BALANCED MODIFIED SYSTEMATIC SAMPLING IN THE PRESENCE OF LINEAR TREND

L. R. Naidoo [1]
University of KwaZulu-Natal
e-mail: *lewereeve@hotmail.com*
and
D. North
University of KwaZulu-Natal
and
T. Zewotir
University of KwaZulu-Natal
and
R. Arnab
University of Botswana and University of KwaZulu-Natal

Key words: Balanced systematic sampling, centered systematic sampling, linear systematic sampling, mean square error, modified systematic sampling, super-population model, Yates' end corrections.

Summary: In the presence of linear trend, linear systematic sampling (LSS) is less efficient than stratified random sampling (STR) and more efficient than simple random sampling (SRS). Consequently, some authors have proposed modifications to the LSS design, which have shown to yield optimal results under certain conditions. In this paper, a further modified design, termed as balanced modified systematic sampling (BMSS), is proposed. BMSS is compared to various well-known modified LSS designs as well as LSS, SRS and STR. If half the sample size is an even integer, then BMSS is optimal. To obtain linear trend free sampling results for the other cases of the sample size, a BMSS with end corrections (BMSSEC) estimator is constructed. The results in this paper suggest that the proposed estimator performs better than all other estimators for odd sample sizes and even sampling intervals. Moreover, the proposed estimator is competitive for all other cases.

1. Introduction

If one wishes to draw a sample of size n from a population of size N, then a simple way to do this would be to randomly select a unit and then to select subsequent units at equally spaced intervals, until a sample of size n is achieved. More specifically, if one randomly selects a unit from the first

[1] Corresponding author.

AMS: 62D05

$k = N/n$ units and every kth unit thereafter, then this sampling design is known as linear systematic sampling (LSS), provided that k is an integer (Cochran, 1977). LSS is advantageous over simple random sampling without replacement (SRS) and stratified random sampling (STR) (based on the random selection of one unit per stratum from n strata, each of size k), owing to its convenience and operational simplicity when implemented.

Consider a finite population $U = (U_1, ..., U_N)$ of size N and let y_q be the value of the study variable of the qth unit of population U, for $q \in \{1, ..., N\}$. Accordingly, the population mean $\overline{Y} = \sum_{q=1}^{N} y_q/N$ is estimated from the sample mean $\bar{y}$. Suppose a population that exhibits linear trend, given by the model A

$$y_q = a + bq + e_q, \qquad q = 1, ..., N \tag{1}$$

where a and b are constants and the e_q's denote the random errors which follows Cochran's (1946) super-population model, i.e. if the function $\mathscr{E}$ denotes the average of all potential finite populations that can be drawn from model A, then

$$\mathscr{E}(e_q) = 0, \qquad \mathscr{E}\left(e_q^2\right) = \sigma^2, \qquad \mathscr{E}(e_q e_z) = 0\,(q \neq z).$$

By using (1), the population mean is given by

$$\overline{Y} = \frac{1}{N}\sum_{q=1}^{N} y_q = \frac{1}{N}\sum_{q=1}^{N} a + \frac{b}{N}\sum_{q=1}^{N} q + \frac{1}{N}\sum_{q=1}^{N} e_q = a + \frac{b(N+1)}{2} + \overline{\overline{e}},$$

where $\overline{\overline{e}} = \sum_{q=1}^{N} e_q/N$ denotes the mean random error of the population. Now, let $\bar{y}_{\text{LSS}}$, $\bar{y}_{\text{SRS}}$, and $\bar{y}_{\text{STR}}$, denote the sample means when conducting LSS, SRS and STR, respectively. Thus, when estimating $\overline{Y}$ under model A, we note that the expected mean square errors (MSEs) of $\bar{y}_{\text{LSS}}$, $\bar{y}_{\text{SRS}}$, and $\bar{y}_{\text{STR}}$, are respectively given by

$$M_{\text{LSS}} = \sigma_e^2 + \frac{b^2\left(k^2-1\right)}{12}, \tag{2}$$

$$M_{\text{SRS}} = \sigma_e^2 + \frac{b^2(N+1)(k-1)}{12}, \tag{3}$$

and

$$M_{\text{STR}} = \sigma_e^2 + \frac{b^2\left(k^2-1\right)}{12n}, \tag{4}$$

where $\sigma_e^2 = \sigma^2(1/n - 1/N)$ represents the minimum expected error variance component, while the second terms on the right hand side represent the linear trend components (Bellhouse, 1988). By comparing Equations (2) through to (4), we obtain

$$M_{\text{STR}} \leq M_{\text{LSS}} \leq M_{\text{SRS}}. \tag{5}$$

Accordingly, some authors have suggested modified LSS designs to remove the linear trend component in Equation (2) and thus improve efficiency. Yates (1948) proposed a corrected estimator which uses the LSS design and is termed as the *Yates' end corrections* (YEC) estimator. This estimate is obtained by applying appropriate weights to the first and the last sampling units. *Centered systematic sampling* (CESS) was first discussed by Madow (1953), where the centrally located linear systematic sample is selected and thus no randomization is required. The centered systematic sample mean is subject to bias, since certain population units have no chance of being selected for the sample (Murthy, 1967). A balanced arrangement reverses the order, with respect to the population unit indices, of every alternative set of k population units. Sethi (1965) considered the application of LSS on this arrangement and this design was later named as *balanced systematic sampling* (BSS) by Murthy (1967, p. 165). Singh, Jindal and Garg (1968) suggested the application of LSS on a modified arrangement, where a subset of units from the end of the population is reversed, with respect to their population unit indices. This sampling design is known as *modified systematic sampling* (MSS). Denote $\bar{y}_{\text{YEC}}$, $\bar{y}_{\text{CESS}}$, $\bar{y}_{\text{BSS}}$, and $\bar{y}_{\text{MSS}}$, as the sample means associated with YEC, CESS, BSS and MSS, respectively. When estimating $\overline{Y}$ under model A, the expected MSEs of these sample means are respectively given by

$$M_{\text{YEC}} = \sigma_e^2 + \frac{\sigma^2\left(k^2-1\right)}{6(n-1)^2k^2}, \tag{6}$$

$$M_{\text{CESS}} = \begin{cases} \sigma_e^2, & \text{if } k \text{ is odd} \\ \sigma_e^2 + b^2/4, & \text{if } k \text{ is even} \end{cases} \tag{7}$$

and

$$M_{\text{BSS}} = M_{\text{MSS}} = \begin{cases} \sigma_e^2, & \text{if } n \text{ is even} \\ \sigma_e^2 + b^2(k^2-1)/12n^2, & \text{if } n \text{ is odd} \end{cases} \tag{8}$$

(Fountain and Pathak, 1989). By referring to Equations (6) to (8), we note that: (*i*) while there is a complete removal of the linear trend component in M_{YEC}, there is a larger error variance component, owing to the uneven weighting of the sampling units; (*ii*) the linear trend component in M_{CESS} is only eliminated when k is odd; and (*iii*) both M_{BSS} and M_{MSS} are equivalent, with the linear trend components being removed only for the case when n is even. Good reviews for these designs are provided by Bellhouse and Rao (1975), Cochran (1977), Fountain and Pathak (1989), Singh (2003) and the corresponding references cited therein.

More recent optimal modified LSS designs for linear trend populations have been suggested by Subramani (2000, 2009, 2010) and Khan, Shabbir and Gupta (in press), while Mukerjee and Sengupta (1990) proposed optimal design-unbiased strategies to estimate $\overline{Y}$. As in the case of the earlier designs, these recent solutions are based on certain assumptions and/or are optimal for linear trend populations under certain conditions.

In the present paper, a modified LSS design, termed as *balanced modified systematic sampling* (BMSS), is proposed. In Section 2, a discussion on the methodology of BMSS is provided. For Section 3, the expected MSE of the BMSS sample mean, is compared to that of M_{LSS}, M_{SRS}, M_{STR}, M_{YEC}, M_{CESS}, M_{BSS} and M_{MSS}. As a result, BMSS is only optimal for the case when $n/2$ is an even integer. A BMSS with end corrections (BMSSEC) estimator is thus constructed, so as to remove the linear trend component in the corresponding expected MSE for the other cases of n. A numerical example on a hypothetical population is then considered in Section 4, before carrying out a simulation study in Section 5. Note that k is assumed to be an integer throughout this paper, i.e. assuming that N is an exact multiple of n, so that sampling is conducted linearly.

2. Balanced Modified Systematic Sampling (BMSS)

A modified arrangement used for BMSS is defined as follows: (a) if n is even, then the order of every alternative set of k population units is reversed, before reversing the order of the first/last $n/2$ sets of k population units; and (b) if n is odd, then the order of every alternative set of k population units is reversed, before reversing the order of the last $(n-1)/2$ sets of k population units. LSS is then applied to this modified arrangement, so as to select the required sample. Note that different arrangements, before applying LSS, will result in different compositions of samples and this paper deals with a specific arrangement, as explained above. By reversing the order of $n/2$ (or $(n-1)/2$) sets of k population units, a balancing effect is obtained which is optimal for populations exhibiting linear trend. Note that MSS reverses the order of the last $n/2$ (or $(n-1)/2$) sets of k population units, without alternating the order of each set, while BSS alternates the order of each set, without reversing the order of the last $n/2$ (or $(n-1)/2$) sets of k population units. Thus, the ordering of BMSS is a mixture of both, the MSS and BSS orderings. Moreover, BMSS reduces to LSS when $n = 2$ and we will thus assume that $n > 2$.

The above-mentioned design is equivalent to selecting sampling units according to the following indices:

(A) if $n/2$ is an even integer, then

$$i+2jk, \qquad 2(j+1)k-i+1, \qquad \text{for } j = 0,...,(n-4)/4$$

and

$$N+i-k-2jk, \qquad N-i-k-2jk+1, \qquad \text{for } j = 0,...,(n-4)/4;$$

(B) if $n/2$ is an odd integer, then

$$i+2jk, \qquad N+i-k-2jk, \qquad \text{for } j = 0,...,(n-2)/4$$

and

$$2(j+1)k-i+1, \qquad N-i-k-2jk+1, \qquad \text{for } j = 0,...,(n-6)/4;$$

(C) if $n = 3$, then

$$i, \qquad 2k-i+1 \qquad \text{and} \qquad N-i+1;$$

(D) if $n \neq 3$ and $(n+1)/2$ is an even integer, then

$$i+2jk, \qquad 2(j+1)k-i+1, \qquad N-i-2jk+1, \qquad \text{for } j = 0,...,(n-3)/4$$

and

$$N+i-2(j+1)k, \qquad \text{for } j = 0,...,(n-7)/4;$$

(E) if $(n+1)/2$ is an odd integer, then

$$\begin{array}{llll} i+2jk, & 2(j+1)k-i+1, & N-i-2jk+1, & N+i-2(j+1)k, \\ \text{for } j=0,\ldots,(n-5)/4 & & \text{and} & i+(n-1)k/2. \end{array}$$

Note that Cases (A) and (B) are sub-cases of n being even, while Cases (C) to (E) are sub-cases of $n>1$ being odd.

The ith ($i \in \{1,\ldots,k\}$) sample mean, denoted by $\bar{y}_{\text{BMSS}}$, is obtained by using the above sampling unit indices for the respective cases, e.g. if we consider Case (A), then the sample mean is given as

$$\bar{y}_{\text{BMSS}} = \frac{1}{n}\sum_{j=0}^{(n-4)/4}\left(y_{i+2jk} + y_{2(j+1)k-i+1} + y_{N+i-k-2jk} + y_{N-i-k-2jk+1}\right).$$

Note that $\bar{y}_{\text{BMSS}}$ is design-unbiased, since BMSS is viewed as an arrangement of units before applying LSS.

3. Expected Mean Square Error (MSE) Comparisons

To compare the expected MSE of the BMSS estimator, to that of M_{LSS}, M_{SRS}, M_{STR}, M_{YEC}, M_{CESS}, M_{BSS} and M_{MSS}, we first need to consider the following theorem.

Theorem 1 If we suppose model B, which is related to model A, given by

$$y_q = a + bq, \qquad q = 1,\ldots,N \tag{9}$$

such that

$$\overline{Y}_{\text{B}} = \frac{1}{N}\sum_{q=1}^{N} y_q = \frac{1}{N}\left[(a+b)+\ldots+(a+Nb)\right] = a + \frac{b\,(N+1)}{2},$$

then by assuming equal weights $(1/n)$ applied to all the sampling units, the expected MSE of any sample mean, when estimating $\overline{Y}$, is given by

$$M_{\text{A}} = \mathscr{E}\text{MSE}\left(\bar{y}_{\text{A}}\right) \triangleq \mathscr{E}\left\{\text{E}\left[\left(\bar{y}_{\text{A}} - \overline{Y}\right)^2\right]\right\} = \sigma_e^2 + \text{Var}\left(\bar{y}_{\text{B}}\right), \tag{10}$$

where $\bar{y}_{\text{B}}$ denotes a linear unbiased estimator of $\overline{Y}_{\text{B}}$, using the sampling design associated with $\bar{y}_{\text{A}}$.

Proof. By using Equations (1) and (9), we obtain $\overline{Y} = \overline{Y}_{\text{B}} + \bar{\bar{e}}$ and $\bar{y}_{\text{A}} = \bar{y}_{\text{B}} + \bar{e}_i$, where $\bar{e}_i = \sum e_i/n$ denotes the mean random error of the sample and $\sum$ denotes the sum over the sample. Using these expressions, it follows that

$$\begin{aligned} M_{\text{A}} &\triangleq \mathscr{E}\left\{\text{E}\left[\left(\bar{y}_{\text{A}} - \overline{Y}\right)^2\right]\right\} \\ &= \mathscr{E}\left\{\text{E}\left[\left(\bar{y}_{\text{B}} - \overline{Y}_{\text{B}}\right)^2 + \left(\bar{e}_i - \bar{\bar{e}}\right)^2\right]\right\} = \mathscr{E}\text{Var}\left(\bar{y}_{\text{B}}\right) + \mathscr{E}\text{Var}\left(\bar{e}_i\right) = \text{Var}\left(\bar{y}_{\text{B}}\right) + \sigma_e^2. \end{aligned}$$

■

If we let $P = 2i - k - 1$, then applying Equation (9) to $\bar{y}_{\text{BMSS}}$ results in

$$\begin{aligned} \bar{y}_{\text{BMSS}} &= a + b(N+1)/2, && \text{for Case (A)} \\ &= a + b\left[N+1+2P/n\right]/2, && \text{for Case (B)} \\ &= a + b\left[N+1-P/n\right]/2, && \text{for Cases (C) to (E).} \end{aligned}$$

Hence, the corresponding variance expression, when using $\bar{y}_{\text{BMSS}}$ to estimate $\overline{Y}_{\text{B}}$, is given by

$$\begin{aligned} \text{Var}\left(\bar{y}_{\text{BMSS}}\right) &= 0, && \text{for Case (A)} \\ &= b^2(k^2-1)/3n^2, && \text{for Case (B)} \\ &= b^2(k^2-1)/12n^2, && \text{for Cases (C) to (E),} \end{aligned} \tag{11}$$

which follows since

$$\text{E}\left(P^2\right) = \frac{1}{k}\sum_{i=1}^{k} P^2 = \frac{\left(k^2-1\right)}{3}.$$

Thus, if we assume model A, then by substituting Equation (11) into Equation (10), we obtain

$$\begin{aligned} M_{\text{BMSS}} &= \sigma_e^2, && \text{for Case (A)} \\ &= \sigma_e^2 + b^2(k^2-1)/3n^2, && \text{for Case (B)} \\ &= \sigma_e^2 + b^2(k^2-1)/12n^2, && \text{for Cases (C) to (E).} \end{aligned} \tag{12}$$

By comparing Equations (12) and (4), we note that $M_{\text{BMSS}} < M_{\text{STR}}$ for all the cases. Thus, by using Equation (5), we conclude that BMSS is more efficient than LSS, SRS and STR. Also, by comparing Equations (12) and (6), we see that $M_{\text{BMSS}} < M_{\text{YEC}}$, for (i) Case (A); (ii) Case (B) (if and only if $\sigma^2 > 2b^2(n-1)^2k^2/n^2$); and (iii) Cases (C) to (E) (if and only if $\sigma^2 > b^2(n-1)^2k^2/2n^2$). In addition, the comparison of Equations (12) and (7) results in:

(i) $M_{\text{BMSS}} = M_{\text{CESS}}$ for Case (A) and if k is odd;

(ii) $M_{\text{BMSS}} < M_{\text{CESS}}$ for Case (A) and if k is even;

(iii) $M_{\text{BMSS}} > M_{\text{CESS}}$ for Cases (B) to (E) and if k is odd;

(iv) $M_{\text{BMSS}} < M_{\text{CESS}}$ for Case (B), if k is even and $4k^2 - 4 < 3n^2$;

(v) $M_{\text{BMSS}} < M_{\text{CESS}}$ for Cases (C) to (E), if k is even and $k^2 - 1 < 3n^2$.

Finally, by comparing Equations (12) and (8), we see that $M_{\text{BMSS}} > M_{\text{BSS}} = M_{\text{MSS}}$ for Case (B), while all other cases result in $M_{\text{BMSS}} = M_{\text{BSS}} = M_{\text{MSS}}$.

Clearly, we only obtain a complete removal of the linear trend component in Equation (12) for Case (A). To remove the linear trend component for the other cases, we next consider the application of weights to the first and last sampling units. Accordingly, the resulting estimator and the corresponding expected MSE are respectively given in the next two theorems.

Theorem 2 The BMSSEC estimator of $\overline{Y}$ with random start i, for $i \in \{1,...,k\}$, is given as

$$\begin{aligned} \bar{y}_{\text{BMSSEC}} &= \bar{y}_{\text{BMSS}} + P(y_i - y_{N+i-k})/[n(N-k)], && \text{for Case (B)} \\ &= \bar{y}_{\text{BMSS}} - P(y_i - y_{N-i+1})/[2n(N-2i+1)], && \text{for Cases (C) and (D)} \\ &= \bar{y}_{\text{BMSS}} + P(y_i - y_{N-i+1})/[2n(N-2i+1)], && \text{for Case (E).} \end{aligned}$$

Proof. See Appendix. ■

Theorem 3 Under model A, the expected MSE of $\bar{y}_{\text{BMSSEC}}$ is given as

$$\begin{aligned} M_{\text{BMSSEC}} &= \sigma_e^2 + 2\sigma^2(k^2-1)/3n^2(N-k)^2, && \text{for Case (B)} \\ &= \sigma_e^2 + \sum_{i=1}^{k} \left\{ P^2\sigma^2/[2(N-2i+1)^2 n^2 k] \right\}, && \text{for Cases (C) to (E).} \end{aligned}$$

Proof. See Appendix. ■

If we compare M_{BMSSEC} to all previous expected MSE expressions, then we note that simple theoretical comparisons are difficult to obtain and we will thus resort to some numerical comparisons in the next two sections. However, one can easily verify that $M_{\text{BSS}} = M_{\text{MSS}} < M_{\text{BMSSEC}} < M_{\text{YEC}}$ for Case (B), while $M_{\text{CESS}} < M_{\text{BMSSEC}}$ if k is odd. Furthermore, just as in the case of the YEC estimator being slightly biased under the assumption of a rough linear trend, owing to the uneven weighting of the sampling units (Murthy, 1967), we obtain the same result for estimator $\bar{y}_{\text{BMSSEC}}$.

4. Numerical Example

Consider the hypothetical linear trend population given by Murthy and Rao (1988, p. 161), which is presented in Table 1. All the possible samples for various values of n when conducting BMSS, which are obtained by using the sampling unit indices in Section 2 for the corresponding cases, are presented in Table 2. The associated MSEs for the various sampling designs mentioned in this paper are given in Table 3. The results suggest that BMSS offers a strict improvement over LSS, SRS and STR, regardless of the sample size. Moreover, if $n/2$ is not a even integer, then we obtain a reduction in estimation error by using the BMSSEC estimator, as opposed to the BMSS estimator. Comparisons amongst the modified LSS designs to either BMSS or the BMSSEC estimator requires further analysis, since we are only considering a single finite population, whereas our theoretical results obtained earlier are based on an infinite super-population. However, we note that in most cases, there is a significant reduction in error when applying any one of the modified LSS designs, as opposed to LSS, SRS and STR.

5. Empirical Comparisons

Three independent simulation studies are carried out to further evaluate the estimator $\bar{y}_{\text{BMSSEC}}$. Monte Carlo simulations are used with the statistical software package R, where 10 000 finite populations are simulated. The expected MSE of each estimator is obtained by averaging the

Table 1: A population of 40 units exhibiting a steady linear trend in the value of a variable y.

U_q	y_q	U_q	y_q	U_q	y_q	U_q	y_q
U_1	0	U_{11}	10	U_{21}	23	U_{31}	41
U_2	1	U_{12}	11	U_{22}	25	U_{32}	43
U_3	2	U_{13}	12	U_{23}	29	U_{33}	46
U_4	3	U_{14}	12	U_{24}	30	U_{34}	50
U_5	4	U_{15}	13	U_{25}	32	U_{35}	52
U_6	5	U_{16}	14	U_{26}	33	U_{36}	53
U_7	7	U_{17}	15	U_{27}	35	U_{37}	57
U_8	7	U_{18}	17	U_{28}	38	U_{38}	59
U_9	8	U_{19}	20	U_{29}	39	U_{39}	62
U_{10}	9	U_{20}	22	U_{30}	40	U_{40}	63

Table 2: For various values of n, the k possible samples (for $i \in \{1,...,k\}$) using BMSS.

Case	n	k	Possible Samples
A	4	10	$S_i = \{U_i, U_{21-i}, U_{31-i}, U_{30+i}\}$
E	5	8	$S_i = \{U_i, U_{17-i}, U_{41-i}, U_{24+i}\} \cup \{U_{16+i}\}$
A	8	5	$S_i = \{U_i, U_{11-i}, U_{36-i}, U_{35+i}, U_{10+i}, U_{21-i}, U_{25+i}, U_{26-i}\}$
B	10	4	$S_i = \{U_i, U_{36+i}, U_{8+i}, U_{28+i}, U_{16+i}, U_{20+i}, U_{9-i}, U_{37-i}, U_{17-i}, U_{29-i}\}$
A	20	2	$S_i = \{U_i, U_{5-i}, U_{38+i}, U_{39-i}, U_{4+i}, U_{9-i}, U_{34+i}, U_{35-i}, U_{8+i}, U_{13-i}, U_{30+i}\}$ $\cup \{U_{31-i}, U_{12+i}, U_{17-i}, U_{26+i}, U_{27-i}, U_{16+i}, U_{21-i}, U_{22+i}, U_{23-i}\}$

Table 3: Mean square errors for a hypothetical population exhibiting a linear trend.

	n				
	4	5	8	10	20
LSS	23.1600	13.6475	6.3288	3.3825	0.4900
SRS	83.2264	64.7316	36.9895	27.7421	9.2474
STR	6.6350	3.1700	0.9625	0.4063	0.0350
YEC	0.4116	0.1887	0.1140	0.0240	0.0134
CESS	0.6400	0.4225	0.0400	0.9025	0.4900
BSS	0.4350	2.2475	0.0288	0.0275	0.0025
MSS	2.4725	0.0575	0.7538	0.2025	0.0400
BMSS	0.1475	0.5775	0.1788	0.2275	0.0025
BMSSEC	N/A	0.0730	N/A	0.0187	N/A

MSEs over the 10 000 populations. The relative expected MSEs of each comparative estimator, with respect to that of estimator $\bar{y}_{\text{BMSSEC}}$, is denoted by $R_\alpha = 100 \times M_{\text{BMSSEC}}/M_\alpha(\%)$, where $\alpha \in \{\text{LSS}, \text{SRS}, \text{STR}, \text{YEC}, \text{CESS}, \text{BSS}, \text{MSS}, \text{BMSS}\}$. Without loss of generality, we suppose that the e_q's are iid $N(0,1)$ random variables and let $a = 5$.

In the first simulation study, Case (B) is examined and arbitrary values of $b = 0.5, 1, 2$ and 4, are assigned while varying n and k. The associated relative expected MSEs are presented in Tables 4 to 7. From Tables 4 to 7, we note that only estimators $\bar{y}_{\text{BSS}}$ and $\bar{y}_{\text{MSS}}$, are marginally subjected to less error than that of estimator $\bar{y}_{\text{BMSSEC}}$. Also, estimator $\bar{y}_{\text{BMSSEC}}$ is always favourable over estimators $\bar{y}_{\text{LSS}}$, $\bar{y}_{\text{SRS}}$ and $\bar{y}_{\text{CESS}}$, with greater discrepancies as n, k and/or b increases. Similarly, we see that estimator $\bar{y}_{\text{BMSSEC}}$ is always preferred over estimator $\bar{y}_{\text{STR}}$, with greater discrepancies as k and/or b increases, while results remain constant as n varies. Finally, we note that estimator $\bar{y}_{\text{BMSSEC}}$ always performs better than estimator $\bar{y}_{\text{BMSS}}$, with greater discrepancies as k and/or b increases and smaller discrepancies as n increases. Thus, $M_{\text{BMSS}} \to M_{\text{BMSSEC}}$ as $n \to \infty$, provided that k and b are relatively small.

Table 4: Simulated relative expected mean square errors for populations exhibiting linear trend under Case (B) ($b = 0.5$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	6	57.85	23.99	90.73	57.85	101.51	102.63	92.40
2	34	19.08	01.02	89.47	19.08	101.07	100.29	98.49
2	130	05.76	00.07	88.25	05.76	102.61	102.51	99.52
2	258	02.99	00.02	88.24	02.99	99.16	99.05	99.87
4	6	28.87	07.50	71.57	66.63	100.70	101.63	79.28
4	34	06.60	00.26	70.74	26.17	100.42	100.33	94.84
4	130	01.82	00.02	70.87	08.53	102.52	99.01	98.84
4	258	00.92	<00.01	70.81	04.45	100.31	99.04	99.18
8	6	10.06	02.01	40.24	70.84	100.66	100.10	50.14
8	34	01.93	00.06	40.08	29.27	99.81	99.85	85.14
8	130	00.51	<00.01	40.10	09.79	100.46	100.15	95.62
8	258	00.26	<00.01	39.82	05.15	100.27	100.55	97.43

For the second simulation study, Cases (C) to (E) (i.e. n is odd) are considered and arbitrary values of $b = 0.5, 1, 2$ and 4, are assigned while varying n and k. The corresponding relative expected MSEs are presented in Tables 8 to 11. From Tables 8 to 11, we note that estimator $\bar{y}_{\text{BMSSEC}}$ performs better than all the estimators considered in this study. In this simulation study, we obtain similar results as those obtained in the previous study. However, estimator $\bar{y}_{\text{BMSSEC}}$ now performs better than estimators $\bar{y}_{\text{BSS}}$ and $\bar{y}_{\text{MSS}}$. Moreover, we see that estimators $\bar{y}_{\text{BSS}}$, $\bar{y}_{\text{MSS}}$ and $\bar{y}_{\text{BMSS}}$, are relatively subject to the same amount of error. Thus, $M_{\text{BSS}}, M_{\text{MSS}}$ and $M_{\text{BMSS}} \to M_{\text{BMSSEC}}$ as $n \to \infty$, provided that k and b are relatively small.

Comparisons between estimators $\bar{y}_{\text{BMSSEC}}$ and $\bar{y}_{\text{YEC}}$ are evaluated in the third simulation study. Because there are no trend components in the expected MSEs of both estimators, an arbitrary value of $b = 4$ is assigned while varying n and k. Also, only Cases (C) to (E) are explored, as it was theoretically shown previously that $M_{\text{BMSSEC}} < M_{\text{YEC}}$ for Case (B). The simulated relative expected MSEs are presented in Table 12. The results suggest that estimator $\bar{y}_{\text{BMSSEC}}$ is only preferred when n and k are small. Otherwise, there are marginal gains when choosing estimator $\bar{y}_{\text{BMSSEC}}$ over estimator $\bar{y}_{\text{YEC}}$.

Table 5: Simulated relative expected mean square errors for populations exhibiting linear trend under Case (B) ($b = 1$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	6	24.95	07.16	66.55	24.95	99.00	99.53	75.38
2	34	05.59	00.25	66.74	05.59	100.84	100.51	94.65
2	130	01.54	00.02	67.78	01.54	100.91	102.35	98.60
2	258	00.80	<00.01	69.26	00.80	105.18	105.60	99.44
4	6	09.25	01.99	38.14	33.49	100.93	101.08	48.34
4	34	01.72	00.06	37.23	08.08	98.99	99.99	83.69
4	130	00.46	<00.01	37.46	02.25	99.82	101.26	95.29
4	258	00.23	<00.01	37.49	01.15	100.23	101.06	97.11
8	6	02.70	00.51	14.28	36.91	100.38	100.58	19.93
8	34	00.49	00.02	14.26	09.31	100.27	100.73	58.78
8	130	00.13	<00.01	14.16	02.61	99.12	99.41	84.22
8	258	00.06	<00.01	14.18	01.33	102.09	100.64	91.56

Table 6: Simulated relative expected mean square errors for populations exhibiting linear trend under Case (B) ($b = 2$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	6	07.68	01.90	33.33	07.68	99.73	101.68	42.20
2	34	01.44	00.06	32.99	01.44	99.04	99.59	80.93
2	130	00.38	<00.01	33.50	00.38	99.90	101.72	94.94
2	258	00.19	<00.01	33.54	00.19	101.66	101.58	96.97
4	6	02.43	00.50	13.03	11.09	100.52	99.60	18.31
4	34	00.44	00.02	13.21	02.18	100.84	99.90	56.22
4	130	00.11	<00.01	12.95	00.57	101.01	99.83	83.06
4	258	00.06	<00.01	13.07	00.29	99.54	99.92	90.80
8	6	00.68	00.13	03.99	12.77	99.65	99.86	05.86
8	34	00.12	<00.01	04.00	02.51	100.08	100.40	25.97
8	130	00.03	<00.01	04.01	00.67	99.76	99.37	57.89
8	258	00.02	<00.01	03.99	00.34	99.41	99.41	72.91

Table 7: Simulated relative expected mean square errors for populations exhibiting linear trend under Case (B) ($b = 4$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	6	02.08	00.49	11.30	02.08	103.38	102.87	16.23
2	34	00.36	00.02	11.05	00.36	102.27	99.00	51.55
2	130	00.10	<00.01	11.16	00.10	101.55	99.06	80.80
2	258	00.05	<00.01	10.83	00.05	99.02	100.08	88.48
4	6	00.62	00.12	03.60	03.02	100.68	99.61	05.31
4	34	00.11	<00.01	03.59	00.54	99.15	99.43	23.98
4	130	00.03	<00.01	03.61	00.14	98.94	99.51	55.27
4	258	00.01	<00.01	03.59	00.07	99.13	100.16	70.01
8	6	00.17	00.03	01.03	03.54	100.02	100.92	01.54
8	34	00.03	<00.01	01.04	00.65	100.75	100.82	08.21
8	130	00.01	<00.01	01.03	00.17	100.02	100.40	25.13
8	258	<00.01	<00.01	01.03	00.08	100.62	99.91	40.06

Table 8: Simulated relative expected mean square errors for populations exhibiting linear trend under Cases (C) to (E) ($b = 0.5$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	3	70.39	53.54	88.52	70.39	98.00	98.41	98.72
2	35	18.35	00.94	87.75	18.35	98.16	98.51	99.57
2	125	05.92	00.08	88.00	05.92	99.01	99.66	99.88
2	255	03.08	00.02	89.77	03.08	99.89	99.99	99.96
4	3	45.40	24.17	72.16	81.30	91.01	89.63	89.99
4	35	06.41	00.24	70.49	25.65	98.96	98.54	98.78
4	125	01.89	00.02	70.82	08.74	99.29	99.36	99.59
4	255	00.93	<00.01	70.67	04.50	99.84	99.99	99.84
8	3	18.50	07.55	40.69	83.34	68.31	67.80	68.28
8	35	01.86	00.06	39.89	28.77	95.34	96.24	95.97
8	125	00.54	<00.01	40.36	10.25	98.78	99.47	98.78
8	255	00.26	<00.01	40.13	05.26	99.40	99.88	99.32

6. Conclusion

A modified LSS design (i.e. BMSS) that depends on an arrangement of population units before applying LSS, which results in the corresponding sample mean being design-unbiased, has been proposed. Results from Sections 3 to 5 indicate that BMSS is more efficient than LSS, SRS and STR, in the presence of linear trend. The optimal case of BMSS is when $n/2$ is an even integer, which results in linear trend free sampling and minimum expected MSE of the corresponding sample mean. For the other cases of BMSS, a modified end corrections estimator, i.e. estimator $\bar{y}_{\text{BMSSEC}}$, has been constructed. Populations exhibiting a rough linear trend result in estimator $\bar{y}_{\text{BMSSEC}}$ being a slightly

Table 9: Simulated relative expected mean square errors for populations exhibiting linear trend under Cases (C) to (E) ($b = 1$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	3	42.66	23.21	70.25	42.66	88.72	88.20	88.28
2	35	05.39	00.24	66.54	05.39	98.63	98.12	98.60
2	125	01.57	00.02	66.46	01.57	98.30	98.48	98.60
2	255	00.78	<00.01	66.48	00.78	99.51	99.97	99.79
4	3	17.03	07.28	38.28	50.89	65.39	66.23	65.88
4	35	01.68	00.06	37.29	07.88	94.72	94.44	95.06
4	125	00.48	<00.01	37.58	02.36	98.88	98.33	98.82
4	255	00.24	<00.01	38.00	01.18	99.99	98.97	99.51
8	3	05.42	02.02	14.71	54.98	34.36	34.26	34.36
8	35	00.48	00.02	14.39	09.24	85.56	85.21	85.64
8	125	00.13	<00.01	14.25	02.71	95.45	96.21	95.36
8	255	00.07	<00.01	14.37	01.36	96.95	97.72	97.57

Table 10: Simulated relative expected mean square errors for populations exhibiting linear trend under Cases (C) to (E) ($b = 2$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	3	14.66	06.79	34.26	14.66	61.90	62.10	61.82
2	35	01.38	00.06	32.71	01.38	95.03	94.90	94.55
2	125	00.39	<00.01	32.60	00.39	98.22	98.37	98.43
2	255	00.19	<00.01	33.09	00.19	98.94	99.75	99.25
4	3	04.88	01.92	13.36	20.68	31.62	31.83	31.91
4	35	00.42	00.02	12.99	02.08	84.01	83.59	84.05
4	125	00.12	<00.01	12.98	00.59	94.68	93.93	94.98
4	255	00.06	<00.01	13.05	00.29	98.13	97.76	97.46
8	3	01.40	00.51	04.09	22.99	11.37	11.35	11.34
8	35	00.12	<00.01	03.98	02.43	59.25	59.13	59.01
8	125	00.03	<00.01	03.99	00.69	83.52	83.49	83.69
8	255	00.02	<00.01	04.03	00.34	91.83	90.98	91.54

biased estimate of $\overline{Y}$ as well as exhibiting an inflated error variance component in the corresponding expected MSE, owing to the uneven weighting of the sampling units.

If $n/2$ is an odd integer, then estimator $\bar{y}_{\text{BMSSEC}}$ is subject to less error than estimators $\bar{y}_{\text{LSS}}$, $\bar{y}_{\text{SRS}}$, $\bar{y}_{\text{STR}}$, $\bar{y}_{\text{YEC}}$ and $\bar{y}_{\text{BMSS}}$, while marginally susceptible to more error than estimators $\bar{y}_{\text{BSS}}$ and $\bar{y}_{\text{MSS}}$, as shown in Sections 3 and 5. In addition, if n is odd, then estimator $\bar{y}_{\text{BMSSEC}}$ is subject to less error than all of the above-mentioned estimators. The simulation study in Section 5 indicates that estimator $\bar{y}_{\text{BMSSEC}}$ performs better than estimator $\bar{y}_{\text{YEC}}$ if n is odd, provided that n and k are small. Otherwise, there are marginal gains when opting to use estimator $\bar{y}_{\text{BMSSEC}}$ over estimator $\bar{y}_{\text{YEC}}$. Under this circumstance, one may opt to use estimator $\bar{y}_{\text{YEC}}$, owing to simplicity.

Table 11: Simulated relative expected mean square errors for populations exhibiting linear trend under Cases (C) to (E) ($b = 4$).

k	n	R_{LSS}	R_{SRS}	R_{STR}	R_{CESS}	R_{BSS}	R_{MSS}	R_{BMSS}
2	3	04.06	01.78	11.28	04.06	27.59	27.81	27.72
2	35	00.35	00.02	11.05	00.35	81.58	81.06	81.51
2	125	00.10	<00.01	11.16	00.10	92.83	93.56	93.80
2	255	00.05	<00.01	11.13	00.05	96.56	97.40	97.06
4	3	01.27	00.49	03.71	06.02	10.41	10.31	10.38
4	35	00.11	<00.01	03.62	00.53	56.26	56.53	56.25
4	125	00.03	<00.01	03.63	00.15	82.43	82.85	82.33
4	255	00.01	<00.01	03.64	00.07	90.81	90.58	90.72
8	3	00.36	00.13	01.06	07.02	03.12	03.11	03.11
8	35	00.03	<00.01	01.04	00.62	26.93	26.93	26.89
8	125	00.01	<00.01	01.03	00.17	56.48	57.07	56.55
8	255	<00.01	<00.01	01.03	00.09	73.05	73.11	72.63

Table 12: Simulated relative expected mean square errors of the YEC sample mean, with respect to that of the MBMSSEC sample mean, for populations exhibiting linear trend under Cases C to E.

	n								
	3	5	7	13	15	29	63	125	255
$k = 2$	86.56	92.63	95.53	97.49	97.83	98.82	99.02	99.59	99.17
$k = 4$	89.23	94.03	96.13	97.69	98.48	98.97	99.41	99.59	99.82
$k = 8$	90.04	94.64	96.38	97.90	98.53	99.16	99.65	99.92	99.95

Finally, we note that estimator $\overline{y}_{\text{BMSSEC}}$ performs better than estimator $\overline{y}_{\text{CESS}}$, provided that k is even, as seen in the simulation study from Section 5. However, if k is odd, then the theoretical results in Section 3 suggest that estimator $\overline{y}_{\text{CESS}}$ is to be the preferred, as CESS is an optimal sampling design for this scenario. Nevertheless, we can expect marginal gains when opting to use estimator $\overline{y}_{\text{CESS}}$ over estimator $\overline{y}_{\text{BMSSEC}}$ when k is odd.

Recommendations for the most appropriate design(s) under various conditions are provided in Table 13. Note that the third column represents a trade-off between estimators $\overline{y}_{\text{YEC}}$ and $\overline{y}_{\text{BMSSEC}}$, where preference is either given to minimum MSE or simplicity.

Note that the end corrections estimators are constructed with the assumption of a perfect linear trend in the population, i.e. model B in Equation (9). The YEC and BMSSEC estimators are thus most suitable under the assumption of a population exhibiting linear trend. The onus then lies on the survey statistician to collect as much information about the population as possible, prior to sampling, so as to estimate the population structure. This may entail the building of appropriate models, where he/she can then use the most suitable design and/or estimator. In addition, it is common practice for a survey statistician to apply auxiliary information for sampling. If the population is arranged in increasing/decreasing order in accordance with an auxiliary variable, then we obtain an approximate linear trend in the population, where the higher the degree of correlation between the auxiliary variable and the characteristic under study, the greater degree of linear trend. Under these circumstances, the theory and results presented in this paper may then apply.

Table 13: Recommended designs for populations exhibiting linear trend.

Case(s)	Condition	Preference	Recommended Design(s)
A	k is even	N/A	BSS, MSS or BMSS
A	k is odd	N/A	CESS, BSS, MSS or BMSS
B	k is even	N/A	BSS or MSS
B	k is odd	N/A	CESS, BSS or MSS
C to E	k is even; n and k are small	Minimum MSE	BMSSEC
C to E	k is even; n and k are small	Simplicity	YEC
C to E	k is even; n and/or k are not small	Minimum MSE	YEC or BMSSEC
C to E	k is even; n and/or k are not small	Simplicity	YEC
C to E	k is odd	N/A	CESS

Acknowledgments

The authors wish to thank the referee for the helpful comments and suggestions as well as the NRF (National Research Foundation) for the funding that was provided during the production of this paper.

Appendix

Theorem 2

Proof. An estimate of $\overline{Y}$ with random start i, for $i \in \{1, \ldots, k\}$, can be written as

$$\bar{y}_{\text{BMSSEC}} = \frac{1}{n}\left[\psi_1 y_{x_1} + \sum_{j=2}^{(n-1)} y_{x_j} + \psi_2 y_{x_n}\right], \tag{13}$$

where ψ_1 and ψ_2 are the weights applied to the first and the last sampling units respectively and $x_1, \ldots, x_n$ are the sampling unit indices, which are arranged in ascending order. By substituting Equation (9) into Equation (13) and then equating this result to $\overline{Y}_{\text{B}}$, we obtain

$$\bar{y}_{\text{BMSSEC}} = \frac{1}{n}\left[\psi_1 (a + bx_1) + \sum_{j=2}^{(n-1)} (a + bx_j) + \psi_2 (a + bx_n)\right] = a + \frac{b(N+1)}{2}. \tag{14}$$

By equating the coefficients of a in Equation (14), it follows that

$$\psi_1 = 2 - \psi_2. \tag{15}$$

Similarly, by equating the coefficients of b in Equation (14), we obtain

$$\frac{1}{n}\left[\psi_1 x_1 + \sum_{j=2}^{(n-1)} x_j + \psi_2 x_n\right] = \frac{N+1}{2}. \tag{16}$$

Substituting Equation (15) into Equation (16) results in

$$2\left[2x_1-\psi_2 x_1+\sum_{j=2}^{(n-1)} x_j+\psi_2 x_n\right]=n\,(N+1)\,,$$

which simplifies to

$$\psi_2=\frac{K-2x_1}{x_n-x_1}, \tag{17}$$

where $K=n(N+1)/2-\sum_{j=2}^{n-1}x_j$. The weight applied to the first sampling unit is thus obtained by substituting Equation (17) into Equation (15), i.e.

$$\psi_1=\frac{2x_n-K}{x_n-x_1}. \tag{18}$$

Substituting Equations (17) and (18) into Equation (13) results in

$$\begin{aligned}\bar{y}_{\text{BMSSEC}}&=\frac{1}{n}\left[\frac{(2x_n-K)}{(x_n-x_1)}y_{x_1}+\sum_{j=2}^{n-1}y_{x_j}+\frac{(K-2x_1)}{(x_n-x_1)}y_{x_n}\right]\\&=\bar{y}_{\text{BMSS}}+\frac{1}{n}\left[\frac{(2x_n-K)}{(x_n-x_1)}y_{x_1}+\frac{(K-2x_1)}{(x_n-x_1)}y_{x_n}-y_{x_1}-y_{x_n}\right]\\&=\bar{y}_{\text{BMSS}}+\frac{[(x_n+x_1)-K]}{n\,(x_n-x_1)}\left(y_{x_1}-y_{x_n}\right).\end{aligned} \tag{19}$$

Now, if we consider Case (B), then $x_1=i$, $x_n=N+i-k$ and

$$\begin{aligned}K&=\frac{n(N+1)}{2}-\sum_{j=2}^{n-1}x_j\\&=\frac{n(N+1)}{2}-\left[\sum_{j=1}^{(n-2)/4}(2i+N-k)+\sum_{j=0}^{(n-6)/4}(k-2i+2+N)\right]\\&=\frac{n(N+1)}{2}-\sum_{j=1}^{(n-2)/4}(2i+N-k+k-2i+2+N)=\frac{(N+1)[n-(n-2)]}{2}=N+1\end{aligned}$$

(refer to the sampling unit indices of Case (B) in Section 2). On substituting these values into Equation (19), we obtain

$$\bar{y}_{\text{BMSSEC}}=\bar{y}_{\text{BMSS}}+\frac{P}{n(N-k)}\left(y_i-y_{N+i-k}\right). \tag{20}$$

We then conclude the proof by finding the values of x_1, x_n and K for the other cases, as shown above, and then substituting these values into Equation (19). ■

Theorem 3

Proof. The expected MSE of $\bar{y}_{\text{BMSSEC}}$ can be written as

$$\begin{aligned}M_{\text{BMSSEC}}&\stackrel{\Delta}{=}\mathscr{E}\left[\mathrm{E}\left(\left\{\bar{y}_{\text{BMSSEC}}-\bar{Y}\right\}^2\right)\right]\\&=\mathrm{E}\left\{\mathscr{E}\left[\left(\bar{y}_{\text{BMSSEC}}-\bar{Y}\right)^2\right]\right\}=\frac{1}{k}\sum_{i=1}^{k}\mathscr{E}\left[\bar{y}_{\text{BMSSEC}}-\bar{Y}\right]^2.\end{aligned} \tag{21}$$

If we consider Case (B) for Equation (1), then

$$\bar{y}_{\text{BMSSEC}}-\overline{Y}=a+\frac{b}{2}\left[N+1+\frac{2P}{n}\right]+\bar{e}_i-\left[a+\frac{b(N+1)}{2}+\bar{\bar{e}}\right]=\frac{bP}{n}+\bar{e}_i-\bar{\bar{e}}. \tag{22}$$

Moreover,

$$y_i-y_{N+i-k}=a+bi+e_i-[a+b(N+i-k)+e_{N+i-k}]=-b[N-k]+e_i-e_{N+i-k}. \tag{23}$$

Using Equations (20), (22) and (23), we obtain

$$\begin{aligned}\mathscr{E}\left[(\bar{y}_{\text{BMSSEC}}-\overline{Y})^2\right]&=\mathscr{E}\left[\left(\frac{bP}{n}+\bar{e}_i-\bar{\bar{e}}+\frac{P[-b(N-k)+e_i-e_{N+i-k}]}{n(N-k)}\right)^2\right]\\&=\mathscr{E}\left[\left(\bar{e}_i-\bar{\bar{e}}+\frac{P[e_i-e_{N+i-k}]}{n(N-k)}\right)^2\right].\end{aligned} \tag{24}$$

Applying Cochran's (1946) super-population model assumptions, results in

$$\mathscr{E}\left(\bar{e}_i^2\right)=\frac{1}{n^2}\left[\sum\mathscr{E}\left(e_i^2\right)+\sum\sum_{j\neq i}\mathscr{E}\left(e_ie_j\right)\right]=\frac{\sigma^2}{n}, \tag{25}$$

$$\mathscr{E}\left(\bar{\bar{e}}^2\right)=\frac{1}{N^2}\left[\sum_{j=1}^{N}\mathscr{E}\left(e_j^2\right)+\sum_{i=1}^{N}\sum_{j\neq i}^{N}\mathscr{E}\left(e_ie_j\right)\right]=\frac{\sigma^2}{N}, \tag{26}$$

$$\mathscr{E}\left(\bar{e}_i\bar{\bar{e}}\right)=\frac{1}{nN}\sum\sum_{j=1}^{N}\mathscr{E}\left(e_ie_j\right)=\frac{n\sigma^2}{nN}=\frac{\sigma^2}{N}, \tag{27}$$

$$\mathscr{E}\left[\bar{e}_i\left(e_i-e_{i+(n-1)k}\right)\right]=\mathscr{E}\left[\left(\frac{1}{n}\sum_{j=1}^{n}e_{\{i+(j-1)k\}}\right)\left(e_i-e_{i+(n-1)k}\right)\right]=0, \tag{28}$$

$$\mathscr{E}\left[\bar{\bar{e}}\left(e_i-e_{i+(n-1)k}\right)\right]=\mathscr{E}\left[\left(\frac{1}{N}\sum_{j=1}^{N}e_j\right)\left(e_i-e_{i+(n-1)k}\right)\right]=0 \tag{29}$$

and

$$\mathscr{E}\left[\left(e_i-e_{i+(n-1)k}\right)^2\right]=\mathscr{E}\left[e_i^2-2e_ie_{i+(n-1)k}+e_{i+(n-1)k}^2\right]=2\sigma^2. \tag{30}$$

Expanding Equation (24) and then substituting Equations (25) through to (30) into this expression, results in

$$\mathscr{E}\left[(\bar{y}_{\text{BMSSEC}}-\overline{Y})^2\right]=\sigma_e^2+\frac{2\sigma^2P^2}{n^2(N-k)^2}. \tag{31}$$

Finally, by substituting Equation (31) into Equation (21), we obtain

$$M_{\text{BMSSEC}}=\sigma_e^2+\frac{2\sigma^2}{n^2k(N-k)^2}\sum_{i=1}^{k}P^2=\sigma_e^2+\frac{2\sigma^2(k^2-1)}{3n^2(N-k)^2}.$$

Similarly, we can obtain $\mathscr{E}\left[(\bar{y}_{\text{BMSSEC}}-\overline{Y})^2\right]$ for Cases (C) to (E) and then substitute these expressions into Equation (21). ■

References

BELLHOUSE, D. R. (1988). Systematic sampling. *In* KRISHNAIAH, P. R. AND RAO, C. R. (Editors) *Handbook of Statistics*, volume 6. Elsevier Science Publishers: Amsterdam, pp. 125–145.

BELLHOUSE, D. R. AND RAO, J. N. K. (1975). Systematic sampling in the presence of a trend. *Biometrika*, **62** (3), 694–697.

COCHRAN, W. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *The Annals of Mathematical Statistics*, **17**, 164–177.

COCHRAN, W. (1977). *Sampling Techniques*. 3rd edition. John Wiley and Sons, Inc.: New York.

FOUNTAIN, R. AND PATHAK, P. (1989). Systematic and nonrandom sampling in the presence of linear trends. *Communications in Statistics - Theory and Methods*, **18** (7), 2511–2526.

KHAN, Z., SHABBIR, J., AND GUPTA, S. N. (in press). Generalized systematic sampling. *Communications in Statistics - Simulation and Computation*.

MADOW, W. G. (1953). On the theory of systematic sampling, iii. comparison of centered and random start systematic sampling. *The Annals of Mathematical Statistics*, **24** (1), 101–106.

MUKERJEE, R. AND SENGUPTA, S. (1990). Optimal estimation of a finite population mean in the presence of linear trend. *Biometrika*, **77** (3), 625–630.

MURTHY, M. (1967). *Sampling: Theory and Methods*. Statistical Publishing Society: Calcutta.

MURTHY, M. N. AND RAO, T. J. (1988). Systematic sampling with illustrative examples. *In* KRISHNAIAH, P. R. AND RAO, C. R. (Editors) *Handbook of Statistics*, volume 6, Sampling. Elsevier Science Publishers: Amsterdam, pp. 147–185.

SETHI, V. K. (1965). On optimal pairing of units. *Sankhyā: The Indian Journal of Statistics, Series B*, **27**, 315–320.

SINGH, D., JINDAL, K. K., AND GARG, J. N. (1968). On modified systematic sampling. *Biometrika*, **55** (3), 541–546.

SINGH, S. (2003). *Advanced Sampling Theory with Applications*, volume 1. Springer: Netherlands.

SUBRAMANI, J. (2000). Diagonal systematic sampling scheme for finite populations. *Journal of the Indian Society of Agricultural Statistics*, **53** (2), 187–195.

SUBRAMANI, J. (2009). Further results on diagonal systematic sampling scheme for finite populations. *Journal of the Indian Society of Agricultural Statistics*, **63** (3), 277–282.

SUBRAMANI, J. (2010). Generalization of diagonal systematic sampling scheme for finite populations. *Model Assisted Statistics and Applications*, **5** (2), 117–128.

YATES, F. (1948). Systematic sampling. *Philosophical Transactions of the Royal Society of London, Series A*, **241** (834), 345–377.

Manuscript received, 2014-04-02, revised, 2015-03-03, accepted, 2015-03-10.