Journal of New Theory

ISSN: 2149-1402

44 (2023) 79-86
Journal of New Theory
https://dergipark.org.tr/en/pub/jnt
Open Access

A Novel Sub-Type Mean Estimator for Ranked Set Sampling with Dual Auxiliary Variables

Eda Gizem Koçyiğit[1]

Article Info
Received: 18 Aug 2023
Accepted: 22 Sep 2023
Published: 30 Sep 2023
doi:10.53570/jnt.1346020
Research Article

Abstract − This research introduces a novel sub-estimator designed to estimate the population mean under ranked set sampling, motivated by the new concept of a recently introduced sub-ratio estimator. The mathematical formulas of the proposed estimator's mean square error and bias are presented and theoretically contrasted with an analogous estimator found in the existing best sub-estimator literature. In addition to the theoretical analysis, empirical evidence is provided to validate the superiority of the proposed estimator. This empirical validation is based on numerical computations using Monte Carlo simulations, encompassing synthetic and real data applications. The results underscore the effectiveness of the proposed estimator. Finally, this study discusses the need for further research.

Keywords *Efficiency, Monte Carlo simulation, ranked set sampling, sub-estimator, survey sampling*

Mathematics Subject Classification (2020) 62D05, 94A17

1. Introduction

Ranked set sampling (RSS) represents a recommended option to simple random sampling (SRS), recognized for its capacity to yield more cost-efficient, time-saving, and efficient outcomes in comparison to SRS [1-3]. The RSS technique hinges on the availability of an auxiliary variable (z), ideally easily accessible and correlated with the study variable (g). This auxiliary variable is pivotal in the ranked process and sample selection procedure. However, once the sample selection is finalized, the estimation process exclusively pertains to the study variable. While it is feasible to incorporate the auxiliary variable into the estimation phase through various estimator types (e.g., ratio, product, or regression type estimators using auxiliary variables) [4-13], such approaches often necessitate knowledge of population parameters. The undeniable impact of utilizing auxiliary variable information in enhancing efficiency is well established. It is widely recognized that employing multiple auxiliary variables has the potential to yield even greater efficiency gains than using a single auxiliary variable [14-17]. However, obtaining these parameters for all variables is frequently impractical, limiting the applicability of such estimators on applications.

A novel method has been proposed, displaying that the auxiliary variable, a crucial component of the RSS method, can be effectively employed in the estimation phase without requiring population parameters. This breakthrough not only broadens the scope of utility for such estimators but also elevates the overall efficiency of the estimation process. These estimators developed for the RSS method are called sub-estimators [18]. The primary objective of this study is to introduce an estimator that surpasses the existing population mean estimators found in the literature in terms of efficiency. For this purpose, a novel sub-estimator using two

[1]eda.kocyigit@deu.edu.tr (Corresponding Author)
[1]Department of Statistics, Faculty of Science, Dokuz Eylül University, İzmir, Türkiye

auxiliary variables without their population parameters is proposed to estimate the population mean. Once the theoretical foundations of the proposed estimator have been established, the subsequent aim is to validate its efficiency through numerical investigations.

2. Estimators in Literature

In the RSS method, a simple random sample of m^2 of the observations z is initially drawn and subsequently divided into m sets. Within each set, a ranking is established based on z, leading to the selection of measurements for the g corresponding to the units situated along the diagonal. If the required sample size cannot be achieved with m, the process is repeated c times, and the desired $n = mc$ is obtained. The fundamental mean estimator of the RSS method and its mean square error (MSE) obtained after the cycle c:

$$\hat{g}_{RSS} = \frac{1}{mc}\sum_{j=1}^{c}\sum_{i=1}^{m} g_{[i,i];j}$$

and

$$\text{MSE}(\hat{g}_{RSS}) = \overline{G}^{-2}\left(\gamma C_g^2 - \omega_g^2\right)$$

where $g_{[i];j}$ is the observation value g of i^{th} ranked in i^{th} set and j^{th} cycle, $\gamma = 1/mc$, C_g^2 is the coefficient of variation of g, $\omega_{g(i)}^2 = \frac{1}{m^2 c\overline{G}^2}\sum_{i=1}^{m}\tau_{g(i)}^2$, $\tau_{g(i)}^2 = \mu_{g(i)} - \overline{G}$, and $\mu_{g(i)}$ represents the mean of the i^{th} order statistics of g. The RSS sub-ratio estimator and its MSE are described as the following equations:

$$\hat{g}_{KK1} = \frac{\overline{g}_{RSS}}{\overline{z}_{RSS}}\overline{Z}_{SUB}$$

and

$$\text{MSE}(\hat{g}_{KK1}) = \overline{G}^{-2}\left[\gamma C_{z_{SUB}}^2 - \omega_{z_{SUB}(i)}^2 + \gamma C_g^2 - \omega_{g(i)}^2 - 2\left(\gamma C_{gz_{SUB}} - \omega_{gz_{SUB}(i)}\right)\right] \tag{1}$$

where $\overline{z}_{RSS} = \frac{1}{mc}\sum_{j=1}^{c}\sum_{i=1}^{m} z_{[i,i];j}$ is the ranked set sample mean of z,

$$\overline{Z}_{SUB} = \frac{1}{m^2 c}\sum_{j=1}^{c}\sum_{k=1}^{m}\sum_{i=1}^{m} Z_{[i,k];j}$$

$C_{z_{SUB}}^2 = \frac{S_{z_{SUB}}^2}{\overline{Z}_{SUB}^2}$ is the coefficient of variation of the Z_{SUB},

$$C_{gz_{SUB}} = \rho C_g C_{z_{SUB}}$$

$$\omega_{z_{SUB}(i)}^2 = \frac{1}{m^2 c\overline{Z}_{SUB}^2}\sum_{i=1}^{m}\tau_{z_{SUB}(i)}^2$$

$$\omega_{gz_{SUB}(i)} = \frac{1}{m^2 c\overline{G}\overline{Z}_{SUB}}\sum_{i=1}^{m}\tau_{gz_{SUB}(i)}^2$$

and $\tau_{gz_{SUB}(i)}^2$ represents the cross-product of the deviation [18]. The RSS sub-exponential ratio type estimator and its MSE are formulated as [18]:

$$\hat{g}_{KK2} = \overline{g}_{RSS}\exp\left(\frac{\overline{Z}_{SUB} - \overline{z}_{RSS}}{\overline{Z}_{SUB} + \overline{z}_{RSS}}\right)$$

and

$$\mathrm{MSE}(\hat{g}_{KK2}) = \overline{G}^2 \left[\gamma C_{z_{SUB}}^2 - \omega_{z_{SUB}(i)}^2 + \frac{1}{4}\left(\gamma C_g^2 - \omega_{g(i)}^2\right) - \left(\gamma C_{gz_{SUB}} - \omega_{gz_{SUB}(i)}\right)\right] \tag{2}$$

The sub-regression type estimator and its MSE are given as [19]:

$$\hat{g}_{KR} = \overline{g}_{RSS} + \hat{b}\left(\overline{Z}_{SUB} - \overline{z}_{RSS}\right)$$

and

$$MSE(\hat{g}_{KR})_{SUB}^2 = \left[\gamma C_g^2 - \omega_g^2 - \frac{\left(\gamma C_{gz_{SUB}} - \omega_{gz_{SUB}}\right)^2}{\gamma C_{z_{SUB}}^2 - \omega_{z_{SUB}}^2}\right]_{\min} \tag{3}$$

where $\hat{b} = R_{SUB} \frac{\gamma C_{gz_{SUB}} - \omega_{gz_{SUB}}}{\gamma C_{z_{SUB}}^2 - \omega_{z_{SUB}}^2}$ and $R_{SUB} = \frac{\overline{G}}{\overline{Z}_{SUB}}$.

3. Proposed Estimator

Building upon the foundation laid by the sub-ratio, sub-exponential, and sub-regression type estimators [18,19], as well as the multiple auxiliary variable estimators [14-17], we introduce a novel sub-estimator for the population mean under ranked set sampling with two auxiliary variables with the following formulation:

$$\hat{g}_{PRO} = \left[\overline{g}_{RSS} \exp\left(\frac{\overline{Z}_{SUB} - \overline{z}_{RSS}}{\overline{Z}_{SUB} + \overline{z}_{RSS}}\right) \exp\left(\frac{\overline{X}_{SUB} - \overline{x}_{RSS}}{\overline{X}_{SUB} + \overline{x}_{RSS}}\right)\right]$$

where x is the second auxiliary variable,

$$\overline{x}_{RSS} = \frac{1}{mc} \sum_{j=1}^{c} \sum_{i=1}^{m} x_{[i,i];j}$$

and

$$\overline{X}_{SUB} = \frac{1}{m^2 c} \sum_{j=1}^{c} \sum_{k=1}^{m} \sum_{i=1}^{m} X_{[i,k];j}$$

To derive the MSE of the proposed estimator, we introduce the following definitions:

$$\overline{g}_{RSS} = \overline{G}\left(e_g + 1\right), \quad \overline{z}_{RSS} = \overline{Z}_{SUB}(e_z + 1), \quad \text{and} \quad \overline{x}_{RSS} = \overline{X}_{SUB}(e_x + 1) \tag{4}$$

where

$$\begin{aligned}
E\left(e_g\right) = E(e_z) = E(e_x) = 0 \\
E\left(e_g^2\right) = \gamma C_g^2 - \omega_g^2 = V_{200} \\
E(e_z^2) = \gamma C_{z_{SUB}}^2 - \omega_{z_{SUB}}^2 = V_{020} \\
E(e_x^2) = \gamma C_{x_{SUB}}^2 - \omega_{x_{SUB}}^2 = V_{002} \\
E\left(e_g e_z\right) = \gamma C_{gz_{SUB}} - \omega_{gz_{SUB}} = V_{110} \\
E\left(e_g e_x\right) = \gamma C_{gx_{SUB}} - \omega_{gx_{SUB}} = V_{101} \\
E(e_z e_x) = \gamma C_{z_{SUB}x_{SUB}} - \omega_{z_{SUB}x_{SUB}} = V_{011} \\
E\left(e_g e_z e_x\right) = \gamma C_{gz_{SUB}x_{SUB}} - \omega_{gz_{SUB}x_{SUB}} = V_{111}
\end{aligned} \tag{5}$$

By utilizing the provided definitions in Equation 4, we can express the estimator given in Equation 5 in linear form using Taylor series and second-degree approximation:

$$\hat{g}_{PRO} = \left[\overline{G}(e_g+1)\exp\left(\frac{\overline{Z}_{SUB}-\overline{Z}_{SUB}(e_z+1)}{\overline{Z}_{SUB}+\overline{Z}_{SUB}(e_z+1)}\right)\exp\left(\frac{\overline{X}_{SUB}-\overline{X}_{SUB}(e_x+1)}{\overline{X}_{SUB}+\overline{X}_{SUB}(e_x+1)}\right)\right]$$

$$= \overline{G}(e_g+1)\exp\left(\frac{-e_z}{e_z+2}\right)\exp\left(\frac{-e_x}{e_x+2}\right)$$

$$= \overline{G}(e_g+1)\exp\left[\frac{-e_z}{2}\left(\frac{e_z}{2}+1\right)^{-1}\right]\exp\left[\frac{-e_x}{2}\left(\frac{e_x}{2}+1\right)^{-1}\right]$$

$$= \overline{G}(e_g+1)\left(1-\frac{e_z}{2}+\frac{3e_z^2}{8}\right)\left(1-\frac{e_x}{2}+\frac{3e_x^2}{8}\right)$$

$$= \overline{G}+\overline{G}e_g\left(1-\frac{e_z}{2}+\frac{3e_z^2}{8}-\frac{e_x}{2}+\frac{3e_x^2}{8}+\frac{e_ze_x}{4}\right)$$

Hence,

$$\hat{g}_{PRO} = \overline{G}+\overline{G}\left(e_g-\frac{e_z}{2}-\frac{e_x}{2}+\frac{3e_z^2}{8}+\frac{3e_x^2}{8}+\frac{e_ze_x}{4}-\frac{e_ge_x}{2}-\frac{e_ge_z}{2}+\frac{e_ge_ze_x}{4}\right) \tag{6}$$

When subtracting from Equation 5 and then taking the expected value, the bias of the estimator is determined as follows:

$$B(\hat{g}_{PRO}) = \overline{G}^2\left(\frac{3V_{020}}{8}+\frac{3V_{002}}{8}+\frac{V_{011}}{4}-\frac{V_{101}}{2}-\frac{V_{110}}{2}+\frac{V_{111}}{4}\right)$$

After subtracting $\overline{G}$ from Equation 6 and subsequently squaring and taking the expected value, upon substituting the equations provided in Equation 5 into their corresponding positions, we arrive at the MSE equation of the estimator as follows:

$$MSE(\hat{g}_{PRO}) = \overline{G}^2\left(V_{200}+\frac{V_{020}}{4}+\frac{V_{002}}{4}-V_{110}-V_{101}+\frac{V_{011}}{2}\right) \tag{7}$$

For theoretical comparisons, we can express Equations 1–3 in the following forms:

$$MSE(\hat{g}_{KK1}) = \overline{G}^2(V_{200}+V_{020}-2V_{110})$$

$$MSE(\hat{g}_{KK2}) = \overline{G}^2\left(V_{200}+\frac{1}{4}V_{020}-V_{110}\right)$$

and

$$MSE(\hat{g}_{KR})\overline{G}^2\left(V_{200}+\frac{V_{110}^2}{V_{020}}\right)_{min}$$

Subsequently, by employing Equation 7, we can perform concluding the disparities between the estimators:

i. If $MSE(\hat{g}_{PRO}) < MSE(\hat{g}_{KK1})$ then $\frac{V_{002}-4V_{101}+2V_{011}}{3V_{020}-4V_{110}} < 1$.

ii. If $MSE(\hat{g}_{PRO}) < MSE(\hat{g}_{KK2})$ then $\frac{V_{002}+2V_{011}}{4V_{110}} < 1$.

iii. If $MSE(\hat{g}_{PRO}) < MSE(\hat{g}_{KR})$ then $-\frac{V_{020}(V_{002}-4V_{110}-4V_{101}+2V_{011})}{V_{020}^2+4V_{110}^2} > 1$.

4. Numerical Study

This section encompasses the numerical computations performed on synthetic and real data sets using the existing and proposed estimator. All calculations are carried out utilizing the R programming language in numerical studies.

4.1. Monte Carlo Simulation

Trivariate random observations (G, Z, X) are generated from a trivariate normal distribution characterized by parameters: $\mu_g = \mu_z = \mu_x \approx 5$, and $\sigma_g = \sigma_z = \sigma_x \approx 1$, correlation coefficients $\rho_{gz} = \rho_{gx} = \rho_{zx} \approx 0.7$, 0.8, and 0.9, along with population sizes $N = 100$ and 1000. In the existing literature, it is recommended that the set size "m" for the RSS method does not exceed 5, and it is often chosen as 3, 4, or 5. Therefore, 100000 RSS samples were drawn from these populations with a set size of $m \in \{3,4,5\}$ and cycle $c \in \{1,2,3,4\}$. The estimators' values were computed based on these selected RSS samples.

MSE values for the estimators are computed using Equation 8, while the relative efficiency (RE) values are obtained via Equation 9. In the context of the RE comparison, the reference variable is designated as the RSS basic mean estimator ($\hat{g}_{RSS}$). The results of these comparisons are tabulated in Tables 1-3, each corresponding to distinct parameter combinations. These tables encapsulate valuable insights into the performance of the estimators under varying conditions.

$$MSE(\hat{g}_h) = \sum_{j=1}^{100000} \frac{\left(\hat{g}_{hj} - \bar{G}\right)^2}{100000}, \quad h \in \{RSS, KK1, KK2, KR, PRO\} \tag{8}$$

$$RE(\hat{g}_l) = \frac{MSE(\hat{g}_{RSS})}{MSE(\hat{g}_l)}, \quad l \in \{KK1, KK2, KR, PRO\} \tag{9}$$

Table 1. RE results for $\rho_{gz} = \rho_{gx} = \rho_{zx} \approx 0.7$

							$N = 1000$					
RE	m	5	5	5	5	4	4	4	4	3	3	3	3
	c	4	3	2	1	4	3	2	1	4	3	2	1
$\hat{g}_{KK1}$		1.0837	1.0957	1.0845	1.0917	1.1028	1.0906	1.0975	1.0943	1.0933	1.0967	1.1006	1.0927
$\hat{g}_{KK2}$		1.1099	1.1161	1.1117	1.1164	1.1238	1.1174	1.1209	1.1198	1.1194	1.1211	1.1235	1.1197
$\hat{g}_{KR}$		1.1176	1.1254	1.1129	1.1092	1.1248	1.1246	1.1222	1.1034	1.1215	1.1210	1.1175	1.0911
$\hat{g}_{PRO}$		**1.2244**	**1.2398**	**1.2246**	**1.2329**	**1.2361**	**1.2176**	**1.2282**	**1.2255**	**1.2059**	**1.2079**	**1.2070**	**1.2031**
							$N = 100$					
RE	m	5	5	5	5	4	4	4	4	3	3	3	3
	c	4	3	2	1	4	3	2	1	4	3	2	1
$\hat{g}_{KK1}$		1.1683	1.1598	1.1639	1.1613	1.1647	1.1512	1.1560	1.1557	1.1478	1.1431	1.1451	1.1420
$\hat{g}_{KK2}$		1.1552	1.1510	1.1528	1.1513	1.1547	1.1483	1.1500	1.1516	1.1456	1.1430	1.1433	1.1433
$\hat{g}_{KR}$		1.1766	1.1764	1.1722	1.1686	1.1707	1.1720	1.1688	1.1646	1.1641	1.1602	1.1586	1.1442
$\hat{g}_{PRO}$		**1.3396**	**1.3222**	**1.3199**	**1.3174**	**1.3066**	**1.2982**	**1.3071**	**1.2931**	**1.2684**	**1.2649**	**1.2643**	**1.2550**

Boldfaced values indicate the "best" performances.

Table 2. RE results for $\rho_{gz} = \rho_{gx} = \rho_{zx} \approx 0.8$

$\boldsymbol{N = 1000}$												
RE $\boldsymbol{m}$	5	5	5	5	4	4	4	4	3	3	3	3
$\boldsymbol{c}$	4	3	2	1	4	3	2	1	4	3	2	1
$\hat{g}_{KK1}$	1.2043	1.1980	1.1993	1.1952	1.2133	1.2088	1.2114	1.2121	1.2072	1.2047	1.2027	1.1977
$\hat{g}_{KK2}$	1.1708	1.1681	1.1675	1.1665	1.1772	1.1751	1.1762	1.1763	1.1729	1.1714	1.1702	1.1680
$\hat{g}_{KR}$	1.2056	1.2044	1.2010	1.1825	1.2071	1.2087	1.2021	1.1875	1.2003	1.1944	1.1877	1.1610
$\hat{g}_{PRO}$	**1.4034**	**1.3979**	**1.3941**	**1.3924**	**1.3828**	**1.3722**	**1.3789**	**1.3789**	**1.3398**	**1.3371**	**1.3369**	**1.3219**
$\boldsymbol{N = 100}$												
RE $\boldsymbol{m}$	5	5	5	5	4	4	4	4	3	3	3	3
$\boldsymbol{c}$	4	3	2	1	4	3	2	1	4	3	2	1
$\boldsymbol{\hat{g}_{KK1}}$	1.1904	1.1820	1.1892	1.1844	1.1875	1.1864	1.1852	1.1791	1.1784	1.1752	1.1751	1.1694
$\boldsymbol{\hat{g}_{KK2}}$	1.2306	1.2281	1.2292	1.2279	1.2263	1.2265	1.2256	1.2218	1.2128	1.2099	1.2095	1.2070
$\boldsymbol{\hat{g}_{KR}}$	1.2418	1.2321	1.2248	1.1898	1.2300	1.2258	1.2133	1.1805	1.2137	1.2088	1.1929	1.1546
$\boldsymbol{\hat{g}_{PRO}}$	**1.3452**	**1.3399**	**1.3403**	**1.3375**	**1.3265**	**1.3268**	**1.3285**	**1.3210**	**1.2986**	**1.2898**	**1.2846**	**1.2841**

Boldfaced values indicate the “best” performances.

Table 3. RE results for $\rho_{gz} = \rho_{gx} = \rho_{zx} \approx 0.9$

$\boldsymbol{N = 1000}$												
RE $\boldsymbol{m}$	5	5	5	5	4	4	4	4	3	3	3	3
$\boldsymbol{c}$	4	3	2	1	4	3	2	1	4	3	2	1
$\hat{g}_{KK1}$	1.3782	1.3816	1.3722	1.3805	1.3722	1.3744	1.3689	1.3659	1.3416	1.3377	1.3333	1.3252
$\hat{g}_{KK2}$	1.2765	1.2799	1.2761	1.2799	1.2735	1.2738	1.2732	1.2700	1.2518	1.2498	1.2495	1.2436
$\hat{g}_{KR}$	1.3762	1.3730	1.3608	1.3385	1.3634	1.3600	1.3537	1.3086	1.3264	1.3100	1.3028	1.2655
$\hat{g}_{PRO}$	**1.5093**	**1.5155**	**1.5074**	**1.5184**	**1.4786**	**1.4852**	**1.4732**	**1.4799**	**1.4243**	**1.4194**	**1.4140**	**1.4109**
$\boldsymbol{N = 100}$												
RE $\boldsymbol{m}$	5	5	5	5	4	4	4	4	3	3	3	3
$\boldsymbol{c}$	4	3	2	1	4	3	2	1	4	3	2	1
$\boldsymbol{\hat{g}_{KK1}}$	1.4533	1.4484	1.4471	1.4368	1.4204	1.4166	1.4186	1.4094	1.3651	1.3651	1.3659	1.3639
$\boldsymbol{\hat{g}_{KK2}}$	1.3240	1.3233	1.3218	1.3190	1.3041	1.3045	1.3053	1.3007	1.2708	1.2693	1.2707	1.2694
$\boldsymbol{\hat{g}_{KR}}$	1.4387	1.4240	1.4113	1.3855	1.4011	1.3968	1.3794	1.33975	1.3455	1.3466	1.3196	1.2751
$\boldsymbol{\hat{g}_{PRO}}$	**1.6143**	**1.6016**	**1.6037**	**1.5920**	**1.5385**	**1.5379**	**1.5418**	**1.5285**	**1.4497**	**1.4537**	**1.4530**	**1.4422**

Boldfaced values indicate the “best” performances.

4.2. Real Data Application

This section extends the simulation study employed in the preceding segment to utilize real rather than synthetic data. The dataset originates from the compilation of data across 81 provinces in Türkiye [20]. The variables used within this context are as follows: $\boldsymbol{G}$: The population of Türkiye in the year 2021; $\boldsymbol{Z}$: The number of registered vehicles in the year 2017, and $\boldsymbol{X}$: The tally of traffic accidents involving fatalities or injuries in 2017. The underlying population parameters are briefly summarized as follows: $\overline{G} = 1045435$, $\overline{Z} = 15400.6$, $\overline{X} = 2255.173$, $\sigma_g = 1914343$, $\sigma_z = 50392.4$, $\sigma_x = 2662.278$, $\rho_{gz} = 0.97$, $\rho_{gx} = 0.87$, and $\rho_{zx} = 0.77$. 100000 RSS samples were drawn from this population ($N = 81$), considering various set sizes $m \in \{3,4,5\}$, and cycle $c \in \{1,2,3,4\}$. After calculating the estimator values from these samples, RE values were computed utilizing Equations 8 and 9. The outcomes of these computations are presented in Table 4, offering insights into the performance of the estimators within the context of real data.

Table 4. RE results for a real data set

RE	$N = 1000$												
	m	5	5	5	5	4	4	4	4	3	3	3	3
	c	4	3	2	1	4	3	2	1	4	3	2	1
$\hat{g}_{KK1}$		0.8853	0.9082	0.9102	0.8641	0.8205	0.8214	0.7860	0.71628	0.7274	0.6975	0.6608	0.5989
$\hat{g}_{KK2}$		3.7009	3.5807	3.3467	3.1058	2.8741	2.8033	2.6894	2.6239	2.2153	2.1860	2.1698	2.2215
$\hat{g}_{KR}$		1.3834	1.2962	1.2149	1.0786	1.7861	1.8303	1.8906	1.9278	**2.5979**	**2.8551**	**3.2944**	**4.1129**
$\hat{g}_{PRO}$		**3.7707**	**3.7500**	**3.6274**	**3.5907**	**2.9717**	**2.9644**	**2.9386**	**3.0829**	2.3119	2.3417	2.4090	2.6944

Boldfaced values indicate the “best” performances.

5. Conclusion

This study aims to introduce a new sub-estimator that surpasses the existing alternatives available in the literature regarding efficiency. A novel sub-estimator is proposed to achieve this objective, which utilizes two auxiliary variables without relying on population parameters. Once the theoretical framework of this novel estimator is established, the subsequent step involves substantiating its efficiency through numerical investigations.

The outcomes of the numerical studies indicate that, across all scenarios, the proposed estimator consistently outperforms other estimators in terms of effectiveness, as evidenced by simulation results derived from trivariate normal distributions. Considering the simulation study conducted on real data, the proposed estimator is the optimal choice in most cases, except for instances where $m = 3$.

In the simulation study and the real data application, a discernible pattern emerged, underscoring that the effectiveness of the estimator exhibited a positive correlation with the parameter “m”, denoting the number of ranked sets employed in the sampling process. Specifically, as the value of “m” increased, there was a notable enhancement in the estimator’s performance.

For prospective research, exploring the proposed estimator’s behavior in the context of skewed distributions and under varied sampling methodologies is recommended. Furthermore, an intriguing avenue for exploration involves devising sub-estimator adaptations under the ranked set sampling (RSS) of the proposed estimators for other sampling methods, subsequently assessing their efficiencies.

Author Contributions

The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

References

[1] Z. Chen, Z. Bai, B. K. Sinha, Ranked Set Sampling: Theory and Applications, Springer, New York, 2004.

[2] D. A. Wolfe, *Ranked Set Sampling*, Wiley Interdisciplinary Reviews: Computational Statistics 2 (4) (2010) 460–466.

[3] G. P. Patil, Ranked Set Sampling. Encyclopedia of Environmetrics, 2006.

[4] S. L. Stokes, *Ranked Set Sampling With Concomitant Variables*, Communications in Statistics-Theory and Methods 6 (12) (1977) 1207–1211.

[5] S. Bhushan, A. Kumar, T. Zaman, A. Al Mutairi, *Efficient Difference and Ratio-Type Imputation Methods under Ranked Set Sampling*, Axioms 12 (6) (2023) 558 22 pages.

[6] S. Bhushan, A. Kumar, *A Novel log Type Class of Estimators under Ranked Set Sampling*, Sankhya B 84 (2022) 421–447.

[7] S. Bhushan, A. Kumar, *On Optimal Classes of Estimators under Ranked Set Sampling*, Communications in Statistics - Theory and Methods 51 (8) (2022) 2610–2639.

[8] E. G. Koçyiğit, C. Kadilar, *Ratio Estimators for Ranked Set Sampling in the Presence of Tie Information*, Communications in Statistics-Simulation and Computation 51 (11) (2022) 6826–6839.

[9] T. Zaman, E. Dünder, A. Audu, D. A. Alilah, U. Shahzad, M. Hanif, *Robust Regression-Ratio-Type Estimators of the Mean Utilizing Two Auxiliary Variables: A Simulation Study*, Mathematical Problems in Engineering 2021 (2021) Article ID 6383927 9 pages.

[10] S. K. Yadav, T. Zaman, *Use of Some Conventional and Non-Conventional Parameters for Improving the Efficiency of Ratio-Type Estimators*, Journal of Statistics and Management Systems 24 (5) (2021) 1077–1100.

[11] M. Ijaz, T. Zaman, H. Bulut, A. Ullah, S. M. Asim, *An Improved Class of Regression Estimators Using the Auxiliary Information*, Journal of Science and Arts 20 (4) (2020) 789–800.

[12] K. U. I. Rather, E. G. Koçyiğit, R. Onyango, C. Kadilar, *Improved Regression in Ratio Type Estimators Based on Robust M-Estimation*, PloS ONE 17 (12) (2022) e0278868 16 pages.

[13] K. U. I. Rather, E. G. Koçyiğit., C. Ünal, *New Exponential Ratio Estimator in Ranked Set Sampling*, Pakistan Journal of Statistics and Operation Research 18 (2) (2022) 403–409.

[14] C. Kadilar, H. Cingi, *A New Estimator Using Two Auxiliary Variables*, Applied Mathematics and Computation 162 (2) (2005) 901–908.

[15] L. Khan, J. Shabbir, S. A. Khan, *Efficient Estimators of Population Mean in Ranked Set Sampling Scheme Using Two Concomitant Variables*, Journal of Statistics and Management Systems 22 (8) (2019) 1467–1480.

[16] S. Bhushan, A. Kumar, *New Efficient Logarithmic Estimators Using Multi-Auxiliary Information under Ranked Set Sampling*, Concurrency, and Computation: Practice and Experience 34 (27) (2022) e7337 23 pages.

[17] S. Bhushan, A. Kumar, N. Alsadat, M. S. Mustafa, M. M. Alsolmi, *Some Optimal Classes of Estimators Based on Multi-Auxiliary Information*, Axioms 12 (6) (2013) 515 25 pages.

[18] E. G. Koçyiğit, C. Kadilar, *Information Theory Approach to Ranked Set Sampling and New Sub-Ratio Estimators*, Communications in Statistics-Theory and Methods (2022) 23 pages.

[19] E. G. Koçyiğit, K. U. I. Rather, *The New Sub-Regression Type Estimator in Ranked Set Sampling*, Journal of Statistical Theory and Practice 17 (2) (2023) 27 14 pages.

[20] TÜİK, https://biruni.tuik.gov.tr/medas/, Accessed 16 Aug 2023.