Special Proceedings: ISBN#: 978-81-950383-2-9
25th Annual Conference, 15–17 February 2023; pp 21–30

How Gainfully can the Additional Units be Used?

Opendra Salam Singh[1], Gurumayum Sandweep Sharma[1] and Bikas K. Sinha[2]
[1]*Manipur University, Imphal, Manipur, India*
[2]*Former Professor, Indian Statistical Institute, Kolkata, India*

Received: 01 June 2023; Revised: 08 June 2023; Accepted: 11 June 2023

Preamble

Bikas K. Sinha [BKS] is the recipient of "VK Gupta Endowment Award for Achievements in Statistical Thinking and Practice - 2023". The Society of Statistics and Computer Applications [SSCA], New Delhi, gave this award — upon receiving recommendation from its Executive Council [EC]. While receiving the award, BKS made an online presentation during SSCA Annual Conference held in Jammu during February 15-17, 2023. This paper originated from that presentation. BKS is happy to induct Opendra Salam Singh and Gurumayum Sandweep Sharma of the Department of Statistics, Manipur University, Imphal as his collaborators. As BKS says "I have chosen to speak on a topic which is simple to state and comprehend. Yet, the technicalities are quite involved." Simply stated, it goes almost like a proverb: "Larger the sample size, more is the precision"! AND we all know that this is indeed true for [SRSWOR (N, n), Sample Mean] Strategy. What about [SRSWR (N, n), Sample Mean] strategy? We must qualify sample mean under srswr: mean based on all units including repeats or mean based on distinct units only? Under both the situations, the claim is valid in some sense. Research scholars may engage themselves for a clear proof.

We intend to discuss some features of this problem of variance reduction *via* enhanced resources in terms of possession of additional population units at a later stage.

Key Words: Sampling designs; Sampling strategies; Unbiased estimators [UEs]; Homogeneous UEs [HUEs]; Linear UEs [LUEs]; HLUEs; SRS WR/WOR schemes; Horvitz-Thompson estimator [HTE]; First and second order inclusion probabilities; Connected sampling designs; Additional units; Improved sampling strategies; Lanke's estimator.

1. Introduction

We start with a Sampling Strategy based on a Fixed-Size (n) [abbreviated as FS(n)] Sampling Design and an HLUE $e(s(n)|Y)$ of a Finite Population Mean $\bar{Y}$ corresponding to a study character Y. Once the sampling design has been chosen and implemented, and a sample $s(n)$ has been chosen and, further, data collection has been completed, we are told about Enhanced Resource in the sense of k additional units! The enhanced sample size now becomes $(n + k)$, once an additional sampling design of fixed size (k) [$FS(k)$] defined over the complement of $s(n)$ is adopted. It leads to the revised HLUE $e(s(n + k)|Y)$ based on the union of the two samples of which $s(n)$ is already at hand.

Corresponding Author: Bikas K. Sinha
Email: bikasksinha118@gmail.com

One pertinent question to be asked is: Whatever be the choice of the initial HLUE $e(s(n)|Y)$ and the choice of the additional sampling design $FS(k)$ [on the complement of $s(n)$], does there exist a suitably defined HLUE $e(s(n + k)|Y)$ which provides uniformly smaller variance than $e(s(n)|Y)$?

Clearly, in the case of *SRSWOR* (N, n), followed by *SRSWOR* $(N - n, k)$, we end up with *SRSWOR* $(N, n + k)$, and hence choice of the corresponding sample means is well understood for the domination result to hold for every $k \geq 1$.

However, in case of *SRSWR* (N, n), the follow-up sampling operation could be

(i) *SRSWR* (N, k) or, (ii) *SRSWR* $(N - v(n); k)$

where $v(n)$ refers to the number of distinct units selected under *SRSWR* (N, n). In a way, under (ii), therefore, the sample is selected under *SRSWR* out of the complement of the units already selected under *SRSWR* (N, n). Whereas the combination under (i) refers to two independent draws from the whole population, under (ii), the two sets of samples are necessarily disjoint. However, within each, units drawn are not necessarily distinct, as we take recourse to *WR* sampling.

The question we ask is: For a given choice of $e(s(n)|Y)$, what is the choice of $e(s(n) \cup s(k)|Y)$ for variance reduction? Here, $s(n) \cup s(k)$ must be understood in the most general sense.

This is apparently not an easy problem to address. There are two choices for $e(s(n)|Y)$ under *SRSWR* (N, n): mean based on all units, and mean based on distinct units [notation $v(n)$]. Note that *SRSWR* $(N - v(n), k)$ excludes the distinct units selected in the first round. So, data analysis is conditional not only on $v(n)$, but also on the actual units selected under $s(n)$. Naturally, these are excluded during the second draw. We leave it for research scholars to ponder over this non-standard inference problem.

Bagchi and Sinha [2022] have addressed a different version of this problem. We do not intend to enter into this matter.

2. Data analysis under FS(.) designs

Consider $FS(.)$ sampling designs – both Initial $D(N; n)$ and Extension $D(N - n; k)$. Denote by $[D(N, n), e(s(n)|Y)]$ the initial sampling strategy for unbiased estimation of a finite population total or mean.

Let $D^*(N - n, k|s(n))$ be the follow-up sampling design $FS(N - n, k)$, conditional on exclusion of $s(n)$.

We ask the question: Given $e(s(n)|Y)$, how would one define $e^*(s(n + k)|Y|s(n))$ – once totally new additional k units are available *via* $s(k)$ - following $FS(N - n, k)$, defined over compliment of $s(n)$, for every $s(n)$ with $P(s(n)) > 0$?

Naturally, we desire:

(i) $E^{*}\,[e^{*}\,(s(n+k)|Y\,|s(n))] = E\,[e(s(n)|Y)]$ (1)

(ii) $V^{*}\,[e^{*}\,(s(n+k)|Y|s(n))] \leq V\,[e(s(n)|Y)]$ (2)

uniformly in Ys, where $E^{*} = E1E2$ and $V^{*} = V1E2 + E1V2$, in usual notations.

This specific problem has been resolved by Lanke (1975) who provided an explicit expression for $e^{*}(s(n+k)|Y)$'s in terms of the $e(s(n)|Y)$'s, provided that $s(n)$ is a subset of $e(s(n+k))$.

We take up this exercise in the sequel.

2.1. Lanke's formula

Lanke (1975) considered extending an arbitrary sampling strategy $[D(N, n), e(s(n)|Y)]$ to another sampling strategy $[D(N, m{=}n+k), e(s(n+k)|Y)]$ *via* $Q(N-n, k)$ so that $[D(N, m = n+k), e(s(n+k)|Y)]$ is better than $[D\,(N, n), e(s(n)|Y)]$, irrespective of the choice of $Q(N-n, k)$.

Lanke proposed the estimator $e(s(n+k)|Y)$ through the relation

$$e(s(n+k)|Y)[P(s(n+k))] = \sum_{s(n)\in s(n+k)} e(s(n)|Y)[P\,(s(n))Q(s(n+k)-s(n)] \quad (3)$$

Here summation is over all $s(n)$ [subsets of $s(n+k)$].

Further,

$$P\,[s(n{+}k)] = \sum_{s(n)\in s(n+k)}[P\,(s(n))Q(s(n+k)-s(n))] \quad (4)$$

summation being over all $s(n)$ [subsets of $s(n+k)$].

It transpires that Lanke basically applied Rao-Blackwellization technique *i.e.*, averaging technique to produce estimator(s) with reduced sum of squares. We display the technical details below. Upon squaring both sides of (3) and rewriting the same, we obtain

$$e^{2}(s(n+k)|Y)[P\,(s(n+k))] =$$

$$[\sum_{s(n)\in s(n+k)} e(s(n)|Y)[P(s(n))Q(s(n+k)-s(n))]]^{2}/P(s(n+k)) \quad (5)$$

By appealing to C-S inequality [elaborated below], we derive from (5):

$$e^{2}(s(n+k)|Y)[P(s(n+k))] \leq [\sum_{s(n)\in s(n+k)} e^{2}(s(n)|Y)[P(s(n))Q(s(n+k)-s(n))]] \quad (6)$$

which further simplifies to

$$[\sum_{s(n)\in s(n+k)} e^{2}(s(n)|Y)[P(s(n))] \quad (7)$$

Hence the domination result follows in appropriate subgroups and hence on the whole.

2.2. Illustrative example: Lanke's formula

Here we take an example to demonstrate the domination result. We start with $N = 10$, $n = 5$, $k = 2$. Let us adopt the initial sampling design in the form:

Table 1: Initial sampling design of fixed size *n* = 5

Sl. No.	*P* (…)
1.	P (1, 2, 3, 4, 5) = 0.075
2.	P (1, 3, 5, 8, 10) = 0.105
3.	P (1, 4, 6, 7, 9) = 0.165
4.	P (4, 6, 7, 8, 10) = 0.135
5.	P (2, 3, 6, 9, 10) = 0.145
6.	P (3, 4, 7, 8, 10) = 0.175
7.	P (5, 6, 7, 8, 9) = 0.180
8.	P (2, 4, 6, 7, 8) = 0.020

The extended design for $k = 2$ must be defined for every sample $s(n)$ [listed above] on its compliment with reference to the whole set of $N = 10$ units. Note that the design shown above is already connected in the sense of positive probability attached to all pairwise units *i.e.*, $P(i, j) > 0$ for all pairs. So, the choice of complimentary samples for the extended design is very simple and we need not restrict to any conditions except that these are complimentary in nature! Of course, the sample size $k = 2$ has to be kept in mind. We take up the following Example of extended design to this effect.

Table 2: Initial description of *s*(*n*) and extension design using *s*(*k*)

Sl. No.	Initial Design	Extension Design
1.	(1, 2, 3, 4, 5)	*P* (6, 7) = 0.4; *P* (6, 9) = 0.3; *P* (8, 10) = 0.3∗
2.	(1, 3, 5, 8, 10)	*P* (2, 4) = 0.7∗; P (4, 7) = 0.3 ∗ ∗
3.	(1, 4, 6, 7, 9)	*P* (2, 3) = 0.5; *P* (5, 8) = 0.5 ∗ ∗ ∗ ∗
4.	(4, 6, 7, 8, 10)	*P* (3, 9) = 0.3 ∗ ∗∗; *P* (3, 5) = 0.7
5.	(2, 3, 6, 9, 10)	*P* (4, 5) = 0.6; *P* (5, 7) = 0.3; *P* (5, 8) = 0.1
6.	(3, 4, 7, 8, 10)	*P* (1,5) = 0.4∗∗; *P* (2, 9) = 0.4; *P* (6, 9) = 0.2∗∗∗
7.	(5, 6, 7, 8, 9)	*P* (1, 4) = 0.3∗∗∗∗; *P* (2, 10) = 0.5; *P* (3, 4) = 0.2
8.	(2, 4, 6, 7, 8)	*P* (3, 9) = 1.00

Remark 1: Note that for the last design [Sl. No. 8], the extension design is degenerate.

Composition of samples based on extended sampling design is shown below:

(1, 2, 3, 4, 5, 6, 7) (8.1)
(1, 2, 3, 4, 5, 6, 9) (8.2)
(1, 2, 3, 4, 5, 8, 10)* (8.3)
(1, 2, 3, 4, 5, 8, 10)* (8.4)
(1, 3, 4, 5, 7, 8, 10)** (8.5)

(1, 2, 3, 4, 6, 7, 9)	(8.6)
(1, 4, 5, 6, 7, 8, 9)***	(8.7)
(3, 4, 6, 7, 8, 9, 10)****	(8.8)
(3, 4, 5, 6, 7, 8, 10)	(8.9)
(2, 3, 4, 5, 6, 9, 10)	(8.10)
(2, 3, 5, 6, 7, 9, 10)	(8.11)
(2, 3, 5, 6, 8, 9, 10)	(8.12)
(1, 3, 4, 5, 7, 8, 10)**	(8.13)
(2, 3, 4, 7, 8, 9, 10)	(8.14)
(3, 4, 6, 7, 8, 9, 10)****	(8.15)
(1, 4, 5, 6, 7, 8, 9)***	(8.16)
(2, 5, 6, 7, 8, 9, 10)	(8.17)
(3, 4, 5, 6, 7, 8, 9)	(8.18)
(2, 3, 4, 6, 7, 8, 9)	(8.19)

Remark 2: There are altogether 19 extended samples formed through the extension formula. However, not all are distinct. For example, the sample (1, 2, 3, 4, 5, 8, 10)* is formed of (i) (1, 2, 3, 4, 5) combined with (8, 10) as well as of (ii) (1, 3, 5, 8, 10) combined with (2, 4). Lanke argued that once the extended sample is available through the extension formula, both the subsets (i) and (ii) are available and they produce $e(s(n)|Y)$ based on initial sample $s(n)$ under both (i) and (ii). Then he suggested the formula shown above in (3) for combining the two estimators. In this example, for the extended sample (1, 2, 3, 4, 5, 8,10)*, the formula yields

$$e((1,2,3,4,5,8,10)*) = [e((1,2,3,4,5))P(1,2,3,4,5)P((8,10)|s(n))+ e((1,3,5,8,10))P(1,3,5,8,10)P((2,4)|s(n))]/ [P(1,2,3,4,5)P((8,10)|s(n)) + P(1,3,5,8,10)P((2,4)|s(n))] \quad (9)$$

In effect, the estimator based on the extended sample is a convex combination of the two initial estimators listed in (i) and (ii) and these are both available whenever the extended sample (1, 2, 3, 4, 5, 8, 10) is realized. It may be noted that for the extended sample (1, 2, 3, 4, 5, 8, 10), P (1, 2, 3, 4, 5, 8, 10) is given by the denominator above in (9).

Towards variance, or equivalently, sum of squares [SS] computation, we find: $e^2((1,2,3,4,5,8,10))P((1,2,3,4,5,8,10)) = [e((1,2,3,4,5))P(1,2,3,4,5)P((8,10)|s(n))+ e((1,3,5,8,10))P(1,3,5,8,10)\mathrm{P}((2,4)|\mathrm{s(n)})]^2/[\mathrm{P}((1,2,3,4,5,8,10))]$. (10)

Set

$$a1 = e(1,2,3,4,5)[P((1,2,3,4,5))P((8,10)|s(n))]^{1/2};$$
$$b1 = [P((1,2,3,4,5))P((8,10)|s(n))]^{1/2} \quad (11)$$
$$a2 = e(1,3,5,8,10)[P((1,3,5,8,10))P((2,4)|\mathrm{s}(n))]^{1/2};$$
$$b2 = [P((1,3,5,8,10))P((2,4)|s(n))]^{1/2} \quad (12)$$

By C-S inequality, we know that

$$[a1b1 + a2b2]^2 \leq [a1^2 + a2^2][b1^2 + b2^2] \tag{13}$$

which leads to

$$\text{RHS of } (10) \leq [e^2((1, 2, 3, 4, 5))P(1, 2, 3, 4, 5)P((8, 10)|s(n)) + e^2((1, 3, 5, 8, 10))P(1, 3, 5, 8, 10)P((2, 4)|s(n))]. \tag{14}$$

Once all the samples are utilized like in the above, we can go back to computation of the upper bound of the sum of squares [*SS*] of the extended estimator $e(s(n + k)|Y)$. This yields, for example, terms like

$$e^2((1, 2, 3, 4, 5))P(1, 2, 3, 4, 5)P((8, 10)|s(n)); \tag{15}$$
$$e^2((1, 2, 3, 4, 5))P(1, 2, 3, 4, 5)P((6, 7)|s(n)); \tag{16}$$
$$e^2((1, 2, 3, 4, 5))P(1, 2, 3, 4, 5)P((6, 9)|s(n)). \tag{17}$$

These three expressions add to $e^2((1, 2, 3, 4, 5))P(1, 2, 3, 4, 5)$, upon obvious simplification. Likewise, we carry on similar computations of the *SS* for the estimators based on extended samples and upon application of C-S inequality, we end up with upper bounds as *SS* based on the samples in the initial sampling design.

Remark 3: It is interesting to note that the three extension designs [(8.1), (8.2) and (8.3)*] arising out of a single initial sample do provide three different estimators for the population parameter. After that, the *SS* for each estimator is examined in the light of the C-S inequality. Taken together, we find that the *SS* for the estimator based on extension design is less than or equal to that of the original estimator. We provide below in Section 4 necessary details to encourage the interested teachers and researchers follow the technicalities in settling the claim.

3. Behavior of Horvitz-Thompson estimator

In as early as 1967, Prabhu-Ajgaonkar discussed the possibility of an hlue based on a sample of size n to outperform an hlue based on an extended design of size $(n + 1)$ - the two estimators belonging to the same class of hlues. The sampling design was chosen to be the Midzuno Sampling Scheme for a sample of size $n = 2$ and it was to be extended to the Generalized Midzuno Sampling Scheme for a sample of size $n = 3$. However, he actually worked out the case of $n = 1$ against $n = 2$ and that was not at all appealing. Our attempt to work for $n = 2$ to $n = 3$ did not go through with the set-up adopted by him.

Starting with an arbitrarily specified initial $FS(N, n)$ design and extending it to a $FS(N, n + k)$ design by increasing the sample size from n to $n + k < N$, confining to the use of the *HTE* in both the designs, one may not succeed in uniformly improving over the *HTE* based on the original design. A quick and tricky proof goes like this. Let $(\pi_i(n)|FS(N, n)); i = 1,2,...,N)$ denote the first order inclusion probabilities based on the original design so that $\sum(\pi_i(n)|FS(N, n)) = n$. Set $Y_i = K\pi_i(n)/n$, $i = 1,2, ...,N$, where K is an arbitrary positive constant.

Then $HTE[s(n)] = \sum_{i\in s(n)} Y_i/\pi_i(n) = K$ for every $s(n)$ with $P(s(n)) > 0$. Hence, $V(HTE) = 0$ at the stated values of Y_i's. In the same vein, evaluated at the same Y-values,

$$HTE(s(n+k)) = \sum_{i\in s(n+k)} Y_i/\pi_i(n+k) \quad (18)$$

$$= [K/n] \sum_{i\in s(n+k)} \pi_i(n)/\pi_i(n+k) \quad (19)$$

Therefore, unless $\pi_i(n)/\pi_i(n+k)$ is the same for all $i = 1,2,...,N$, the second estimator has a strictly positive variance. Therefore, uniform domination is not possible using *HTE* in both the situations.

For the case where the extended sampling design $Q(N-n, k|s(n))$ is SRSWOR, Sinha (1980) presented simple conditions on the first and second order inclusion probabilities of the original sampling design $FS(N, n)$ so that $HTE(s(n+t)|Y)$ is better than $HTE(s(n+t-1)|Y)$ simultaneously for all $t = 1,2, ...,k$ for any arbitrary choice of $k < N-n$.

Sengupta (1982) extensively studied the properties of Lanke's estimator for various choices of $e(s(n)|Y)$ [based on $FS(N, n)$] and its extensions. In particular, he observed that (i) Lanke's estimator, even though it improves over the estimator $e(s(n)|Y)$, may itself turn out to be inadmissible, and (ii) if the estimator $e(s(n)|Y)$ is the sample mean (or *HTE*) then there may not exist an extended sampling design such that Lanke's estimator based on e^* is again the sample mean (or *HTE*). He also showed that when $e(s(n)|Y)$ is the sample mean and the extended sampling design is $SRSWOR(N-n, k)$, Lanke's estimator will again be the sample mean if and only if the initial sampling design $FS(N, n)$ is itself *SRSWOR*.

Some other features of uses of additional resources are discussed in Sengupta *et al.* (1987). Another interesting and related paper on finding admissible estimators is Patel and Dharmadhikari (1977).

4. Variance comparison and effect of sample size

With reference to the example taken up above, we will examine the effect of sample sizes n versus $n + k$ by computing 'Efficiency per Unit Observation'. Note that in general terms, efficiency is defined as the reciprocal of variance and efficiency per unit observation is to be computed as reciprocal of

$$n \times V[e(s(n)|Y)] \text{ as against } (n+k) \times V[e(s(n+k)|Y)]. \quad (20)$$

We fix the population Y-values as

[1,2,3, ...,10] with a total of 55 and mean of 5.5.

We now opt for the *HTE* [for the population total] based on the original design. In Table 3, we display all the initial samples and the *HTE*-values based on them. Also, we show the corresponding probabilities.

Table 3: *s*(*n*) *P*(s(*n*)) *e*(*s*(*n*)|*Y*)

Sl. No.	*s*(*n*)	*P*(s(*n*))	*e*(*s*(*n*)|*Y*)
1.	(1,2,3,4,5)	0.075	38.3934
2.	(1,3,5,8,10)	0.105	53.6526
3.	(1,4,6,7,9)	0.165	48.2112
4.	(4,6,7,8,10)	0.135	57.8106
5.	(2,3,6,9,10)	0.145	59.8600
6.	(3,4,7,8,10)	0.175	54.5083
7.	(5,6,7,8,9)	0.180	64.9370
8.	(2, 4, 6, 7, 8)	0.020	48.2868

Computations yield for the *HTE* of the population total based on the initial design:

(1) *E* [*HTE*] = 55.1453

(2) *V*(*HTE*) = 52.6695

Next, towards computation of Lanke's estimator, we obtain the following:

First, we show *s*(*n* + *k*), next follows *e*(*s*(*n* + *k*)), lastly we show *P*(*s*(*n* + *k*)).

1 (1,2,3,4,5,6,7) *e*(1,2,3,4,5) 0.0300
2 (1,2,3,4,5,6,9) *e*(1,2,3,4,5) 0.0225
(1, 2) *combined* *e*(1, 2, 3, 4, 5) = 38.3934 0.0525

3 (1,2,3,4,5,8,10)∗ *e*(1,2,3,4,5) 0.0225
4 (1,2,3,4,5,8,10)∗ *e*(1,3,5,8,10) 0.0735
(3,4) *combined* (1, 2, 3, 4, 5, 8, 10)∗ [0.0225×*e*(1,2,3,4,5)+
0.0735×*e*(1,3,5,8,10)]/0.0960=50.0762 0.0960

5 (1,3,4,5,7,8,10)∗∗ *e*(1,3,5,8,10) 0.0315
13 (1,3,4,5,7,8,10)∗∗ *e*(3,4,7,8,10) 0.0700
(5,13) *combined* (1,3,4,5,7,8,10)∗∗ [0.0315×*e*(1,3,5,8,10)+
0.070×*e*(3,4,7,8,10)]/0.1015=54.2427 0.1015

6 (1,2,3,4,6,7,9)) *e*(1,4,6,7,9) = 48.2112 0.0825
7 (1,4,5,6,7,8,9)∗∗∗ *e*(1,4,6,7,9) 0.0825
16 (1,4,56,7,8,9)∗∗∗ *e*(5,6,7,8,9) 0.0540
(7,16) *combined* (1,4,5,6,7,8,9)∗∗∗ [0.0825×*e*(1,4,6,7,9) +
0.0540 × *e*(5, 6, 7, 8, 9)]/0.1365 = 54.8280 0.1365

8 (3,4,6,7,8,9,10))∗∗∗∗ *e*(4,6,7,8,10) 0.0405
15 (3,4,6,7,8,9,10)∗∗∗∗ *e*(3,4,7,8,10) 0.0350
(8,15) *combined* (3,4,6,7,8,9,10))∗∗∗∗ [0.0405×*e*(4,6,7,8,10)+
0.035 × *e*(3, 4, 7, 8, 10)]/0.0755 = 56.2797 0.0755

9 (3,4,5,6,7,8,10) *e*(4,6,7,8,10) = 57.8106 0.0945

10	(2,3,4,5,6,9,10)	*e*(2,3,6,9,10)	0.0870
11	(2,3,5,6,7,9,10)	*e*(2,3,6,9,10)	0.0435
12	(2,3,5,6,8,9,10)	*e*(2,3,6,9,10)	0.0145
(10, 11, 12)	*combined*	*e*(2, 3, 6, 9, 10) = 59.86	0.1450
14	(2,3,4,7,8,9,10)	*e*(3,4,7,8,10) = 54.5083	0.0700
17	(2,5,6,7,8,9,10)	*e*(5,6,7,8,9)	0.0900
18	(3,4,5,6,7,8,9)	*e*(5,6,7,8,9)	0.0360
(17, 18)	*combined*	*e*(5, 6, 7, 8, 9) = 64.9370	0.1260
19	(2,3,4,6,7,8,9)	*e*(2,4,6,7,8) = 48.2868	0.0200

Remark 4: It may be noted that we started with a total of 8 samples for the sample size $n = 5$ and after extension, we ended up with a total of 19 samples. However, for the estimator in the above, we have ended up with a total of 11 samples. Computations yield:

(1) $E[e(n + k)|Y] = 55.1453$

(2) $V(e(n + k)|Y) = 38.2284$

Therefore, Lanke's estimator performs better with the use of additional units. Finally, referring to (4.1), we work out efficiency of the extended estimator by comparing 5 × *V* (*HTE*) with 7 × *V* (*extended estimator*). The quantities are respectively 263.3475 and 267.5988. Therefore, according to this criterion, Lanke's extension formula fails to provide a more efficient estimator.

5. Concluding remarks

Mr. Sharma [Research Scholar in Statistics] and Dr. Singh [Statistics Faculty] express their gratitude to Prof. K. K. Singh Meitei, Head, Department of Statistics, Manipur University, Imphal, for providing excellent academic atmosphere towards conducting collaborative research and for creating opportunities for Prof. Sinha's multiple visits for collaborative research with the faculty and students of this department.

We raised the issue of effective use of additional resources. In general terms, Lanke's estimator serves this purpose. However, this estimator itself may not be admissible in the class of competing estimators [Sengupta *et al*. (1987)]. Further, though variance reduction is achieved, efficiency per unit observation may not necessarily increase with enhanced resources. This area of research still holds rich rewards for those who wish to venture into the perplex question of profitable use of additional resources.

References

Bagchi, S. B. and Sinha, B. K. (2022). Some inferential aspects of mixture sampling designs. *Thai Statistician*, **20**, 233 - 239.

Lanke, J. (1975). *Some Contributions to Theory of Survey Sampling*. Ph.D. Thesis, University of Lund, Sweden.

Patel, H. C. and Dharmadhikari, S. W. (1977). On linear invariant unbiased estimators in survey sampling. *Sankhya C*, **39**, 21-27.

Prabhu-Ajgaonkar, S. G. (1967). The effect of increasing sample size on the precision of an estimator. American Statistician, **21**, 26-28.

Sengupta, S. (1982). *Further Studies on Some Strategies in Sampling Finite Populations*. Unpublished Ph.D. Thesis, Calcutta University.

Sengupta, S. N., Sinha, B. K. and Sastry, G. P. (1987). Some inferential aspects of finite population sampling with additional resources. *Journal of Statistical Planning and Inference*, **16**, 203 - 211.

Sinha, B. K. (1980). *On the Concept of Admissible Extensions of Sampling Designs and Some Related Problems*. Technical Report, No. 14/80, Stat-Math. Division, Indian Statistical Institute, Calcutta.