1940). The estimation of the yields of cereal experiments by sam- pling for the ratio of grain to total produce. *The Journal of Agricultural Science*, **30**(2), 262–275.

[3] Cochran, W.G. (1963). Sampling Techniques. Second edition, Wiley Publications in Statistics, John Wiley Sons.

[4] Cochran, W.G. (1977). *Sampling Techniques*. John Wiley & Sons, New York.

[5] Gujarati, D.N. (2009). *Basic Econometrics*. Tata McGraw-Hill Education, New York.

[6] Haq, A., Khan, M., and Hussain, Z. (2017). A new estimator of finite popula- tion mean based on the dual use of the auxiliary information. *Communications in Statistics - Theory and Methods*, **46**(9), 4425–4436.

[7] Hussain, A., Ullah, K., Cheema, S. A., Ali Khan, A., Hussain, Z. (2022). Empirical distribution function based dual use of auxiliary information for the improved estimation of finite population mean. *Concurrency and Computation: Practice and Experience*, **34**(27), e7346.

[8] Kadilar, C., and Cingi, H. (2006). Improvement in estimating the population mean in simple random sampling. *Applied Mathematics Letters*, **19**(1), 75-79.

[9] Khare, B.B., and Srivastava, S. (1997). Transformed ratio type estimators for the population mean. *Communications in Statistics-Theory and Methods*, **26**(7), 1779- 1791.

[10] Koyuncu, N., and Kadilar, C. (2009). Efficient estimators for the population mean.

*Hacettepe Journal of Mathematics and Statistics*, **38**(2), 217-225.

[11] Mak, T. K., and Kuk, A. (1993). A new method for estimating finite-population quantiles using auxiliary information. *Canadian Journal of Statistics*, **21**(1), 29-38.

[12] Murthy, M.N. (1964). Product method of estimation. *Indian Journal of Statistics, Series A*, **26**(1), 69-74.

[13] Murthy, M.N. (1967). *Sampling Theory and Methods*. Statistical Publication Soci- ety, Calcutta.

[14] Pandey, A.K., Usman, M., and Singh, G.N. (2021). Optimality of ratio and regres- sion type estimators using dual of auxiliary variable under non response. *Alexandria Engineering Journal*, **60**(5), 4461-4471.

[15] Riaz, S., Diana, G., and Shabbir, J. (2014). Improved classes of estimators for population mean in presence of non-response. *Pakistan Journal of Statistics*, **30**(1).

[16] Singh, H. P., and Solanki, R. S. (2013). An efficient class of estimators for the population mean using auxiliary information. *Communications in Statistics-Theory and Methods*, 42(1), 145-163.

[17] Srivastava, S. K. (1967). An Estimator Using Auxiliary Information in Sample Serveys. Calcutta

Statistical Association Bulletin, **16**(2-3), 121-132.

[18] Yaqub, M., Shabbir, J., and Gupta, S.N. (2017). Estimation of population mean based on dual use of auxiliary information in non-response. *Communications in Statistics-Theory and Methods*, **46**(24), 12130-12151.

[19] Zaman, T., and Kadilar, C. (2021). New class of exponential estimators for finite population mean in two-phase sampling. *Communications in Statistics-Theory and Methods*, **50**(4), 874-889.