mathematics

MDPI

Article

An Efficient Ratio-Cum-Exponential Estimator for Estimating the Population Distribution Function in the Existence of Non-Response Using an SRS Design

Ayesha Khalid [1], Aamir Sanaullah [1,*], Mohammed M. A. Almazah [2,3] and Fuad S. Al-Duais [4,5]

[1] Department of Statistics, COMSATS University Islamabad, Lahore Campus, Islamabad 45550, Pakistan
[2] Department of Mathematics, College of Science and Arts (Muhyil), King Khalid University, Muhyil 61421, Saudi
[3] Department of Mathematics and Computer, College of Sciences, Ibb University, Ibb 70270, Yemen
[4] Mathematics Department, College of Humanities and Science, Prince Sattam Bin Abudulaziz University, Al Aflaj 16278, Saudia Arabia
[5] Administrative Department, Administrative Science College, Thamar University, Thamar 87246, Yemen
* Correspondence: draamirsanaullah@cuilahore.edu.pk; Tel.: +92-3335541499

Abstract: To gain insight into various phenomena of interest, cumulative distribution functions (CDFs) can be used to analyze survey data. The purpose of this study was to present an efficient ratio-cum-exponential estimator for estimating a population CDF using auxiliary information under two scenarios of non-response. Up to first-order approximation, expressions for the bias and mean squared error (MSE) were derived. The proposed estimator was compared theoretically and empirically, with the modified estimators. The proposed estimator was found to be better than the modified estimators based on present-relative efficiency PRE and MSE criteria under the specific conditions.

Keywords: auxiliary information; exponential estimator; sub-sampling of non-respondents; cumulative distribution function; non-response

MSC: 62D05; 62G05

check for updates

Citation: Khalid, A.; Sanaullah, A.; Almazah, M.M.A.; Al-Duais, F.S. An Efficient Ratio-Cum-Exponential Estimator for Estimating the Population Distribution Function in the Existence of Non-Response Using an SRS Design. *Mathematics* **2023**, *11*, 1312. https://doi.org/10.3390/math11061312

Academic Editor: María Del Mar Rueda

Received: 22 January 2023
Revised: 19 February 2023
Accepted: 3 March 2023
Published: 8 March 2023

1. Introduction

It is a well-accepted fact in survey sampling that, under certain conditions, auxiliary information can provide precise estimates of population parameters such as the mean, median, standard deviation, totals, quantiles, and the cumulative distribution function (CDF), etc. If a linear and higher correlation is observed between the study variable Y and auxiliary variable X, researchers often use a traditional estimator for the mean, like ratio, product, and regression estimators, to estimate a population mean. The literature includes a significant amount of work for estimating different parameters of a population, for example, see [1–7]. These studies propose improved ratio-, product-, and regression-type estimators for estimating the mean and variance of a population using auxiliary variables.

Non-response is an issue that cannot be avoided in complex sample surveys, and it can be found in surveys that involve human responses. Language problems, inaccurate return addresses, a lack of information, and the sensitivity of the survey question(s), among many other reasons, can play a part in causing this issue. For example, an individual may be reluctant to provide salary information. Non-response in sample surveys is more prevalent and pervasive in postal surveys than in special canvasser surveys. Therefore, the term non-response refers to the inability to measure part of the units in a sample survey. Non-response can compromise estimator accuracy and increase its bias.

To cope with the non-response problem, several measures are proposed by various researchers, such as the weighting adjustment approach as, suggested by Oh and Scheuren [8]; imputation techniques provided by Kalton [9] and Kalton and Maligalig [10]; and the approach of sub-sampling non-respondents as recommended by Hansen and Hurwitz [11]. Various researchers have attempted to reduce the bias and to improve the effectiveness of the estimators of a population mean in the existence of non-response. Some significant references on estimating the population mean utilizing auxiliary variables in the existence of non-response include [12–17], etc.

Although there is extensive literature on estimating different estimators, it is noted that an auxiliary information-based estimation of a population CDF is less emphasized. It is becoming increasingly significant in survey sampling when statisticians are frequently interested in the proportion of a variable's domain under examination. For example, policymakers may want to know where the percentage of educated women is higher or equal to 50% in Pakistan, the proportion of individuals having a weekly income of 100 USD or more in a developing country, etc. Similarly, a psychiatrist may be interested in knowing how many children spend one or more hours with their mobile activities or what proportion of children spend one or more hours on their phones. Several studies have revealed that spending more than an hour a day on phones or smart devices has a significant relation with psychological problems among childern, such as anxiety, loneliness, and depression.

Hence, it has become necessary to estimate the finite population CDF. Therefore, Singh et al. [18], Muñoz et al. [19], Yaqub and Shabbir [20,21], Hussain et al. [22], and Hussain et al. [23] have put their efforts on estimating the population CDF using auxiliary information.

2. Sampling Design and Notations

2.1. Notations for the CDF under SRS

Consider a finite population $\mathcal{U} = \{\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3, \ldots, \mathcal{U}_N\}$ of N distinct units, let $(y_i, x_i)\,\epsilon\,\mathcal{U}_i$ be the values of research variable Y and auxiliary variable X, respectively, on the i^{th} unit. For every index t_y and t_x where $(t_y, t_x)\,\epsilon\,\mathcal{U}$, the population CDFs of Y and X are defined, respectively, by,

$$F_Y(t_Y) = \frac{\sum_{i=1}^{N} I(Y_i \leq t_y)}{N}, \qquad F_X(t_X) = \frac{\sum_{i=1}^{N} I(X_i \leq t_x)}{N},$$

where $I(.)$ is an indicator variable. It is an average of the Bernoulli distributed variable, such that

$$I(Y_i \leq t_y) = \begin{cases} 1 & \text{for } (Y_i \leq t_y) \\ 0 & \text{and } (Y_i > t_y). \end{cases}$$

Theorem 1. *In SRS, $\ell\,\hat{F}_y(t_y) = \sum_{i=1}^{\ell} I(y_i \leq t_y) = I(Y \leq t_y)$ is a hyper-geometrically distributed variable with expected mean $E(.)$ and variance $V(.)$ for Y, respectively,*

$$\begin{aligned} E\big(\hat{F}_y(t_y)\big) &= F_Y(t_y), \\ V\big(\hat{F}_y(t_y)\big) &= \frac{N-\ell}{\ell(N-1)}[F_y(t_y)(1-F_1(t_y))], \\ \mathrm{Cov}\big(\hat{F}_y(t_y), \hat{F}_x(t_x)\big) &= \frac{N-\ell}{\ell N}\left(\frac{N_{11}N_{22} - N_{12}N_{21}}{N^2}\right), \end{aligned}$$

where we have the following:
N_{11} = the number of units in the population that belong to $I(Y_i \leq t_y)$ and $I(X_i \leq t_x)$;
N_{12} = the number of units in the population that belong to $I(Y_i \leq t_y)$ and $I(X_i > t_x)$;
N_{21} = the number of units in the population with $I(Y_i > t_y)$ and $I(X_i \leq t_x)$; and
N_{22} = the number of units in the population that belong to $I(Y_i > t_y)$ and $I(X_i > t_x)$.

Theorem 1. can be proved easily along the lines of García et al. [24].

Lemma 1. *For a large sample size ℓ, the variance of $\hat{F}_y(t_y)$ is defined as*

$$V\left(\hat{F}_y(t_y)\right) = \frac{N-\ell}{\ell N}\left[F_y(t_y)\left(1 - F_y(t_y)\right)\right].$$

Let us consider that $S^2_{F_{Y(t_y)}} = \left[F_Y(t_y)(1 - F_Y(t_y))\right]$ and $S^2_{F_{X(t_x)}} = \left[F_X(t_x)(1 - F_X(t_x))\right]$ are the population variances of $I(Y \leq t_y)$ and $I(X \leq t_x)$, respectively.

Let $S_{(FY(t_y),F_{X(t_x)})} = \mathrm{Cov}\left(F_Y(t_y), F_X(t_x)\right)$ be the population covariance between $I(Y \leq t_y)$ and $I(X \leq t_x)$, then we have the following:

$C_{F_Y(t_y)} = \dfrac{S_{F_Y(t_y)}}{F_Y(t_y)}$ is the population coefficient of variation of $I(Y \leq t_y)$, and

$C^2_{F_Y(t_y)} = \dfrac{1 - F_Y(t_y)}{F_Y(t_y)}$;

$C_{F_X(t_x)} = \dfrac{S_{F_X(t_x)}}{F_X(t_x)}$ is the population coefficient of variation of $I(X \leq t_x)$, and

$C^2_{F_X(t_x)} = \dfrac{1 - F_X(t_x)}{F_X(t_x)}$;

$\rho_{(F_Y(t_y),F_X(t_x))} = \dfrac{S_{(F(t_y),F(t_x))}}{(S_{F_Y(t_y)})(S_{F_X(t_x)})}$ is the phi-coefficient of correlation between $I(Y \leq t_y)$ and $I(X \leq t_x)$.

2.2. Notation for the CDF with Non-Response under an SRS Design

Consider the case where a finite population of N units is divided into two groups: a respondent's group of N_1 units and another non-respondent's group of N_2 units, where $N = N_1 + N_2$. Consider the case where a sample of size ℓ is drawn from a target population using SRSWOR, and it is further assumed that only ℓ_1 out of ℓ units respond, while $\ell_2 = \ell - \ell_1$ units do not. Now, a sub-sample, also referred to as the 2^{nd} phase sample, of $q = \ell_2/k$ units, where $k > 1$, is taken from the group of non-respondents of size ℓ_2 for interviewing. This way of dealing with non-respondents to obtain responses from them is also referred to as the canvasser method. Hence, the total number of responses is $(\ell_1 + q)$, collected from ℓ units, and only $(\ell_2 - q)$ units are left as non-respondents who are not selected in the 2^{nd} phase sample. Following Hansen and Hurwitz [11] a population CDF in the existence of non-response can be defined as follows:

$$F^*_Y(t_y) = \mathcal{W}_1 F_{Y(1)}(t_y) + \mathcal{W}_2 F_{Y(2)}(t_y). \tag{1}$$

Similarly, let

$$F^*_X(t_x) = \mathcal{W}_1 F_{X(1)}(t_x) + \mathcal{W}_2 F_{X(2)}(t_x),$$

where $\mathcal{W}_j = \dfrac{N_j}{N}$ and $j = 1, 2$. In addition, we have the following:

$F_{Y(1)}(t_y) = \sum_{i=1}^{N_1} I(Y_i \leq t_y)/N_1$ is the population CDF of $I(Y \leq t_y)$ for the response group;
$F_{Y(2)}(t_y) = \sum_{i=1}^{N_2} I(Y_i \leq t_y)/N_2$ is the population CDF of $I(Y \leq t_y)$ for the non-response group;
$F_{X(1)}(t_x) = \sum_{i=1}^{N_1} I(X_i \leq t_x)/N_1$ is the population CDF of $I(X \leq t_x)$ for the response group;
$F_{X(2)}(t_x) = \sum_{i=1}^{N_2} I(X_i \leq t_x)/N_2$ is the population CDF of $I(X \leq t_x)$ for the non-response group.

Yaqub and Shabbir [20] briefly studied the unbiased estimator of the population CDF of the research variable when there was non-response in the sample.

Let the sample CDF $\left(\hat{F}^*_y(t_y), \hat{F}^*_x(t_x)\right)$ be the unbiased estimators of the population CDF $\left(F^*_Y(t_y), F^*_X(t_x)\right)$, based on ℓ units in the existence of non-response.

By using the Hansen and Hurwitz [11] approach, $\hat{F}_y^*(t_y)$ is defined as

$$\hat{F}_y^*(t_y) = \omega_1 \hat{F}_{y(1)}(t_y) \ + \ \omega_2 \hat{F}_{y(2\mathbf{q})}(t_y), \tag{2}$$

where $\omega_1 = \dfrac{\ell_1}{\ell}$ and $\omega_2 = \dfrac{\ell_2}{\ell}$. In addition, we have the following:
$\hat{F}_{y(1)}(t_y) = \sum_{i=1}^{\ell_1} I(Y_i \leq t_y)/\ell_1$ denotes the sample CDF based on ℓ_1 responding units out of ℓ units;
$\hat{F}_{(2q)}(t_y) = \sum_{i=1}^{q} I(Y_i \leq t_y)/q$ denotes the sample CDF based on q responding units out of ℓ_2 non-response units.

Theorem 2. *The mean and variance of $\hat{F}_y^*(t_y)$ is defined as follows:*

- *$\hat{F}_y^*(t_y)$ is an unbiased estimator of $F_Y(t_y)$, i.e., $E(\hat{F}_y^*(t_y)) = F_Y(t_y)$;*
- *The variance of $\hat{F}_y^*(t_y)$ is defined by*

$$Var\ (\hat{F}_Y^*(t_y)) \ = \ \left[\delta_1 F_{Y(1)}(t_y)\Big(1 - F_{Y(1)}(t_y)\Big) + \delta_2 F_{Y(2)}(t_y)\Big(1 - F_{Y(2)}(t_y)\Big)\right], \tag{3}$$

where $\delta_1 = \dfrac{(N-\ell)}{N\ell}$ and $\delta_2 = \dfrac{W_2(k-1)}{\ell}$. Theorem 2. can be proved along the lines of [20].

Similarly, for the supplemental variable X, the estimator $\hat{F}_x^*(t_x)$ is defined as

$$\hat{F}_x^*(t_x) = \omega_1 \hat{F}_{x(1)}(t_x) \ + \ \omega_2 \hat{F}_{x(2\mathbf{q})}(t_x).$$

In addition, we have the following:
$\hat{F}_{x(1)}(t_x) = \sum_{i=1}^{\ell_1} I(X_i \leq t_x)/\ell_1$ denotes the sample CDF based on ℓ_1 responding units out of ℓ units;
$\hat{F}_{(2q)}(t_x) = \sum_{i=1}^{q} I(X_i \leq t_x)/q$ denotes the sample CDF based on q responding units out of ℓ_2 non-response units.

Lemma 2. *On the lines of Theorem 2, the mean and variance of $\hat{F}_x^*(t_x)$ are defined as follows:*

- *$\hat{F}_x^*(t_x)$ is an unbiased estimator of $F_X(t_x)$, i.e., $E\big(\hat{F}_x^*(t_x)\big) = F_X(t_x)$;*
- *The variance of $\hat{F}_x^*(t_x)$ is defined by*

$$Var\ (\hat{F}_x^*(t_x)) \ = \ \left[\delta_1\ F_{X(1)}(t_x)\Big(1 - F_{X(1)}(t_x)\Big) + \delta_2\ F_{X(2)}(t_{x_2})\Big(1 - F_{X(2)}(t_{x_2})\Big)\right];$$

.

- *The covariance between $\hat{F}_y^*(t_y)$ and $\hat{F}_x^*(t_x)$ is given by*

$$\text{Cov}\Big(\hat{F}_y^*(t_y), \hat{F}_x^*(t_x)\Big) = \left[\delta_1\left(\frac{N_{11}N_{22} - N_{12}N_{21}}{N^2}\right) + \delta_2\left(\frac{N_{11(2)}N_{22(2)} - N_{12(2)}N_{21(2)}}{N_{(2)}^2}\right)\right].$$

(For the proof see [20]).

In addition, let us define the following:
$S^2_{F_{Y(1)}(t_y)} = F_{Y(1)}(t_y)\Big(1 - F_{Y(1)}(t_y)\Big)$ is the population variance of $I(Y \leq t_y)$ for the response group;
$S^2_{F_{Y(2)}(t_y)} = F_{Y(2)}(t_y)\Big(1 - F_{Y(2)}(t_y)\Big)$ is the population variance of $I(Y \leq t_y)$ for the non-response group;
$S^2_{F_{X(1)}(t_x)} = F_{X(1)}(t_x)\Big(1 - F_{X(1)}(t_x)\Big)$ is the population variance of $I(X \leq t_x)$ for the re-

sponse group;
$S^2_{F_{X(2)}(t_x)} = F_{X(2)}(t_x)\Big(1 - F_{X(2)}(t_x)\Big)$ is the population variance of $I(X \le t_x)$ for the non-response group;
$C_{F_{Y(1)}(t_y)} = S_{F_{Y(1)}(t_y)}/F_{Y(1)}(t_y)$ is the population coefficient of variation of $I(Y \le t_y)$ for the response group;
$C_{F_{Y(1)}(t_y)} = S_{F_{Y(1)}(t_y)}/F_{Y(1)}(t_y)$ is the population coefficient of variation of $I(Y \le t_y)$ for the response group;
$C_{F_{X(2)}(t_x)} = S_{F_{X(2)}(t_x)}/F_{X(2)}(t_x)$ is the population coefficient of variation of $I(X \le t_x)$ for the non-response group;
$C_{F_{X(2)}(t_x)} = S_{F_{X(2)}(t_x)}/F_{X(2)}(t_x)$ is the population coefficient of variation of $I(X \le t_x)$ for the non-response group;
$S_{F_{Y(1)}(t_y)F_{X(1)}(t_x)} = \mathrm{Cov}\Big(F_{Y(1)}(t_y), F_{X(1)}(t_x)\Big)$ is the population covariance between $I(Y \le t_y)$ and $I(X \le t_x)$ for the response group;
$S_{F_{Y(2)}(t_y)F_{X(2)}(t_x)} = \mathrm{Cov}\Big(F_{Y(2)}(t_y), F_{X(2)}(t_x)\Big)$ is the population covariance between $I(Y \le t_y)$ and $I(X \le t_x)$ for the non-response group.

The following relative error terms are taken into account to determine the biases and MSEs of the existing and proposed estimators:

$$e_0^* = \frac{\hat{F}_y^*(t_y) - F_Y(t_y)}{F_Y(t_y)}, \qquad e_1^* = \frac{\hat{F}_x^*(t_x) - F_X(t_x)}{F_X(t_x)}, \quad \text{and} \qquad e_2 = \frac{\hat{F}_x(t_x) - F_X(t_x)}{F_X(t_x)},$$

such that $E(e_i^*) = 0 = E(e_2)$ for $i = 0, 1$, where $E(\cdot)$ is mathematical expectation. Utilizing approximation up to the first order we have the following:

$$E(e_0^{*2}) \cong \frac{1}{F_Y^2(t_y)}\Big[\delta_1 F_{Y(1)}(t_y)\Big(1 - F_{Y(1)}(t_y)\Big) + \delta_2 F_{Y(2)}(t_y)\Big(1 - F_{Y(2)}(t_y)\Big)\Big];$$
$$E(e_0^{*2}) \cong \frac{1}{F_Y^2(t_y)}\Big[\delta_1\, S^2_{F_{Y(1)}(t_y)} + \delta_2\, S^2_{F_{Y(2)}(t_y)}\Big] \cong V^*_{200};$$
$$E(e_1^{*2}) \cong \frac{1}{F_X^2(t_x)}\Big[\delta_1 F_{X(1)}(t_x)\Big(1 - F_{X(1)}(t_x)\Big) + \delta_2 F_{X(2)}(t_x)\Big(1 - F_{X(2)}(t_x)\Big)\Big];$$
$$E(e_1^{*2}) \cong \frac{1}{F_X^2(t_x)}\Big[\delta_1\, S^2_{F_{X(1)}(t_x)} + \delta_2\, S^2_{F_{X(2)}(t_x)}\Big] \cong V^*_{020};$$

$$E(e_0^* e_1^*) \cong \frac{1}{F_Y(t_y)F_X(t_x)}\left[\delta_1\left(\frac{N_{11}N_{22} - N_{12}N_{21}}{N^2}\right) + \delta_2\left(\frac{N_{11(2)}N_{22(2)} - N_{12(2)}N_{21(2)}}{N^2_{(2)}}\right)\right];$$
$$E(e_0^* e_1^*) \cong \frac{1}{F_Y(t_y)F_X(t_x)}\Big[\delta_1\, S_{F_{Y(1)}(t_y)F_{X(1)}(t_x)} + \delta_2\, S_{F_{Y(2)}(t_y)F_{X(2)}(t_x)}\Big] \cong V^*_{110};$$

$$E(e_2^2) \cong \frac{1}{F_X^2(t_x)}[\delta_1 F_X(t_x)(1 - F_X(t_x))] \cong V_{002}; \quad \text{and}$$
$$E(e_0^* e_2) \cong \frac{1}{F_Y(t_y)F_X(t_x)}\Big[\delta_1\, S_{(1)F(t_y)F(t_x)}\Big] \cong V^{*'}_{101}.$$

There are two scenarios under consideration in the existence of non-response: **Scenario I** refers to non-response on both the study and auxiliary variables, whereas **Scenario II** solely refers to non-response on the study variable.

3. Some Modified Estimators for the CDF under Non-Response

3.1. Modified Estimators in Scenario I

In this section, some existing estimators for population mean estimation are modified for the estimation of a population CDF using SRS under Scenario I, i.e., non-response is present in both the study and the auxiliary variables. Furthermore, the biases and MSEs of the modified estimators are derived to the first order of approximation.

1. The Cochran [25] ratio estimator is modified for $F_Y^*(t_y)$, and given by

$$\hat{F}_R^*(t_y) \;=\; \left(\frac{\hat{F}_y^*(t_y)}{\hat{F}_x^{\,*}(t_x)}\right)F_X(t_x). \tag{4}$$

To the first order of approximation, the bias and MSE of Equation (4) are

$$\text{Bias}\big(\hat{F}_R^*(t_y)\big) \cong F_Y(t_y)(V_{020}^* - V_{110}^*) \quad \text{and}$$

$$\text{MSE}\Big(\hat{F}_R(t_y^*)\Big) \cong F_Y^2(t_y)(V_{200}^* + V_{020}^* - 2V_{110}^*). \tag{5}$$

2. Singh et al. [26] proposed an exponential estimator under non-response on both the study and auxiliary variables along the lines of Bahl and Tuteja [27]. The modified form of [26] for estimating the CDF is given by

$$\hat{F_{S_1}}(t_y^*) = \hat{F}_y^*(t_y)\exp\left(\frac{F_X(t_x) - \hat{F}_x^{\,*}(t_x)}{F_X(t_x) + \hat{F}_x^{\,*}(t_x)}\right). \tag{6}$$

The bias and MSEs of Equation (6) to the first order of approximation are given as

$$\text{Bias}(\hat{F_{S_1}}(t_x)) \cong F_Y(t_y)\left(\frac{3}{8}V_{020}^* - \frac{1}{2}V_{110}^*\right), \quad \text{and}$$

$$\text{MSE}\big(F_{S_1}(t_x)\big) \cong \frac{F_Y^2(t_y)}{4}(4V_{200}^* + V_{020}^* + 4V_{110}^*). \tag{7}$$

3. The modified regression estimator for $F_Y^*(t_y)$ is provided as

$$\hat{F}_{Reg}(t_y) \;=\; \hat{F_Y}^*(t_y) + \mathcal{B}^*\Big(F_X(t_x) - \hat{F_x}^*(x)\Big), \tag{8}$$

where $\mathcal{B}^*$ is the regression co-efficient. Moreover, Equation (8) is an unbiased estimator of $F_Y(t_y)$.

In addition, at the optimum value $\mathcal{B}^*_{(\text{opt})} = (S^*_{F_{Y(t_y)}F_{X(t_x)}})/(S^{*2}_{F_{X(t_x)}})$, the minimum variance of $\hat{F}^*_{Reg}(t_y)$ is given as

$$\text{MSE}_{\min}\Big(F^*_{Reg}(t_x)\Big) \;=\; F_Y^2(t_y)V_{200}^*\left(1 - \frac{V_{110}^{*2}}{V_{200}^*V_{020}^*}\right). \tag{9}$$

3.2. Modified Estimators in Scenario II

In this section, some of the existing estimators used to estimate the mean of a population, are modified for the estimation of the population CDF under Scenario II, i.e., non-response is present only on the study variable. Furthermore, the biases and MSEs of these modified estimators are obtained to their first order approximation. Let $\hat{F}'_{(.)}(t_y)$ be the estimator of the population CDF under Scenario II.

1. The Cochran [25] ratio estimator is modified for $F_Y^*(t_y)$ under Scenario II, and given as

$$\hat{F}_R'(t_x) = \left(\frac{\hat{F}_y^*(t_y)}{\hat{F}_x(t_x)}\right) F_X(t_x). \quad (10)$$

Up to the first order of approximation, the bias and MSE of Equation (10) are

$$\text{Bias}\big(\hat{F}_R'(t_y)\big) \cong F_Y(t_y)\Big(V_{020}^* - V_{101}^{*\prime}\Big), \quad \text{and}$$

$$\text{MSE}\big(\hat{F}_R'(t_y)\big) \cong F_Y^2(t_y)\Big(V_{200}^* + V_{020} - 2V_{101}^{*\prime}\Big). \quad (11)$$

2. The exponential estimator of Singh et al. [26] is modified for estimating $F_Y^*(t_y)$ and is provided as

$$\hat{F_{S_2}}'(t_y) = \hat{F}_y^*(t_y)\exp\left(\frac{F_X(t_x) - \hat{F}_x(t_x)}{F_X(t_x) + \hat{F}_x(t_x)}\right). \quad (12)$$

The bias and MSEs of Equation (12) up to the first order of approximation are given as

$$\text{Bias}(\hat{F'_{S_2}}(t_y)) \cong F_Y(t_y)\left(\frac{3}{8}V_{002} - \frac{1}{2}V_{101}^{*\prime}\right), \quad \text{and}$$

$$\text{MSE}\Big(F_{S_2}'(t_y)\Big) \cong \frac{F_Y^2(t_y)}{4}\Big(4V_{200}^* + V_{002} + 4V_{101}^{*\prime}\Big). \quad (13)$$

3. The modified regression estimator for $F_Y^*(t_y)$ in Scenario II is provided as

$$\hat{F}'_{Reg}(t_y) = \hat{F_y}^*(t_y) + \mathcal{B}'\big(F_X(t_x) - \hat{F}_x(t_x)\big), \quad (14)$$

where $\mathcal{B}'$ is said to be the regression coefficient. Moreover, Equation (14) is an unbiased estimator of $\hat{F_Y}'(t_y)$. In addition, at the optimum value $\mathcal{B}'_{(\text{opt})} = (S^*_{F_{Y(t_y)}F_{X(t_x)}})/(S^2_{F_{X(t_x)}})$, the minimum variance of $\hat{F}'_{Reg}(t_y)$ is given as

$$\text{Var}_{\min}\Big(\hat{F_{Reg}}'(t_y)\Big) = F_Y^2(t_y)V_{200}^*\left(1 - \frac{V_{101}^{*\prime 2}}{V_{200}^* V_{002}}\right), \quad (15)$$

or

$$\text{Var}_{\min}\Big(\hat{F'_{Reg}}(t_y)\Big) = F_Y^2(t_y)\,\Big(\delta_1 C^2_{F_{Y(1)(t_y)}}(1 - \rho^2_{F_y(t_y)F_x(t_y)}) + \delta_2(C^2_{F_{Y(2)}(t_y)})\Big).$$

4. Proposed Estimators

4.1. The Proposed Estimator in Scenario I

Following [17], an estimator for estimating a population CDF of the study variable under Scenario I is defined as

$$\hat{F}^*_{prop1}(t_y) = \hat{F_y}^*(t_y)\left(\frac{\hat{F}_x^*(t_x)}{F_X(t_x)}\right)^{\alpha}\exp\left(\frac{\big(F_X(t_x) - \hat{F_x}^*(t_x)\big)}{\big(F_X(t_x) + \hat{F_x}^*(t_x)\big)}\right), \quad (16)$$

where α $(-\infty < \alpha < +\infty)$ is unknown and needs to be estimated such that the MSE is minimum.

Theorem 3. *The bias and MSE of* (16) *are given, respectively, as follows:*

$$Bias\left(\hat{F}^*_{prop1}(t_y)\right) = F_Y(t_y)\left[\left(\frac{\alpha^2}{2} - \alpha + \frac{3}{8}\right)V^*_{020} + \left(\alpha - \frac{1}{2}\right)V^*_{110}\right], \quad and$$

$$MSE(\hat{F}^*_{prop1}(t_y)) = F^2_Y(t_y)\left[V^*_{200} + \left(\alpha^2 - \alpha + \frac{1}{4}\right)V^*_{020} + (2\alpha - 1)V^*_{110}\right]. \tag{17}$$

Proof. Equation (16) is expressed in error terms as

$$\hat{F}^*_{prop1}(t_y) = F_Y(t_y)(1 + e^*_0)(1 + e^*_1)^\alpha \exp\left(\frac{(-e^*_1)}{(2 + e^*_1)}\right). \tag{18}$$

Expanding Equation (18) up to the first order of approximation, yields

$$\hat{F}^*_{Prop1}(t_y) = F_Y(t_y)(1 + e^*_0)\left(1 + \alpha\, e^*_1 + \frac{\alpha\,(\alpha - 1)}{2}\, e^{*2}_1\right)\left(1 - \frac{1}{2}\, e^*_1 + \frac{3}{8}e^{*2}_1 - \cdots\right). \tag{19}$$

Keeping the terms up to the second power and extending the above equation, we get the following:

$$\hat{F}^*_{prop1}(t_y) = F_Y(t_y)\left[1 - \frac{e^*_1}{2} + \frac{3e^{*2}_1}{8} + \alpha\, e^*_1 - \frac{\alpha}{2}\, e^{*2}_1 + \frac{\alpha^2}{2}e^{*2}_1 - \frac{\alpha}{2}e^{*2}_1 + e^*_0 - \frac{e^*_0 e^*_1}{2} + \alpha e^*_0 e^*_1\right]$$

and

$$\hat{F}^*_{prop1}(t_y) - F_Y(t_y) = F_Y(t_y)\left[e^*_0 + e^*_1\left(\alpha - \frac{1}{2}\right) + \frac{3e^{*2}_1}{8} - \alpha\, e^{*2}_1 + \frac{\alpha^2}{2}e^{*2}_1 - \frac{e^*_0 e^*_1}{2} + \alpha\, e^*_0 e^*_1\right]. \tag{20}$$

After simplifying the expectation on both sides of Equation (20), we obtain the bias of (16):

$$\text{Bias}\left(\hat{F}^*_{prop1}(t_y)\right) = F_Y(t_y)\left[\left(\frac{\alpha^2}{2} - \alpha + \frac{3}{8}\right)V^*_{020} + \left(\alpha - \frac{1}{2}\right)V^*_{110}\right].$$

Squaring (20) and applying the expectation yield MSE of $\hat{F}^*_{prop1}(t_y)$ we obtain

$$\left(\hat{F}^*_{prop1}(t_y) - F_Y(t_y)\right)^2 = F^2_Y(t_y)\left(e^{*2}_0 + \alpha^2 e^{*2}_1 + \frac{1}{4}e^{*2}_1 + 2\alpha e^*_0 e^*_1 - e^*_0 e^*_1 - \alpha e^{*2}_1\right).$$

The MSE of (16) is obtained as

$$\text{MSE}(\hat{F}^*_{prop1}(t_y)) = F^2_Y(t_y)\left[V^*_{200} + \left(\alpha^2 - \alpha + \frac{1}{4}\right)V^*_{020} + (2\alpha - 1)V^*_{110}\right]. \tag{21}$$

□

Hence the theorem is proved.

Theorem 4. *The minimum MSE of* $\hat{F}^*_{prop1}(t_y)$ *is given as follows:*

$$MSE_{min}(\hat{F}^*_{prop1}(t_y)) = F^2_Y(t_y)\left[V^*_{200} - \frac{(V^*_{110})^2}{V^*_{020}}\right]. \tag{22}$$

Proof. Differentiating Equation (17) with respect to α and simplifying it to obtain the optimal value of α for minimal MSE, we get

$$\alpha_{(\text{opt})} = \frac{V^*_{020} - 2V^*_{110}}{2V^*_{020}}.$$

Substituting $\alpha_{(opt)}$ into (17) we obtain the minimal MSE of $\hat{F}^*_{prop1}(t_y)$, such that

$$\text{MSE}_{min}(F^*_{prop1}(t_y)) = F_Y^2(t_y)\left[V^*_{200} - \frac{(V^*_{110})^2}{V^*_{020}}\right]. \tag{23}$$

□

Hence the theorem is proved.

4.2. Proposed Estimator in Scenario II

Motivated by [28], to estimate the population CDF in the presence of non-response under Scenario II, an estimator is proposed as

$$F'_{prop2}(t_y) = \hat{F}_y{}^*(t_y)\left(\frac{\hat{F}_x(t_x)}{F_X(t_x)}\right)^{\alpha_1} \exp\left(\frac{(F_X(t_x) - \hat{F}_x(t_x))}{(F_X(t_x) + \hat{F}_x(t_x))}\right), \tag{24}$$

where α_1 $(-\infty < \alpha_1 < +\infty)$ is an unknown and needs to be estimated such that the MSE is minimum.

Theorem 5. *The Bias$\left(\hat{F}'_{prop2}(t_y)\right)$ and MSE$(\hat{F}'_{prop2}(t_y))$ are given, respectively, by the following:*

$$Bias\left(\hat{F}'_{prop2}(t_y)\right) = F_Y(t_y)\left[\left(\frac{\alpha_1^2}{2} - \alpha_1 + \frac{3}{8}\right)V_{002} + \left(\alpha_1 - \frac{1}{2}\right)V^{*'}_{101}\right] \quad and$$

$$MSE(\hat{F}'_{prop2}(t_y)) = F_Y^2(t_y)\left[V^*_{200} + \left(\alpha_1^2 - \alpha_1 + \frac{1}{4}\right)V_{002} + (2\alpha_1 - 1)V^{*'}_{101}\right]. \tag{25}$$

Proof. In error terms Equation (24) can be expressed as

$$\hat{F}'_{prop2}(t_y) = F_Y(t_y)(1+e_0^*)(1+e_2)^{\alpha_1} \exp\left(\frac{(-e_2)}{(2+e_2)}\right). \tag{26}$$

Expanding Equation (26) to the first order of approximation yields

$$\hat{F}'_{prop2}(t_y) = F_Y(t_y)(1+e_0^*)\left(1 + \alpha_1 e_2 + \frac{\alpha_1(\alpha_1 - 1)}{2} e_2^2\right)\left(1 - \frac{1}{2}e_2 + \frac{3}{8}e_2^2 - \cdots\right). \tag{27}$$

Keeping the terms up to the second power and extending the above equation, we get

$$\hat{F}'_{prop2}(t_y) = F_Y(t_y)\left[1 - \frac{1}{2}e_2 + \frac{3}{8}e_2^2 + \alpha_1 e_2 - \frac{\alpha_1}{2}e_2^2 + \frac{\alpha_1^2}{2}e_2^2 - \frac{\alpha_1}{2}e_2^2 + e_0^* - \frac{e_0^* e_2}{2} + \alpha_1 e_0^* e_2\right]$$

and

$$\hat{F}'_{prop2}(t_y) - F_Y(t_y) = F_Y(t_y)\left[e_0^* + \left(\alpha_1 - \frac{1}{2}\right)e_2 + \left(\frac{\alpha_1^2}{2} - \alpha_1 + \frac{3}{8}\right)e_2^2 + \left(\alpha_1 - \frac{1}{2}\right)e_0^* e_2\right]. \tag{28}$$

After simplifying the expectation on both sides of Equation (28), we obtain the bias of (24) as

$$\text{Bias}\left(\hat{F}'_{prop2}(t_y)\right) = F_Y(t_y)\left[\left(\frac{\alpha_1^2}{2} - \alpha_1 + \frac{3}{8}\right)V_{002} + \left(\alpha_1 - \frac{1}{2}\right)V^{*'}_{101}\right].$$

Squaring (28) and applying the expectation after simplification we obtain

$$E\left(\hat{F}'_{prop2}(t_y) - F_Y(t_y)\right)^2 = F_Y^2(t_y)\, E\left(e_0^{*2} + \alpha_1^2 e_2^2 + \frac{1}{4}e_2^2 + 2\alpha_1 e_0^* e_2 - e_0^* e_2 - \alpha_1 e_2^2\right). \tag{29}$$

The MSE of $\hat{F}'_{prop2}(t_y)$ is determined as

$$\text{MSE}\left(\hat{F}'_{prop2}(t_y)\right) = F_Y^2(t_y)\left[V^*_{200} + \left(\alpha_1^2 - \alpha_1 + \frac{1}{4}\right)V_{002} + (2\alpha_1 - 1)V^{*'}_{101}\right].$$

□

Hence the theorem is proved.

Theorem 6. *The Minimum MSE of $\hat{F}'_{prop2}(t_y)$ is given as*

$$MSE_{min}\left(\hat{F}'_{prop2}(t_y)\right) = F_Y^2(t_y)\left[V^*_{200} - \frac{(V^{*'}_{101})^2}{V_{002}}\right]. \quad (30)$$

Proof. Differentiating Equation (25) with respect to α_1, equating it to zero, and simplifying it to obtain the optimal value of α_1 for the minimum MSE, we get

$$\alpha_{1(opt)} = \frac{V_{002} - 2V^{*'}_{110}}{2V_{002}}. \quad (31)$$

Substituting $\alpha_{1(opt)}$ into (25) we obtain the minimum MSE of $\hat{F}'_{prop2}(t_y)$ as

$$\text{MSE}_{min}\left(\hat{F}'_{prop2}(t_y)\right) = F_Y^2(t_y)\left[V^*_{200} - \frac{(V^{*'}_{101})^2}{V_{002}}\right]. \quad (32)$$

□

Hence the theorem is proved.

5. Efficiency Comparisons

The MSEs of the modified and proposed estimators $\hat{F}^*_{prop1}(t_y)$ are compared in this section.

5.1. Efficiency Comparisons under Scenario I

The proposed estimator under Scenario I is more efficient if we have the following:

(a) From (5) and (22),

$$\text{MSE}(\hat{F}^*{}_R(t_y)) > \text{MSE}_{min}(\hat{F}^*{}_{prop1}(t_y)) \text{ if}$$
$$\text{MSE}(\hat{F}^*{}_R(t_y)) - \text{MSE}_{min}(\hat{F}^*{}_{prop1}(t_y)) > 0, \text{ or}$$
$$\frac{V^*_{110}}{V^*_{020}} < 1;$$

(b) From (7) and (22),

$$\text{MSE}(\hat{F}^*{}_{S1}(t_y)) > \text{MSE}_{min}(\hat{F}^*{}_{prop1}(t_y)) \text{ if}$$
$$\text{MSE}(\hat{F}^*{}_{S1}(t_y)) - \text{MSE}_{min}(\hat{F}^*{}_{prop1}(t_y)) > 0, \text{ or}$$
$$\frac{(V^*_{110} + V^*_{020})^2}{2V^*_{020}V^*_{110}} > -1;$$

(c) From (9) and (22), it can be shown iff $\alpha = \frac{V^*_{020} - 2V^*_{110}}{2V^*_{020}}$,

$$\text{MSE}_{min}\left(\hat{F}^*_{prop1}(t_y)\right) = \text{MSE}_{min}\left(\hat{F}^*_{Reg}(t_y)\right).$$

5.2. Efficiency Comparisons under Scenario II

The proposed estimator under Scenario II is more efficient compared to existing modified estimators if we have the following:

(a) From (11) and (30),

$$\begin{aligned} &\mathrm{MSE}(\hat{F}'_R(t_y)) > \mathrm{MSE}_{min}(\hat{F}'_{prop2}(t_y)) \quad \text{if} \\ &\mathrm{MSE}(\hat{F}'_R(t_y)) - \mathrm{MSE}_{min}(\hat{F}^*{}_{prop2}(t_y)) > 0, \text{ or} \\ &\frac{V^{*\prime}_{101}}{V_{002}} < 1; \end{aligned}$$

(b) From (13) and (30),

$$\begin{aligned} &\mathrm{MSE}(\hat{F}'_{S2}(t_y)) > \mathrm{MSE}_{min}(\hat{F}'_{prop2}(t_y)) \quad \text{if} \\ &\mathrm{MSE}(\hat{F}'_{S2}(t_y)) - \mathrm{MSE}_{min}(\hat{F}'_{prop2}(t_y)) > 0, \text{ or} \\ &\frac{(V^{*\prime}_{101} + V_{002})^2}{2V_{002}V^{*\prime}_{101}} > -1; \end{aligned}$$

(c) From (15) and (30), it can be shown iff $\alpha_1 = \frac{V_{002} - 2V^{*\prime}_{110}}{2V_{002}}$,

$$\mathrm{MSE}_{min}\left(\hat{F}'_{prop2}(t_y)\right) = \mathrm{MSE}_{min}\left(\hat{F}'_{Reg}(t_y)\right).$$

6. Numerical Study

An empirical evaluation is presented to evaluate the performance of the proposed estimators and some of the existing estimators, by using three different populations. The summary statistics for these populations are shown in Tables 1, 2 and 3 respectively:

Table 1. Summary statistics for Population I.

Parameter	Value	Parameter	Value
N	69	S_2	0.50361
n	20	S_3	7058.99
δ_1	0.03550	S^2_{44}	20.0602
$\bar{X}$	4954.44	R_{12}	0.94202
$\bar{R}$	35.0000	R_{13}	−0.57090
$F_Y(t_y)$	0.50725	R_{23}	−0.57274
$F_X(t_x)$	0.50725	R_{14}	−0.85584
S_1	0.50361	R_{24}	−0.86603
Non-response			
Parameter	**Value**	**Parameter**	**Value**
$N_{(2)}$	17	$S_{4(2)}$	5.04975
W_2	0.24638	$R_{12(2)}$	0.88273
δ_2	0.01231	$R_{13(2)}$	−0.62943
$S_{1(2)}$	0.51449	$R_{23(2)}$	−0.57830
$S_{3(2)}$	5920.86	$R_{24(2)}$	−0.85391

Population I (source: [29]).

Table 2. Summary statistics for Population II.

Parameter	Value	Parameter	Value
N	50	S_2	0.5051
n	15	S_3	21.805
λ	0.04667	S_4	15756
$\bar{R}$	25.5000	R_{13}	0.2518
$F_Y(t_y)$	0.5000	R_{23}	−0.7952
$F_X(t_x)$	0.5000	R_{14}	0.2134
S_1	0.5051	R_{24}	−0.8663

Population II (source: [30]).

Table 3. Summary statistics for Population III.

Parameter	Value	Parameter	Value
N	50	S_2	0.50508
n	15	S_3	21.3175
δ_1	0.04667	S_4	14.5756
$\bar{X}$	78.2900	R_{12}	−0.12000
$\bar{R}$	25.5000	R_{13}	0.22925
$F_Y(t_y)$	0.50000	R_{23}	−0.78936
$F_X(t_x)$	0.50000	R_{14}	0.18435
S_1	0.50508	R_{24}	−0.86630
Non-response			
Parameter	**Value**	**Parameter**	**Value**
N_2	12	$S_{4(2)}$	3.60555
w_2	0.24000	$R_{12(2)}$	−0.16903
δ_2	0.01600	$R_{13(2)}$	0.25695
$S_{1(2)}$	0.51493	$R_{23(2)}$	−0.81369
$S_{2(2)}$	0.52223	$R_{14(2)}$	0.22034
$S_{3(2)}$	18.2593	$R_{24(2)}$	−0.86905

Population III (source: [30]).

MSEs of the modified estimators and the proposed estimators are given in Tables 4 and 5, with respect to Scenario I and Scenario II, respectively.

Table 4. MSEs under Scenario I.

Estimator	Data 1	Data 2	Data 3
$\text{MSE}\hat{F}_y^*(t_y)$	0.02942125	0.03513561	0.04084997
$\text{MSE}\hat{F}_R^*(t_y)$	0.10046990	0.11875580	0.13704170
$\text{MSE}\hat{F}_{S1}^*(t_y)$	0.05427591	0.06176173	0.05427591
$\text{MSE}\hat{F}_{Reg}^*(t_y)$	0.01471663	0.01847541	0.02223419
$\text{MSE}\hat{F}_{prop1}^*(t_y)$	0.01469690	0.01840932	0.02210629

Table 5. MSEs under Scenario II.

Estimator	Data 1	Data 2	Data 3
$\text{MSE}\hat{F}_y'(t_y)$	0.02942125	0.03513561	0.04084997
$\text{MSE}\hat{F}_R'(t_y)$	0.08789827	0.09361262	0.09932698
$\text{MSE}\hat{F}_{S2}(t_y)$	0.05273304	0.05844739	0.06416175
$\text{MSE}\hat{F}_{Reg}'(t_y)$	0.01667221	0.02238657	0.02810093
$\text{MSE}\hat{F}_{prop2}'(t_y)$	0.01667221	0.02238657	0.02810093

The proposed estimators and the modified estimators of a population CDF were compared to the variance of $\hat{F}_y(t_y)$ under both scenarios in terms of their percent-relative efficiencies (PREs) by using the following formula:

$$\text{PRE}(\cdot) = \frac{\text{MSE}(\cdot)}{\text{MSE}\hat{F}_y(t_y)} \times 100\%.$$

PREs of the proposed estimators, and the modified estimators, are given in Tables 6 and 7, with respect to Scenario I and Scenario II.

Table 6. PREs of estimators under Scenario I.

Estimator	Data 1	Data 2	Data 3
$\text{PRE}\hat{F}^*_y(t_y)$	100	100	100
$\text{PRE}\hat{F}^*_R(t_y)$	29.28	29.59	29.81
$\text{PRE}\hat{F}^*_{S1}(t_y)$	54.21	56.89	75.26
$\text{PRE}\hat{F}^*_{Reg}(t_y)$	199.99	190.18	184.73
$\text{PRE}\hat{F}^*_{prop1}(t_y)$	200.19	190.86	184.79

Table 7. PREs of the estimators under Scenario II.

Estimator	Data 1	Data 2	Data 3
$\text{PRE}\hat{F}'_y(t_y)$	100	100	100
$\text{PRE}\hat{F}'_R(t_y)$	33.47	37.53	41.13
$\text{PRE}\hat{F}'_{S2}(t_y)$	55.79	60.11	63.67
$\text{PRE}\hat{F}_{Reg}(t_y)$	176.47	156.95	145.37
$\text{PRE}\hat{F}'_{prop2}(t_y)$	176.47	156.95	145.37

From Tables 6 and 7 we have the following:

- It was observed that the PREs corresponding of the estimators $\hat{F}^*_R(t_y)$, $\hat{F}^*_{S1}(t_y)$, $\hat{F}'_R(t_y)$, and $\hat{F}'_{S2}(t_y)$ declined.
- The PREs corresponding to the proposed estimators, $\hat{F}^*_{prop1}(t_y)$ and $\hat{F}'_{prop2}(t_y)$, and the modified regression estimators, $\hat{F}^*_{Reg}(t_y)$ and $\hat{F}'_{Reg}(t_y)$, showed that the proposed estimators were efficient estimators under both scenarios of non-response.

7. Conclusions

This study proposed an improved class of estimators for the estimation of a finite population CDF under two different scenarios of non-response using SRS. From theoretical and empirical comparisons, the proposed estimators were found to be perform better, based on large PRE and smaller MSE criteria. Therefore, our study suggests the use of the proposed estimators for estimating the CDF in the presence of non-response. Limitations of this study are provided in the Appendix A.

Author Contributions: Conceptualization, A.K. and A.S.; methodology, A.K. and A.S.; software, A.K.; formal analysis, A.K.; investigation, A.K. and A.S., F.S.A.-D. and M.M.A.A.; data curation, A.K.; writing—original draft preparation, A.K.; writing—review and editing, A.K. and A.S.; supervision, A.S.; funding acquisition, F.S.A.-D. and M.M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data utilized to support the numerical conclusions of this study are available within the article. The data can also be retrieved by searching the provided data sources.

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through the Larg Groups Program, under the grant number (RGP.2/23/44). This study was supported via funding from Prince Sattam bin Abdulaziz University, under project number (PSAU/2023/R/1444).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CDF	cumulative distribution function;
SRS	simple random sampling;
SRSWOR	simple random sampling without replacement;
MSE	mean square error;
PRE	percent-relative efficiency .

Appendix A. Limitations

The proposed class of estimators was designed to perform well under certain conditions. When these conditions deviated, the proposed estimator did not outperform. For example, we have the following:

1. The performance of the estimator relies on the relationship between the study variable and the auxiliary variable. If this relationship is weak the proposed estimator did not perform well.
2. The proposed estimator assumes that the underlying distribution of the study variable, as well as auxiliary variable, has a certain shape, such as exponential or normal. In situations there are large gaps in the data, the distribution is highly skewed, or risk behavior of the distribution is non-additive, the proposed estimators may not perform well.
3. The choice of an auxiliary variable can have a significant impact on the performance of the estimator. The auxiliary variable should be strongly related to the variable of interest and should be easy to measure for all units in the sample.
4. The proposed estimator is expected to perform well when the sample size is large enough to provide sufficient precision in the estimates.

References

1. Shabbir, J.; Gupta, S. Estimation of finite population mean in simple and stratified random sampling using two auxiliary variables. *Commun. Stat.-Theory Methods* **2017**, *46*, 10135–10148. [CrossRef]
2. Haq, A.; Khan, M.; Hussain, Z. A new estimator of finite population mean based on the dual use of the auxiliary information. *Commun. Stat.-Theory Methods* **2017**, *46*, 4425–4436. [CrossRef]
3. Kaur, H.; Brar, S.S.; Sharma, M. Efficient ratio type estimators of population variance through linear transformation in simple and stratified random sampling. *Int. J. Stat. Reliab. Eng.* **2018**, *4*, 144–153.
4. Ahmad, S.; Shabbir, J. Use of extreme values to estimate finite population mean under pps sampling scheme. *J. Reliab. Stat. Stud.* **2018**, *43*, 99–112.
5. Ahmad, S.; Hussain, S.; Yasmeen, U.; Aamir, M.; Shabbir, J.; El-Morshedy, M.; Al-Bossly, A.; Ahmad, Z. A simulation study: Using dual ancillary variable to estimate population mean under stratified random sampling. *PLoS ONE* **2022**, *17*, e0275875. [CrossRef] [PubMed]
6. Niaz, I.; Sanaullah, A.; Saleem, I.; Shabbir, J. An improved efficient class of estimators for the population variance. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e6620. [CrossRef]
7. Shabbir, J.; Onyango, R. Use of an efficient unbiased estimator for finite population mean. *PLoS ONE* **2022**, *17*, e0270277. [CrossRef]
8. Oh, H.; Scheuren, F. Weighting adjustments for unit non-response. *Incomplete Data Sample Surv.* **1983**, *2*, 143–184.
9. Kalton, G. Handling wave nonresponse in panel surveys. *J. Off. Stat.* **1986**, *2*, 303–314.
10. Kalton, G.; Maligalig, D.S. A comparison of methods of weighting adjustment for nonresponse. In Proceedings of the 1991 Annual Research Conference, Arlington, TX, USA, 17–20 March 1991; US Bureau of the Census: Washington, DC, USA; 1991; Volume 409428.
11. Hansen, M.H.; Hurwitz, W.N. The problem of non-response in sample surveys. *J. Am. Stat. Assoc.* **1946**, *41*, 517–529. [CrossRef]

12. Sanaullah, A.; Noor-ul Amin, M.; Hanif, M. Generalized exponential-type ratio-cum-ratio and product-cum-product estimators for population mean an the presence of non-response under stratified two-phase random sampling. *Pak. J. Stat.* **2015**, *31*, 71–94.
13. Ahmed, S.; Shabbir, J.; Gupta, S. Use of scrambled response model in estimating the finite population mean in presence of non response when coefficient of variation is known. *Commun. Stat.-Theory Methods* **2017**, *46*, 8435–8449. [CrossRef]
14. Saleem, I.; Sanaullah, A.; Hanif, M. A Generalized class of estimators for estimating population mean in the presence of non-Response. *J. Stat. Theory Appl.* **2018**, *17*, 616–626. [CrossRef]
15. Makhdum, M.; Sanaullah, A.; Hanif, M. A modified regression-cum-ratio estimator of population mean of a sensitive variable in the presence of non-response in simple random sampling. *J. Stat. Manag. Syst.* **2020**, *23*, 495–510. [CrossRef]
16. Bhushan, S.; Pandey, A.P. An efficient estimation procedure for the population Mean under non-response. *Statistica* **2020**, *79*, 363–378.
17. Ünal, C.; Kadilar, C. Exponential type estimator for the population mean in the presence of non-response. *J. Stat. Manag. Syst.* **2020**, *23*, 603–615. [CrossRef]
18. Singh, H.P.; Singh, S.; Kozak, M. A family of estimators of finite-population distribution function using auxiliary information. *Acta Appl. Math.* **2008**, *104*, 115–130. [CrossRef]
19. Muñoz, J.; Arcos, A.; Álvarez, E.; Rueda, M. New ratio and difference estimators of the finite population distribution function. *Math. Comput. Simul.* **2014**, *102*, 51–61. [CrossRef]
20. Yaqub, M.; Shabbir, J. Estimation of population distribution function in the presence of non-response. *Hacet. J. Math. Stat.* **2018**, *47*, 471–511. [CrossRef]
21. Yaqub, M.; Shabbir, J. Estimation of population distribution function involving measurement error in the presence of non response. *Commun. Stat.-Theory Methods* **2020**, *49*, 2540–2559. [CrossRef]
22. Hussain, S.; Ahmad, S.; Saleem, M.; Akhtar, S. Finite population distribution function estimation with dual use of auxiliary information under simple and stratified random sampling. *PLoS ONE* **2020**, *15*, e0239098. [CrossRef] [PubMed]
23. Hussain, S.; Akhtar, S.; El-Morshedy, M. Modified estimators of finite population distribution function based on dual use of auxiliary information under stratified random sampling. *Sci. Prog.* **2022**, *105*, 00368504221128486. [CrossRef] [PubMed]
24. García, M.R.; Cebrián, A.A.; Rodríguez, E. Quantile interval estimation in finite population using a multivariate ratio estimator. *Metrika* **1998**, *47*, 203–213. [CrossRef]
25. Cochran, W. The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *J. Agric. Sci.* **1940**, *30*, 262–275. [CrossRef]
26. Singh, R.; Kumar, M.; Chaudhary, M.K.; Smarandache, F. *Estimation of Mean in Presence of Non Response Using Exponential Estimator*; Infinite Study: Dubai, United Arab Emirates, 2009. [CrossRef]
27. Bahl, S.; Tuteja, R.K. Ratio and product type exponential estimators. *J. Inf. Optim. Sci.* **1991**, *12*, 159–164. [CrossRef]
28. Singh, G.; Usman, M. Ratio-to-product exponential-type estimators under non-response. *Jordan J. Math. Stat. (JJMS)* **2019**, *12*, 593–616.
29. Singh, S. *Advanced Sampling Theory with Applications: How Michael 'Selected' Amy, Volume 1*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2003.
30. Gujarati, D.N. *Basic Econometrics*; Tata McGraw-Hill Education: New York, NY, USA, 2009.