Applied Science Innovation (ASI)
Published by College of Physical and Applied Sciences, MOUAU

- *Received:* 14 May 2026
- *Revised*: 28 May 2026
- *Accepted:* 11 June 2026
- *Published:* 16 June 2026

ON AN EXTENDED ESTIMATOR FOR POPULATION MEAN CHANGE IN SUCCESSIVE SAMPLING WITH A UXILIARY INFORMATION

Charles Kelechi Ekezie[1], Emmanuel John Ekpenyong[2], and Chinelo Okezie-Okamgba[3]

[*1]**Department of Statistics, Michael Okpara University of Agriculture, Umudike Abia State, Nigeria.**
[2]**Department of Mathematics and Statistics, Federal University of Technology, Ikot Abasi, Nigeria.**
[3]**Department of Mathematics, Michael Okpara University of Agriculture, Umudike Abia State, Nigeria.**

***Corresponding author:** kelekezie@gmail.com

ABSTRACT

This study extends the estimator proposed by Tiwari *et al.* (2023) to estimating the change in population mean under successive sampling, using auxiliary information. Expressions for the bias and mean square error (MSE) of the proposed estimator were derived, and its performance was evaluated through simulation under varying sample sizes and correlation structures (weak, moderate, and strong correlations). The results show that estimator efficiency improves with increasing correlation between the study and auxiliary variables and with larger sample sizes. Under strong correlation, the estimated change closely matched the true population mean (10.0571), with negligible bias (-0.0055 to -0.0022) and the lowest MSE values (0.0031 to 0.0011). By increasing the sample size from 150 to 350 reduced MSE by approximately 65%. The findings demonstrate that the proposed estimator provides accurate and reliable estimates of population change and is particularly effective when strong auxiliary information is available. The estimator has potential applications in repeated surveys and may be extended to more complex sampling designs and survey conditions

Keywords: Successive sampling, Population mean change, Auxiliary information, Bias, Mean square error (MSE).

1. Introduction

Successive sampling is widely used in agriculture, economics, and public health to estimate population parameters and monitor changes over time. The efficiency of estimators in this framework is often improved by incorporating auxiliary information that is highly correlated with the study variable. Foundational works

by Jessen (1942), Patterson (1950), and Cochran (1977) established the basis for successive sampling, while later studies such as Sen (1971) enhanced estimation techniques using auxiliary variables.

Recent contributions, including Tiwari *et al.* (2023) and Bhushan and Pandey (2024), have proposed improved estimators for the current population mean by combining matched and unmatched samples to minimize mean square error *(MSE).* Other studies include, Priyanka and Trisandhya (2019) and Yadav *et al.(* 2022) have developed ratio, regression, and generalized estimators to further improve efficiency under different sampling conditions. Additionally, Ekezie *et al.* (2025) compared the performance of various proposed estimators through simulation studies, while Ailobhio *et al.* (2025) developed estimation procedures for the change in successive sampling.

Despite these advances, relatively little attention has been devoted to the estimation of change in the population mean, which is one of the primary objectives of successive sampling. Estimating change is crucial because it enables researchers and policymakers to quantify how a population characteristic evolves over time. Specifically, it provides information on the magnitude and direction of change in the population average, facilitates the assessment of trends such as growth, decline, or stability, and supports the evaluation of policies, programs, or interventions by comparing population means across successive occasions. In addition, reliable estimates of mean change provide valuable information for planning and decision-making and can serve as a basis for predicting future behavior of the population characteristic under study.

Motivated by this gap in the literature, the present study extends the estimator proposed by Tiwari *et al.* (2023) for the estimation of change in the population mean under successive sampling. The bias and mean square error *(MSE)* of the proposed estimator are derived analytically, and its performance is evaluated through simulation studies under varying sample sizes and correlation structures. The results provide further insights into efficient estimation of population change and contribute to literature on successive sampling.

2 Methodology

2.1 Review of Tiwari *et al.* (2023) Estimator

Tiwari *et al*. (2023) proposed an estimator T_{KSS} for estimating the population mean $\left(\bar{Y}\right)$of the study variable y on current (second) occasion for successive sampling, using auxiliary variable z. The proposed estimator is given by

$$T_{KSS} = \phi_{KSS} T_{KSSm} + (1-\phi_{KSS}) T_{KSSu} \tag{1}$$

where ϕ_{KSS} is a constant determined from minimum mean square error, T_{KSSm} and T_{KSSu} are estimators of matched and unmatched portion of the sample.

The estimator for unmatched portion of the sample is given by

$$T_{KSSu} = \bar{y}_u \left(\frac{\bar{Z}}{\bar{z}_u}\right) \exp\alpha\left(\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right) \tag{2}$$

where α is real scalar

The bias and *MMSE* of the unmatched estimators is given as

$$B(T_{KSSu}) = \frac{\bar{Y}\lambda_u C_y^2}{4}\left(\rho_{yz}^2 - 1\right) \tag{3}$$

and

$$MMSE(T_{KSSu}) = \bar{Y}^2 \lambda_u C_y^2 \left(1 - \rho_{yz}^2\right) \tag{4}$$

The estimator for matched portion of the sample is given by

$$T_{KSSm} = \left[\bar{y}_m + \alpha_1\left(\bar{x}_n - \bar{x}_m\right)\right]\exp\alpha_2\left(\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_n}\right) \tag{5}$$

where α_1 and α_2 are constants chosen to minimize the MSE of T_{KSSm}.

The bias and *MMSE* of the matched estimators is given as

$$B(T_{KSSm}) = \bar{Y}\left[\frac{\alpha_2^2}{8}\lambda_m C_z^2 + \frac{\alpha_2}{4}\lambda_n C_z^2 - \frac{\alpha_2}{2}\lambda_m C_{yz} - \frac{\alpha_1\alpha_2}{2}\lambda_n C_{xz} + \frac{p\alpha_1\alpha_2}{2}\lambda_m C_{xz}\right] \tag{6}$$

$$\alpha_{1opt} = \frac{\lambda_m C_y\left(\rho_{yx} - \rho_{zx}\rho_{yz}\right)}{pC_x\left[\lambda_m - \left(\lambda_m - \lambda_n\right)\rho_{zx}^2\right]}$$

$$\alpha_{2opt} = \frac{2C_y\left[\lambda_m\rho_{yz} - \left(\lambda_m - \lambda_n\right)\rho_{zx}\rho_{yx}\right]}{C_z\left[\lambda_m - \left(\lambda_m - \lambda_n\right)\rho_{zx}^2\right]}$$

$$p = \frac{\bar{X}}{\bar{Y}}, \textbf{and}$$

$$MMSE(T_{KSSm}) = \bar{Y}^2\lambda_m C_y^2\left[1 - \rho_{yz}^2 - \frac{\left(\lambda_m - \lambda_n\right)\left(\rho_{yx} - \rho_{xz}\rho_{yz}\right)^2}{\lambda_m - \rho_{xz}^2\left(\lambda_m - \lambda_n\right)}\right] \tag{7}$$

The *MMSE* of the combined estimators of T_{KSSm} and T_{KSSu} are given by

$$MMSE(T_{KSS}) = \frac{S_y^2}{n}\frac{\delta_0\left[\delta_0(1-\theta_0\rho_{zx}^2) - \theta_0 V^2\right]}{\left[\delta_0(1-\theta_0\rho_{zx}^2) - \theta_0^2 V^2\right]} \tag{8}$$

where $\delta_0 = 1 - \rho_{yz}^2$, $V = \rho_{yx} - \rho_{zx}\rho_{yz}$ and θ_0 is the optimal fraction of sample taken afresh, and it is given by

$$\theta_0 = \frac{\delta_0 \pm \sqrt{\delta_0^2(1-\rho_{zx}^2) - \delta_0 V^2}}{V^2 + \delta_0\rho_{zx}^2} \tag{9}$$

$\rho_{yx}, \rho_{yz}, \rho_{xz}$, are the correlation coefficient between the variables shown in suffices, n is the sample size S_y^2 is the sample variance.

3. Results and Discussion

3.1 Bias of Change in Population Mean

This is one of the objectives of repetitive survey, especially in successive sampling. The change in mean is given as

$$\begin{aligned}\Delta_{T_{KSS}} &= T_{KSS} - \bar{y}_{1m} \\ &= \phi_{KSS} T_{KSSm} + (1-\phi_{KSS}) T_{KSSu} - \bar{y}_{1m}\end{aligned} \tag{10}$$

T_{KSSm} and T_{KSSu} is defined in Eqns. (2) and (5) above

$\bar{y}_{1m} = \frac{1}{m}\sum_{i=1}^{m} y_i$ mean of matched portion of the sample on the first occasion

3.2 Bias of Change in Population Mean

he bias of change in population mean is

$$B\left(\Delta_{T_{KSS}}\right) = \phi_{KSS} B\left(T_{KSSm}\right) + (1-\phi_{KSS}) B\left(T_{KSSu}\right) - B\left(\bar{y}_{1m}\right) \tag{11}$$

For simple random sampling, $B\left(\bar{y}_{1m}\right) = 0$, hence, the bias of the estimator become

$$B\left(\Delta_{T_{KSS}}\right) = \phi_{KSS} B\left(T_{KSSm}\right) + (1-\phi_{KSS}) B\left(T_{KSSu}\right) \tag{12}$$

Where the bias of T_{KSSm} and T_{KSSu} is defined Eqns. (3) and (6). Thus, the Eqn. (12) is now rewriting as

$$B\left(\Delta_{T_{KSS}}\right) = \left[\begin{array}{l} \bar{Y}\left(\frac{\alpha_2^2}{8}\lambda_m C_z^2 + \frac{\alpha_2}{4}\lambda_n C_z^2 - \frac{\alpha_2}{2}\lambda_m C_{yz} - \frac{\alpha_1\alpha_2}{2}\lambda_n C_{xz} + \frac{p\alpha_1\alpha_2}{2}\lambda_m C_{xz}\right)\phi_{KSS} \\ + \frac{\bar{Y}\lambda_u C_y^2}{4}\left(\rho_{yz}^2 - 1\right)(1-\phi_{KSS}) \end{array}\right] \tag{13}$$

3.3 Mean Square Error of Estimation of Change in Population Mean

By definition, the mean square error of the estimation of change in population mean is given as

$$MSE\left(\Delta_{T_{KSS}}\right) = \phi_{KSS}^2 MSE\left(T_{KSSm}\right) + (1-\phi_{KSS})^2 MSE\left(T_{KSSu}\right) - Var\left(\bar{y}_{1m}\right) - 2(1-\phi)Cov\left(T_{KSSm}, \bar{y}_{1m}\right) \tag{14}$$

Tiwari *et al.* (2023) and Bhushan and Pandey (2024) assumed that T_{KSSm}, T_{KSSu} and $\bar{y}_{1m}$ are independent and uncorrelated, and covariance between them is negligible for large sample sizes. Hence Eqn. (14) is reduced as

$$MSE\left(\Delta_{T_{KSS}}\right) = \phi_{KSS}^2 MSE\left(T_{KSSm}\right) + (1-\phi_{KSS})^2 MSE\left(T_{KSSu}\right) - Var\left(\bar{y}_{1m}\right) \tag{15}$$

Where $Var\left(\bar{y}_{1m}\right)=\frac{S_y^2}{m}$ is the variance of the matched portion of the sample mean on the first occasion.

The *MMSE* of T_{KSSm} and T_{KSSu} is defined Eqns. (4) and (8). Thus, the Eqn. (15) is now rewriting as

$$MMSE\left(\Delta_{T_{KSS}}\right)=\left\{\begin{array}{l}\phi_{KSS}^2\bar{Y}^2\lambda_m C_y^2\left[1-\rho_{yz}^2-\frac{\left(\lambda_m-\lambda_n\right)\left(\rho_{yx}-\rho_{xz}\rho_{yz}\right)^2}{\lambda_m-\rho_{xz}^2\left(\lambda_m-\lambda_n\right)}\right] \\ +(1-\phi_{KSS})^2\bar{Y}^2\lambda_u C_y^2\left(1-\rho_{yz}^2\right)-\frac{S_y^2}{m}\end{array}\right\} \quad (16)$$

$$\phi_{KSS}=\frac{MSE\left(T_{KSSu}\right)}{MSE\left(T_{KSSu}\right)+MSE\left(T_{KSSm}\right)} \quad (17)$$

3.4 Simulation Study

A simulation study was conducted to evaluate the performance of the proposed estimator under varying correlation structures and sample sizes using the R statistical software. Three artificial populations representing weak, moderate, and strong positive correlations between the study variables and the auxiliary variable were generated from a multivariate normal distribution. The correlation coefficients were set at $\rho=0.2$ for weak, $\rho=0.6$ for moderate and $\rho=0.9$ for strong positive correlation respectively, and the populations were generated using the mvrnorm function from the MASS package.

For each correlation scenario, a finite population of size $N=1000$ was created. Samples of sizes $n=250,$ and $n=350,$ were then selected using simple random sampling without replacement (SRSWOR). In the successive sampling framework, each sample was partitioned into matched and unmatched components, and the corresponding estimators for both portions were computed according to the proposed methodology. The sampling and estimation procedure was repeated a large number of times to ensure the stability and reliability of the simulation results. For each replication, the estimate of the change in population mean was obtained, and its performance was assessed using measures such as bias and mean square error (MSE). The functional relationships adopted for generating the study variables under the different correlation scenarios are presented in Table 3.1, while the summary statistics computed from the simulated results, including the bias and MSE, are reported in Tables 3.2 and 3.3.

Table 3.1: Population Generation Functions across Correlation Scenarios in the Simulation Study

variable	Strong +ve Correkation	Moderate Positive Correlation	Weak Positive Correlation
Auxiliary (z)	$Z = 0.2 * X + 0.2 * Y + rnorm(\mu = 0, \sigma = 1)$	$Z = 0.5 * X + sqrt(1 - 0.5\text{^}2) * rnorm(\mu = 0.3, \sigma = 2)$	$Z = 0.2 * X + sqrt(1 - 0.2\text{^}2) * rnorm(\mu = 0, \sigma = 1)$
First occasion (x)	$X = rnorm(\mu = 10, \sigma = 5)$	$X = sqrt(1 - 0.4) * rnorm(\mu = 0.3, \sigma = 2)$	$X = rnorm(\mu = 0, \sigma = 1)$
Second occasion (y)	$Y = X + rnorm(\mu = 0, \sigma = 1)$	$Y = 0.9 * X + 0.6 * Z + rnorm(\mu = 0, \sigma = 2)$	$Y = 0.2 * X + 0.2 * Z + sqrt(1 - 2 * 0.2\text{^}2) * rnorm(\mu = 0, \sigma = 1)$

Table 3.2: Statistics for varying correlation structures and sample sizes

Values	Pop. I (Weak)			Population II (Moderat)			Population III (Strong)		
	n=150	n=250	n=350	n=150	n=250	n=350	n=150	n=250	n=350
$Var(\bar{y}_{1m})$	1.1107	1.0348	0.9974	1.1662	1.2202	1.1908	25.5034	28.6377	26.4076
$\bar{y}_m$	-0.2970	-0.2769	-0.1533	0.1197	0.1693	0.0181	10.2684	9.5146	9.9451
$\bar{y}_u$	0.1056	0.0965	-0.0344	-0.0635	0.0602	0.0734	10.6235	10.0243	9.9906
$\bar{x}_m$	0.1264	-0.0550	0.0594	0.0916	0.0827	0.0302	10.2688	9.5691	9.9768
$\bar{x}_u$	0.0509	0.0955	0.0957	-0.1358	0.0003	0.0037	10.7072	9.9116	10.0755
$\bar{z}_n$	0.0733	-0.0335	0.1076	0.0786	0.0843	-0.0225	4.1153	3.7890	3.9647
$\bar{z}_u$	0.1118	-0.0587	-0.1154	-0.1565	0.0397	0.0921	4.4458	3.9234	4.0048
$\bar{Y}$	-0.0653	-0.0653	-0.0653	0.0478	0.0478	0.0478	10.0571	10.0571	10.0571
$\bar{X}$	0.0541	0.0541	0.0541	0.0147	0.0147	0.0147	10.0457	10.0457	10.0457
$\bar{Z}$	0.0034	0.0034	0.0034	0.0360	0.0360	0.0360	4.0063	4.0063	4.0063
S_y^2	0.9983	0.9983	0.9983	1.2944	1.2944	1.2944	25.8300	25.8300	25.8300
S_x^2	1.0165	1.0165	1.0165	1.0355	1.0355	1.0355	24.9569	24.9569	24.9569
S_z^2	0.9660	0.9660	0.9660	1.0699	1.0699	1.0699	5.0321	5.0321	5.0321
ρ_{xy}	0.2654	0.2654	0.2654	0.6426	0.6426	0.6426	0.9798	0.9798	0.9798
ρ_{zy}	0.2888	0.2888	0.2888	0.6429	0.6429	0.6429	0.8882	0.8882	0.8882
ρ_{zx}	0.2714	0.2714	0.2714	0.5175	0.5175	0.5175	0.8892	0.8892	0.8892

Population III, strong positive correlation with varying sample sizes $n = 350,\ 250,\ 150$

Population II, moderate positive correlation with varying sample sizes $n = 350,\ 250,\ 150$

Population I, weak positive correlation with varying sample sizes $n = 350,\ 250,\ 150$

Table 3.3: Estimate of change in mean in two successive sampling

correlation structures	N	n	ϕ_{KSS}	T_{KSS}	$\bar{y}_{1m}$	$\Delta_{T_{KSS}}$	$B(\Delta_{T_{KSS}})$	$MMSE(\Delta_{T_{KSS}})$
Weak	1000	150	0.6818	-0.3435	-0.2520	-0.0915	0.0692	0.0055
		250	0.6181	-0.4132	-0.2330	-0.1802	0.0383	0.0031
		350	0.5914	-0.1614	-0.1773	0.0159	0.0254	0.0021
Moderate	1000	150	0.6936	0.0138	0.1388	-0.1250	-0.1266	0.0044
		250	0.6353	0.0339	0.1135	-0.0797	-0.0674	0.0025
		350	0.6121	0.0739	0.0600	0.0139	0.0431	0.0017
Strong	1000	150	0.6987	9.8822	10.0207	-0.1385	-0.0055	0.0031
		250	0.6456	10.1301	10.0612	0.0689	-0.0032	0.0017
		350	0.6275	10.0257	9.9851	0.0406	-0.0022	0.0011

3.5 Discussion of Results

The results presented in Table 3.3. show that the performance of the proposed estimator improves as the correlation between the study and auxiliary variables increases. Estimation accuracy is reflected in lower bias and mean square error (MSE), while larger sample sizes further enhance precision.

Under weak correlation, the estimated changes are relatively unstable, with larger bias and higher MSE, indicating limited efficiency in capturing population change. For moderate correlation, the estimates become consistent and closer to the true population values, with noticeable reductions in bias and MSE as sample size increases. The best performance is observed under strong correlation, where the estimated changes closely match the true values, bias is negligible, and MSE is minimized, demonstrating the high efficiency of the proposed estimator.

These findings confirm the importance of auxiliary information in successive sampling. As the strength of the relationship between the study and auxiliary variables increases, the estimator is better to exploit this information, resulting in substantial gains in accuracy and precision. The results are consistent with the findings of Tiwari *et al.* (2023), Bhushan and Pandey (2024) and Ailobhio *et al.* (2025), all of whom reported improved estimation efficiency when strong auxiliary information is available.

From a practical perspective, the estimator appears robust across different sample sizes and correlation structures. However, in real survey applications, factors such as non-response and measurement errors may

affect performance. Therefore, practitioners should select auxiliary variables that are strongly related to the study variable, maintain adequate sample sizes, and implement effective data-quality control measures to minimize survey errors.

4. Conclusion

This study extends the estimator proposed by Tiwari *et al.* (2023) to estimating the change in population mean under successive sampling. Unlike previous studies that primarily focus on estimating the current population mean, the proposed approach explicitly addresses the estimation of population change between successive occasions. Expressions for the bias and mean square error were derived, and the estimator's performance was evaluated through simulation under varying correlation structures and sample sizes.

The results show that estimation accuracy improves with increasing correlation between the study and auxiliary variables and with larger sample sizes. The estimator performs optimally under strong correlation, producing negligible bias and the lowest MSE. These findings demonstrate that the proposed estimator provides reliable and efficient estimates of population change when suitable auxiliary information is available.

Although the study was conducted using simulated populations, the results suggest that the estimator has considerable potential for practical applications in repeated surveys. Future research may extend the methodology to more complex sampling designs, investigate its performance under non-response and measurement errors, and explore the use of multiple auxiliary variables to further improve estimation efficiency.

Acknowledgments

The authors sincerely thank the reviewers for their insightful comments, constructive suggestions, and positive feedback, which greatly contributed to improving the quality and clarity of this manuscript.

References

Ailobhio, D. T., Ikughur, J. A., Nwoasu, S. C., & Uba, T. (2025). Estimation of change in successive sampling. *Journal of Statistical Sciences and Computational Intelligence, 1*(3), 241–247. https://doi.org/10.64497/jssci.55

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. https://doi.org/10.1080/00949655.2023.2282741

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

Ekezie, C. K., Ekpenyong, E. J., & Adiele, D. F. (2025). An empirical and simulation-based evaluation of existing class estimators in two-occasion successive sampling. *Journal of Applied Science, Computing and Energy, 3*(3), 471–481.

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

Patterson, H. D. (1950). Sampling on successive occasions with partial replacement of units. *Journal of the Royal Statistical Society: Series B (Methodological, 12*(2), 241–255. https://doi.org/10.1111/j.2517-6161.1950.tb00084.

Priyanka, K., & Trisandhya, P. (2019). Some classes of estimators for sensitive population mean on successive moves. *J Stat Pract, 13*(9), 1–26. https://doi.org/10.1007/s42519-018-0008-5

Sen, A. R. (1971). Successive sampling with partial replacement and its applications. *Sankhyā: The Indian Journal of Statistics, Series B, 33*(1), 1–14.

Tiwari, K. K., Bhougal, S., & Kumar, S. (2023). On efficient estimation in successive sampling over two occasions. *The Thailand Statistician, 21*(2), 305–313. https://ph02.tci-thaijo.org/index.php/thaistat/article/view/249003

Yadav, A., & Shukla, D. (2021). A class of ratio-type estimators using first occasion data in successive sampling. *Communications in Statistics - Simulation and Computation, 50*(4), 1072–1088. https://doi.org/10.1080/03610918.2019.1686000