Applied Science, Computing and Energy, 2025, 3(3), 471-481

An Empirical and Simulation-Based Evaluation of Existing Class Estimators in Two-Occasion Successive Sampling

Charles Kelechi Ekezie, Emmanuel John Ekpenyong and David Friday Adiele
Received: 19 August 2025/Accepted: 12 October 2025/Published: 25 October 2025

Abstract*: This study presents an empirical and simulation-based comparison of four established estimators for estimating the population mean in two-occasion successive sampling. Artificial populations have been generated under varying correlation structures (strong, moderate, and weak) and different sample sizes to evaluate estimator their performances using percent relative efficiency (PRE) and the optimum replacement policy. The results reveal that estimators' efficiencies increase with increase in correlation strength and sample size. Real-data applications supported the simulation outcomes, confirming the superior and consistent performance of some estimators over others across multiple populations. Overall, no single estimator dominated across all conditions, emphasizing that the choice of estimator should depend on the expected correlation structure and sampling design.*

Keywords: *Successive sampling, Population mean, Estimators, Efficiency, Simulation study, Correlation strength*

Charles Kelechi Ekezie
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
Email: kelekezie@gmail.com
https://orcid.org/0000-0000-0000-0000

Emmanuel John Ekpenyong
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
Email: ej.ekpenyong@mouau.edu.ng
https://orcid.org/0000-0000-0000-0000

David Friday Adiele
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
Email: david.adiele@mouau.edu.ng
https://orcid.org/0000-0000-0000-0000

1.0 Introduction

Successive sampling on two occasions has gained significant attention in recent years due to its application in repeated surveys and longitudinal population studies. This method involves retaining a part of the sample from the first occasion and supplementing it with a new portion in the second occasion, allowing the exploitation of inter-occasion correlation for more efficient estimation (Cochran, 1977; Jessen, 1942).

Several researchers have proposed improved estimators to enhance efficiency in successive sampling. Singh and Pal (2017) introduced a generalized class of estimators for population mean estimation using auxiliary information. Beevi (2018) developed a modified ratio-type estimator that integrates correlation information more effectively. Later, Tiwari *et al*. (2023) and Bhushan and Pandey (2024) extended this approach to handle non-response and model-based scenarios, respectively, with improved Mean Squared Error (MSE) performance.

Recent works by Ailobhio *et al*. (2025) and Ikughur *et al.* (2024) further emphasized the need for comparative empirical evaluations of these estimators across different correlation strengths and sample size scenarios, as performance may vary considerably under changing population structures. Despite these advances, empirical evidence comparing these estimators under simulation and real-world datasets remains limited.

Thus, this study aims to fill this gap by performing a detailed simulation-based comparison of four selected existing estimators those of Singh and Pal (2017), Beevi (2018), Tiwari *et al.* (2023), and Bhushan and Pandey (2024) using different correlation strengths and sample sizes. Additionally, their performance is validated using four real-world datasets.

2. 0 Sampling Procedure and notations

Given a finite population $\Omega = (\Omega_1, \Omega_2, \cdots, \Omega_N)$ of N units, which has been sampled over two occasions, the variables under study are denoted by $x(y)$ on the first (second) occasions respectively. We will assume that the information on an auxiliary variable z (with known population mean), is available on both the occasions and is positively correlated with x and y on the first and second occasions respectively. Let a simple random sample (without replacement) of size n be drawn on the first occasion, and random sub-sample of size $m = n\lambda$ is retained (matched) from the sample selected on the first occasion for its use on the second occasion, while a fresh sample (unmatched sample) of size $u = (n-m) = n\mu$ is selected on the second occasion from the remaining population $(N-n)$ by simple random sampling (without replacement) method so that the sample size on the second occasion is also n, λ and μ are the fractions of the matched and fresh sample, respectively, at the current (second) occasion. They satisfy the following conditions, $(0 \le \mu \le 1)$, $(0 \le \lambda \le 1)$ and $(\lambda + \mu = 1)$.

$\bar{Y}$: The population mean of study variable y on the second occasion

$\bar{X}$: The population means of study variable x on the first occasion

$\bar{Z}$: The population means of auxiliary variable z

$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i$, $\bar{z}_n = \frac{1}{n}\sum_{i=1}^{n} z_i$ $\bar{x}_m = \frac{1}{m}\sum_{i=1}^{m} x_i$, $\bar{y}_m = \frac{1}{m}\sum_{i=1}^{m} y_i$, $\bar{z}_m = \frac{1}{m}\sum_{i=1}^{m} z_i$ $\bar{x}_u = \frac{1}{u}\sum_{i=1}^{u} x_i$, $\bar{y}_u = \frac{1}{u}\sum_{i=1}^{u} y_i$ and $\bar{z}_u = \frac{1}{u}\sum_{i=1}^{u} z_i$ are the sample means of study variables showed in suffices

$\rho_{yx}, \rho_{yz}, \rho_{xz}$: The correlation coefficient between the variables shown in suffices.

$S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$, S_y^2, S_z^2 are Population mean squares of x, y, z respectively.

$S_{yx} = (N-1)^{-1}\sum_{i=1}^{N}(x_i - \bar{X})(y_i - \bar{Y})$, S_{yz}, S_{xz} are Covariance between variables showed in suffices

$C_y = \frac{S_y}{\bar{Y}}$, C_x, C_z are Coefficients of variation for the variables shown in suffices.

$f_1 = \frac{n_2}{n_{2h}}$, $f = \frac{n}{N}$, $\lambda_u = \frac{1}{u} - \frac{1}{N}$, $\lambda_m = \frac{1}{m} - \frac{1}{N}$ and $\lambda_n = \frac{1}{n} - \frac{1}{N}$

When there is no auxiliary variable, the usual unbiased estimator $\hat{\bar{y}}_n = \bar{y}_n$ is used to estimate the population means (Mukhopadhyay *et al.*, 2020). The variance is given by

$$\text{var}(\hat{\bar{y}}_n) = \frac{s_y^2}{n} \tag{1}$$

Also, when there is no matching. The variance is given by

$$V(\hat{\bar{Y}}) = \left(\frac{1}{2}\right)\left[1 + \sqrt{\left(1 - \rho_{yx}^2\right)}\right]\frac{S_y^2}{n} \tag{2}$$

3.0 Review of Selected Class of Estimators under Study

i) Singh Pal (2017). Proposed an exponential method for estimating the population mean in successive sampling. The estimator is given as;

$$T_{SP} = \eta T_u + (1-\eta)T_m \tag{3}$$

where η is constant to be determined from minimum mean square error. T_u and T_m are estimators of unmatched and matched portion and they are defined as followed;

$$T_u = \left[\bar{y}_u + b_{yz(u)}(\bar{Z} - \bar{z}_u)\right]\exp\left\{\frac{\delta a\left(\bar{z}_u - \bar{Z}\right)}{a\left(\bar{z}_u - \bar{Z}\right) + 2ab}\right\}$$

$$T_m = \left[\bar{y}_m + b_{yx(m)}(\bar{x}_n - \bar{x}_m) + b_{yz(m)}(\bar{Z} - \bar{z}_u)\right]\exp\left\{\frac{\delta a\left(\bar{z}_n - \bar{Z}\right)}{a\left(\bar{z}_n + \bar{Z}\right) + 2ab}\right\}$$

where $b_{yz(u)}$ is the regression coefficient of y and z based on the sample u unmatched portion, while $b_{yx(m)}$ and $b_{yz(m)}$ are the regression coefficient of y on x and y on z respectively based on the sample m matched portion. a & b are suitably chosen scalars and δ is a scalar taking value -1 and $+1$ for generating exponential ratio type and exponential product type estimator respectively .

The optimal unmatched proportion is given as;

$$\mu_0 = \frac{\left(\alpha_3 \pm \sqrt{\alpha_1 \alpha_3}\right)}{\alpha_2} \tag{4}$$

The minimum mean square error $(MMSE)$ were derived to the first degree of approximation and given by;

$$MMSE(T_{SP}) = \frac{\alpha_3\left[(1-f)\alpha_3 - \mu_0\alpha_2 + \mu_0^2\alpha_2 f\right]}{\left[\alpha_3 - \mu_0^2\alpha_2\right]}\frac{S_y^2}{n} \tag{5}$$

where;

$$\alpha_1 = 1 - \rho_{yx}^2 - \rho_{yz}^2 + 2\rho_{yx}\rho_{yz}\rho_{xz}; \alpha_2 = \rho_{yx}^2 - 2\rho_{yx}\rho_{yz}\rho_{xz} + \delta^2\theta^2$$

$$\alpha_3 = 1 - \rho_{yz}^2 + \delta^2\theta^2; \theta = a\bar{Z}/2(a\bar{Z} + b);$$

Singh & Pal (2017) concluded that the estimator T_{SP} is more efficient than the usual the estimator $\hat{\bar{y}}_n$ and the difference type estimator $\hat{\bar{Y}}$

ii) Beevi (2005). Proposed a dual to ratio estimators for mean estimation in successive sampling using auxiliary information on two occasion. The estimator is given as;

$$T_B = \psi T_u + (1-\psi)T_m \tag{6}$$

where ψ is constant to be determined from minimum mean square error. T_u and T_m are estimators of unmatched and matched portion and they are defined as follow;

$$T_u = \bar{y}_u \frac{\bar{z}_u}{\bar{Z}}, \quad T_m = \bar{y}_m \frac{\bar{x}_m}{\bar{x}_n}\frac{\bar{z}_n}{\bar{Z}}$$

where, $\bar{z}_u = (1+g)\bar{Z} - g\bar{z}; \bar{x}_m = (1+g)\bar{X} - g\bar{x}; \bar{x}_n = (1+g)\bar{X}; g = \frac{n}{N-n}$

The optimal unmatched proportion is given as;

$$\mu_0 = \frac{-k_1 \pm \sqrt{k_1^2 + k_1 k_2}}{k_2} \quad (7)$$

The minimum mean square error $(MMSE)$ were derived to the first degree of approximation and given by;

$$MMSE(T_B) = \left[\frac{k_1^2 + \mu_0 k_1 k_2}{k_1 + \mu_0^2 k_2}\right]\frac{1}{n} \quad (8)$$

$$k_1 = 1 + g^2 - 2g\rho_{yz}; k_2 = 2g(\rho_{yx} - \rho_{yz})$$

Beevi (2018) concluded that the use of an auxiliary variable in estimating the population mean in successive sampling is justified. The proposed estimator performed better when the auxiliary information was strongly and positively correlated with the study variable. Therefore, there is a need to review these claims and examine the efficiency of the proposed estimator as well as the cost of sample replacement on the second occasion.

iii) Tiwari *et al*. (2023), proposed estimator on efficient Estimation in successive sampling over two occasion. The proposed estimators is given by

$$T_{KSS} = \phi T_m + (1-\phi)T_u \quad (9)$$

where ϕ is a constant to be determined from minimum mean square error. T_u and T_m are estimators of unmatched and matched portion and they are defined as follow;

$$T_u = \bar{y}_u\left(\frac{\bar{Z}}{\bar{z}_u}\right)\exp\alpha\left(\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\right), \quad T_m = \left[\bar{y}_m + \alpha_1(\bar{x}_n - \bar{x}_m)\right]\exp\alpha_2\left(\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_n}\right)$$

where α is real scalar, α_1 and α_2 are constants derived from MSE.

The optimal unmatched proportion is given as;

$$\theta_0 = \frac{\delta_0 \pm \sqrt{\delta_0^2(1-\rho_{zx}^2) - \delta_0 V^2}}{V^2 + \delta_0 \rho_{zx}^2} \quad (10)$$

The minimum mean square error $(MMSE)$ were derived to the first degree of approximation and given by;

$$MMSE(T_{KSS}) = \frac{S_y^2}{n}\frac{\delta_0\left[\delta_0(1-\theta_0\rho_{zx}^2) - \theta_0 V^2\right]}{\left[\delta_0(1-\theta_0\rho_{zx}^2) - \theta_0^2 V^2\right]} \quad (11)$$

where $\delta_0 = 1 - \rho_{yz}^2$, $V = \rho_{yx} - \rho_{zx}\rho_{yz}$ and θ_0 is optimal unmatched proportion (fraction of sample taken afresh).

However, Tiwari *et al*. (2023), concluded that their estimator works better than the estimators of Shabbir *et al*.(2005) and Singh and Pal (2016) in terms of efficiency gain.

iv) Bhushan and Pandey (2024). Developed an effective class of estimators for population mean estimation in successive sampling using simulation approach. The estimator is given as;

$$T_{BP} = \phi T_m + (1-\phi)T_u \quad (12)$$

where ϕ is a constant to be determined for MSE. The estimators for the unmatched and matched portion are as follows

$$T_u = \bar{y}_u + w_5(\bar{Z} - \bar{z}_u),\ T_m = w_1 \bar{y}_m + w_2(\bar{x}_n - \bar{x}_m) + w_3(\bar{z}_n - \bar{z}_m) + w_4(\bar{Z} - \bar{z}_n)$$

where $w_i, i = 1,2,3,4,5$ are suitably chosen constants

The optimal unmatched proportion is given as;

$$\theta_{opt} = \frac{1 \pm \sqrt{1 - \rho_{yx.z}^2}}{\rho_{yx.z}^2} \tag{13}$$

Also, the $MMSE$ of the estimator T_{BP} to the first degree of approximation is given by

$$MMSE(T_{BP}) = \frac{\min V^*(T_g)}{1 + \frac{\min V^*(T_g)}{\bar{Y}^2}} \tag{14}$$

where $\min V(T_g) = \frac{S_y^2}{n} \frac{(1 - \rho_{yz}^2)\left[(1 - \rho_{yz}^2) + \theta(\rho_{yz}^2 - \rho_{yx.z}^2)\right]}{\left[(1 - \rho_{yz}^2) + \theta^2(\rho_{yz}^2 - \rho_{yx.z}^2)\right]}$: $\rho_{yx.z} = \frac{(\rho_{yx} - \rho_{yz}\rho_{xz})}{\sqrt{1 - \rho_{zx}^2}\sqrt{1 - \rho_{yz}^2}}$

$$\min V^*(T_g) = \frac{1}{2n}(1 + \sqrt{1 - \rho_{yx.z}^2})S_y^2$$

Bhushan and Pandey (2024) concluded that for large population, the proposed estimator is always better for survey practitioners than other estimators proposed by Singh and Vishwakarma (2007), Singh and Pal (2016) and Bhushan *et al.* (2020), in terms of efficiency gain.

In summary, from the reviewed literature, several estimators have been proposed for two-occasion successive sampling; however, their comparative performances in terms of efficiency under varying levels of correlation and sample sizes remain unclear. Existing empirical studies are often limited to specific conditions and populations, creating uncertainty about which estimator performs best in practical situations. Consequently, survey practitioners lack clear guidance on selecting the most efficient estimator for a given survey conditions, particularly when the correlation between variables are moderate or weak. Hence, this study focuses on addressing these gaps by conducting a comprehensive comparative evaluation of selected prominent estimators under different correlation structures and sample sizes.

4. 0 Optimum Replacement Policy

The optimum unmatched proportion (μ_0) is a fraction of sample taken afresh on the second occasion. It is obtained such that the population mean $\bar{Y}$ is estimated with minimum mean square error (with Maximum precision), thereby playing the role of reducing cost of survey. The real values of μ_0 exists, when $0 \le \mu_0 \le 1$ and admissible for only positive values. On the other hand, two real values of μ_0 are obtained, but the one that lies within the interval $0 \le \mu_0 \le 1$ is chosen. If both values lie within the specified interval, the value closed to zero is selected, since the smaller value of μ_0 minimize the cost of survey.

5. 0 Efficiency Comparison

The Percent Relative Efficiency (PRE) is employed to assess the efficiency of the estimators. The PRE is expressed as:

$$E_1 = PRE(T_i, \bar{y}_n) = \frac{V(\bar{y}_n)}{MMSE(T_i)_{opt}} \times 100 \quad , \quad E_2 = PRE(T_i, \hat{\bar{Y}}) = \frac{V(\hat{\bar{Y}})}{MMSE(T_i)_{opt}} \times 100 \qquad (15)$$

Where;

$$V(\bar{y}_n) = \frac{s_y^2}{n} \; , \quad V(\hat{\bar{Y}}) = \left(\frac{1}{2}\right)\left[1 + \sqrt{\left(1 - \rho_{yx}^2\right)}\right]\frac{S_y^2}{n}$$

A PRE value greater than 100 indicates a gain in efficiency of the estimator i, while a PRE value less than 100 indicates a loss in efficiency.

6. 0 Simulation Design

In this section, a simulation study was performed using R software to generate an artificial populations taking into consideration of different correlation strength (strong, moderate and weak) and sample sizes (large, medium and small), to evaluate the performance of the estimators of Singh and Pal (2017), Beevi (2018), Tiwari *et al*. (2023) and Bhushan and Pandey (2024).

The descriptive statistics for the simulated data are summarized in Table 1. While the results obtained from the analyses are summarized and presented in Tables 2 to 4.

Table 1: Statistics for Different Correlation Strengths with Varying Sample Sizes

Correlation Strength	N	n	ρ_{xy}	ρ_{zy}	ρ_{zx}	S_y^2	$\bar{Y}$	$\bar{X}$	$\bar{Z}$
Strong Positive	1000	120,60,30	0.9887	0.9041	0.9121	0.8451	0.0374	0.0541	0.0454
Moderate Positive	1000	120,60,30	0.6695	0.5787	0.6613	7.3472	0.3362	0.3213	0.1793
Weak Positive	1000	120,60,30	0.2654	0.2888	0.2714	0.9983	0.0653	0.0541	-0.0034

7. 0 Application to Real Data Sets

To further assess the performance of the considered estimators, four real data sets were extracted from Mukhopadhyay *et al.* (2020). The characteristics of these populations are summarized in Table 5, while the results obtained are presented in Table 6.

Descriptions of the Data Sets:

(i) Population I: Wheat area in India (Sukhatme and Sukhatme, 1970);*y*: Area under wheat (1937); *x*: Area under wheat (1936); *z*: Total cultivated area (1931).

Table 2: Summary Statistics of Real Populations

Population	$\bar{Y}$	$\bar{X}$	$\bar{Z}$	N	n	ρ_{yx}	ρ_{yz}	ρ_{zx}	C_y	C_x	C_z
I	201.41	218.41	765.35	34	15	0.93	0.83	0.90	0.74	0.76	0.61
II	5182.60	5182.60	1126.50	80	30	0.91	0.99	0.94	0.35	0.94	0.75
III	76.20	68,04	68.59	34	15	0.98	0.99	0.99	0.61	0.62	0.18
IV	19.93	19.93	20.55	51	25	0.97	0.60	0.57	0.48	0.26	0.30

Source: Mukhopadhyay *et al.* (2020)

Table 3: When the Study and Auxiliary Variables are Strongly Positively Correlated

Estimators	**Population I, for** $n=120$			**Population II, for** $n=80$			**Population III, for** $n=60$		
	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
$\hat{y}_n$		100	100		100	100		100	100
T_{SP}	0.47	156.98	90.26	0.47	150.53	86.55	0.47	147.50	84.80
T_B	0.01	109.42	62.91	0.01	99.37	57.13	0.01	95.09	54.67
T_{KSS}	0.87	609.35	350.35	0.87	609.35	350.35	0.87	609.35	350.35
T_{BP}	0.74	1096.53	630.45	0.74	1348.27	775.19	0.74	1600.01	919.93

Table 4: When the Study and Auxiliary Variables are Moderately Positively Correlated

Estimators	Population I, for $n=120$			Population II, for $n=80$			Population III, for $n=60$		
	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
$\hat{y}_n$		100	100		100	100		100	100
T_{SP}	0.56	110.66	96.43	0.56	105.23	91.70	0.56	102.72	89.51
T_B	0.01	853.21	743.48	0.01	810.01	705.83	0.01	789.79	688.22
T_{KSS}	0.60	157.07	136.87	0.60	157.07	136.87	0.60	157.07	136.87
T_{BP}	0.53	198.57	173.03	0.53	225.65	196.63	0.53	252.73	220.23

Table 5: When the Study and Auxiliary Variables are Weakly Positively Correlated

Estimators	Population I, for $n=120$			Population II, for $n=80$			Population III, for $n=60$		
	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
$\hat{y}_n$		100	100		100	100		100	100
T_{SP}	0.56	90.99	89.36	0.56	86.53	84.98	0.56	82.94	82.94
T_B	**	-	-	**	-	-	**	-	-
T_{KSS}	0.51	110.20	108.23	0.51	110.20	108.23	0.51	110.20	108.23
T_{BP}	0.51	302.97	297.54	0.51	400.52	393.34	0.51	498.07	489.14

Note, '' represent Inadmissible, which implies μ_0 does not exist**

Table 6: Summary of results for real data set

Estimators	Population I			Population II			Population III			Population IV		
	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2	μ_0	E_1	E_2
$\hat{y}_n$		100	100		100	100		100	100		100	100
T_{SP}	0.50	217.70	148.86	0.43	247.38	174.97	0.42	258.02	154.68	0.64	300.63	186.86
T_B	0.23	14.57	4471	**	-	-	**	-	-	0.39	145572.5	6189.59
T_{KSS}	0.78	342.33	234.08	0.76	5087.73	3598.57	0.88	5025.13	3012.56	0.80	218.60	135.87
T_{BP}	0.60	296.98	205.57	0.53	2747.83	1943.84	0.50	14085.66	1566.56	0.77	241.02	142.00

Note, '' represent Inadmissible, which implies μ_0 does not exist**

(ii) Population II: Agricultural data from Murthy (1967);*y*: Area under wheat (1964); *x*: Area under wheat (1963); *z*: Cultivated area (1961).

(iii) Population III: Literacy rate of India (Census, 2011)

y, *x*, *z*: Literacy rates of India (2011, 2001, and female literacy 2011).

(iv) Population IV: Abortion rates in the United States (CDC, 2008–2011);

y, *x*, *z*: Abortion rates across U.S. states (2008, 2007, and 2005).

8. 0 Results and Discussion

Simulation Results

Tables 3 to 4 summarized the performances of the four estimators across different correlation strengths.

(i) Strong Correlation between the study and auxiliary variables

From Table 3, when the correlation between the study and auxiliary variables is strong, the estimator of Bhushan and Pandey (2024) consistently yields a higher percent relative efficiency (PRE) with a moderate optimum unmatched proportion, followed by Tiwari *et al*. (2023). This indicates that both estimators are highly efficient in situations where auxiliary variables are strongly associated with the study variable. The estimators of Beevi (2018) and Singh and Pal (2017) show smaller optimum unmatched proportion values and lower PREs, implying less precision under strong correlations.

(ii) Moderate Correlation between the study and auxiliary variables

Under moderate correlation (Table 4), Beevi (2018) exhibits the minimum optimum unmatched proportion with a relatively higher PRE for all sample sizes. This suggests a reduction in survey cost accompanied by higher efficiency gains. It is followed by the estimators of Bhushan and Pandey (2024), Tiwari *et al.* (2023), and Singh and Pal (2017).

(iii) Weak Correlation between the study and auxiliary variables

In Table 5, where the correlation is weak, Beevi (2018) becomes inadmissible, suggesting instability or non-existence under such conditions. Bhushan and Pandey (2024) maintains higher efficiency, followed by Tiwari *et al.* (2023), with both estimators having the same optimal replacement values. Singh and Pal (2017) perform poorly as the correlation weakens.

Real Data Results

The real-data application results (Table 6) reinforce the findings from the simulation study. The estimators of Tiwari *et al.* (2023) and Bhushan and Pandey (2024) demonstrated consistent performance across all four populations, particularly under high-correlation datasets (Populations I, II, and III), followed by Singh and Pal (2017). Beevi (2018) exhibited higher efficiency for Population IV but produced inadmissible values for Populations II and III, indicating sensitivity to the correlation structure.

9.0 Conclusion

This study provides a comprehensive empirical comparison of four prominent estimators for population mean estimation under two-occasion successive sampling. The findings reveal that:

(i) Estimators' efficiencies improve with increasing correlation strengths and sample sizes.

(ii) The Bhushan and Pandey (2024) estimator performs best under both strong and weak positive correlations, followed by the estimator of Tiwari *et al*. (2023).

(iii) Beevi (2018) exhibits robustness under moderate correlations, followed by Bhushan and Pandey (2024).

(iv) The efficiency of Tiwari *et al.* (2023) remains stable across all sample sizes, implying consistent performance.

(v) Generally, estimators' efficiencies decline as correlation weakens.
(vi) No single estimator dominates under all conditions; thus, the choice of estimator should be guided by the expected correlation structure and data characteristics.

These findings provide practical guidance for survey practitioners and researchers involved in the design and analysis of successive sampling studies.

10.0 References

Ailobhio, D. T., Ikughur, J. A., Nwoasu, S. C., & Uba, T. (2025). An improved estimation procedure for two-occasion successive sampling. *Journal of the Royal Statistical Society – Nigeria Group*, 2(1), 222–236.

Bandyopadhyay, S., & Singh, H. P. (2020). Improved ratio-type estimators in successive sampling. *Journal of Applied Statistics*, 47(3), 411–427.

Beevi, A. (2018). Improved ratio estimators for two-occasion successive sampling. *International Journal of Statistical Techniques and Applications*, 7(2), 155–166.

Bhushan S. & Pandey S. (2024). An effective class of estimators for population mean estimation in successive sampling using simulation approach. *Journal of Statistical Computation and Simulation* 94(6), 1204–1235.

Cochran, W. G. (1977). Sampling Techniques (3rd ed.). *Wiley*.

Gupta, P. C., & Singh, D. (1983). Ratio estimators using auxiliary information in successive sampling: *The Indian Journal of Statistics*, 45(B), 324–332.

Jessen, R. J. (1942). Statistical investigation of a sample survey for obtaining farm facts. *Iowa Agricultural Experiment Station Research Bulletin*, 304, 1–104.

Mukhopadhyay, P., Singh, H. P., & Bandyopadhyay, S. (2020). Use of auxiliary information in successive sampling. *Communications in Statistics – Theory and Methods*, 49(14), 3421–3437.

Patterson, H. D. (1950). Sampling on successive occasions with partial replacement of units. *Journal of the Royal Statistical Society, Series B*, 12(2), 241–255.

Rao, J. N. K., & Graham, J. E. (1964). Rotation designs for sampling on successive occasions. *Journal of the American Statistical Association*, 59(306), 492–509.

Sen, A. R. (1971). Successive sampling with partial replacement and its applications. *Sankhya, Series B*, 33(1), 1–14.

Singh, D., Chauhan, P. S., & Tracy, D. S. (1973). Successive sampling and use of auxiliary information. *Journal of the Indian Society of Agricultural Statistics*, 25, 67–83.

Singh, R., & Pal, S. (2017). A class of modified estimators in two-occasion successive sampling. *Journal of Statistics Applications & Probability*, 6(3), 479–490.

Srivastava, S. K., & Jhajj, H. S. (1981). Unbiased classes of estimators in successive sampling. *Sankhya, Series B*, 43, 249–256.

Tiwari K. K., Bhougal S. & Kumar S. (2023). On Efficient Estimation in Successive Sampling Over Two Occasions. *Thailand Statistician*, 21(2): 305-313

Declaration

Competing interests

There are no known financial competing interests to disclose

Ethical Consideration

Ethical consideration is not applicable to this study because it is a conceptua paper

Funding:

There was no external financial sponsorship for this study

Availability of data and materials:

The data supporting the findings of this study can be obtained from the corresponding author upon request

Authors' Contributions

Charles Kelechi Ekezie designed the study, developed the simulation framework, analyzed data, and drafted the manuscript. Emmanuel John Ekpenyong contributed to estimator formulation, validated results, and revised the manuscript for statistical accuracy. David Friday Adiele assisted with literature review, result interpretation, and final editing for clarity and consistency. All authors approved the final manuscript for publication.