COMPARISON OF ESTIMATES ON TWO OCCASIONS INVOLVING STRATIFICATION USING SUCCESSIVE SAMPLING

Olayiwola O.M., Apantaku F. S., Ajayi A.O.

Department of Statistics, College of Physical Sciences, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria

Corresponding author: Olayiwola O. M., laniyimathew@yahoo.com

ABSTRACT: Population with large number of elements remains unchanged in several occasions but the value of units changes. There are existing works on changes in population estimates from one occasion to another using double sampling approach. There is dearth of information on the estimation of population estimate using double sampling approach with stratification. This study proposed an estimator to determine the population estimates of two occasions involving stratification using successive sampling with two different secondary data. A random sample of size 120 was selected from 260 prices of milled rice and 40 from 97 immunized children on the first occasion. On the second occasion, the samples were stratified and subsamples of 60 from the price of milled rice and 20 from the immunized children were selected from each stratum which was then supplemented with a post stratified fresh sample of 60 and 40 for both data respectively. The variance of double sampling estimator and minimum variance for the successive sampling for the existing and proposed estimators were obtained. The variance of the existing estimator for double sampling for the price of milled rice and the number of immunized children were 40,914.20 and 1,870,842.00 respectively, while the variance of the proposed estimator were 26,198.11 and 799,064.40 respectively. The minimum variances for the existing estimator for the price of milled rice and the number of immunized children were 25,572.38 and 677,564.40 respectively, while the minimum variances for the proposed estimator were 21,513.36 and 577,275.30 respectively. The variance of the proposed estimator is less than that of the existing estimator, hence more efficient.
KEYWORD: Estimator Matched and Unmatched part, Double sampling, Stratification.

1. INTRODUCTION

It is often seen that sample surveys are not limited to one time inquiries. If the value of study character of a finite population is subject to change over time, a survey carried out on a single occasion will provide information about the characteristics of the surveyed population for the given occasion only and cannot give any information on the nature or the rate of change of the characteristic over all occasions or more recent occasion. Data regarding changing properties of the populations of cities or counties, such as unemployment statistics, are collected regularly on a sample basis, to estimate the changes from one occasion to the next or to estimate the average over a certain period. An important aspect of continuous surveys is the structure of the sample on each occasion. To meet these requirements, successive sampling provides a strong tool for generating the reliable estimates at different occasions. This work aims to compare estimates of successive sampling, when there is stratification of the unmatched portion of the second occasion and when there is the full stratification of unmatched and matched portion of the second occasion.

2. DATA

Two sets of secondary data were considered. The first set of data is the record of Number of Children Immunized against polio in Local Government/Local Council Development Areas of Lagos state. It was collected from the Abstract of Local Government Statistics (2012), Lagos Bureau of Statistics, ministry of Economic, Planning and Budget. The second set is the price of milled rice, collected from five different markets in Nigeria (Giwa in Kaduna state, Bida in Niger state, Bodija in Oyo state, Monday market in Borno and Umuahia in Abia state). These data were broadcasted across different radio stations in Nigeria and were collated by the department of Agricultural Extension and Rural Development, Federal University of Agriculture, Abeokuta.

3. LITERATURE

The problem of sampling on two successive occasions with a partial replacement of units was first considered by Jessen ([Jes42]). Yates ([Yat49]) extended Jessen's scheme to the situation where the population of the variable is estimated on each one of the h≥2 occasions from a rotational sample design. Kulldufo ([Kul63]) modified Jessen's scheme of the sampling by selecting the unmatched sample from the units not selected on the first occasion. He considered in details the optimum choice of the matching fraction under the most general form of cost function apart from fixed costs. Sen ([Sen72]) generalized Jessen's work by using a double sampling multivariate ratio estimate using P-auxiliary Variable (p >1) from the matched portion of the sample. Expressions for optimum matching fraction and the combined estimate and its error have also been derived. Okafor ([Oka87]) compared some estimators of the population total in two–stage successive sampling using auxiliary variable.

Artes and Garcia ([AG01]) worked on estimating the current mean in successive sampling using a product estimate. Garcia Luengo ([GA04]) considered the problem of estimation of a finite population mean and for the current occasion based on the sample selected over two occasions for the case when, several auxiliary variables are correlated with the main variable. A double sampling multivariate product estimate from the matched portion of the sample is presented. Expressions for optimum estimator and its error have been derived. The gain in efficiency of the combined estimate over the direct estimate using no information gathered on the first occasion was computed. Artes et al. ([ARA05]) also worked on successive sampling using a product estimate but, they considered the case when the auxiliary variables are negatively correlated and double sampling product estimate from the matched portion of the sample was presented.

Expression for optimum estimator and its variance have been desired. Reuda et al. ([R+06]) talked on estimating quantitative under sampling on two occasions with arbitrary sample designs. Rueda et al. ([RMA07]) further extended the work on successive sampling in estimating quartiles with p-auxiliary variables.

They mainly discussed the estimation of quartiles for the current occasion based on sampling on two occasions and using p-auxiliary variables obtained from the previous occasions. A multivariate ratio estimate from the matched portion was used to provide the optimum estimate of a quartile by weighing the estimate inversely to derive optimum weight.

Trivedi and Shukla ([TS08]) looked into the efficient estimator in successive sampling using post stratification. They stated that it is often seen that a population having large number of elements remains unchanged in several occasions but the value units change. In their work, they introduced an estimator under successive survey, the estimator is unbiased and efficient over post stratification estimations, the minimum variance of the optimum estimator was derived and comparative study incorporated. Housila et al. ([H+10]) looked into estimation of population variance in successive sampling and proposed a class of estimators of finite population variance in successive sampling on two occasions and analyzed its properties.

Housila et al. ([H+07]) looked into the problem of estimating a finite population quantile in successive sampling on two occasions. They aimed at providing the optimum estimates by combining.

(i) Three double sampling estimator viz. ratio-type, product-type and regression-type estimator, from the matched portion of the sample and.
(ii) A sample quantile based on a random sample from the unmatched portion of the sample on the second occasion.

A simulation study was carried out in order to compare the three estimators and it is found that the performance of the regression-type estimator is the best among all the estimators discussed.

Olayiwola et al. ([O+12]) worked "on double sampling approach for comparing estimates of students' enrolment in Oyo state public secondary Schools". Also Amahia et al. ([E+11]) successive sampling for regression estimation was used to determine the current estimate of the mean, minimum variance, maximum precision, estimate of change between the two successive occasions under consideration and estimate of average over the period of the two occasions. Anieting and Ezegwu ([AE13]) also used successive sampling to determine the current estimate of the mean, minimum variance, estimate of change between the two successive occasions and estimate of average over the period of the two occasions. The estimates of change and estimate of average over time give their optimum variance when $\rho = 0$ and $\mu = 1$ under varying values of ρ and μ. Also the variance where found to be smaller in the second occasion as compared to the variance in the first occasion.

Singh and Priyanka ([SP10]) emphasized the role of several varying auxiliary variates at both the occasions to improve the precision of estimates at current occasion in two-occasion successive sampling. Two different efficient estimators were proposed and their theoretical properties examined. Relative comparison of efficiencies of the proposed estimators with the sample mean estimator when there is no matching from previous occasion, and the optimum successive sampling estimator when no auxiliary information is used have been incorporated. Empirical studies are significantly justifying the composition of proposed estimators. Kowalczyk ([Kow03]) gave the theory for the population total estimation on the current occasion under two stages sampling with SRSWOR in both stages and rotation unit in the second stage. The paper considered a problem when the same population of *N* elements is sampled on two successive occasions using two-stage sampling design.

4. METHODOLOGY

Consider a population of size N that is sampled over two occasions. On the first occasion a sample of n units is selected from N units in the population by SRSWOR. On the second occasion a sub sample of λn units, ($0<\lambda<1$), is selected from the first occasion sample by SRSWOR. This is supplemented by a fresh sample of μn units ($\mu+\lambda=1$), selected by SRSWOR from N or from (N-n) units. Let X and Y represent the variables measured in the first and second occasions respectively.
Schematically, the selection can be represented as below:

$\bar{x}''$	$\bar{x}'$	
	$\bar{y}'$	$\bar{y}''$

where $\bar{x}''$ and $\bar{x}'$ are the means of the matched and unmatched samples on the first occasion, while $\bar{y}'$ and $\bar{y}''$ are the means of the matched and unmatched samples on the second occasion.

4.1. STRATIFICATION OF SAMPLE IN SECOND OCCASION

Trivedi and Shukla ([TS08]) stratify the unmatched part of the second occasion and proposed an unbiased estimator for the mean of the second occasion. He proposed an estimator that is unbiased and efficient over post-stratification estimator, also minimum variance of the optimum estimator was derived.
For Estimating $\bar{Y}_2$ based on successive sampling using post stratification scheme, an estimator was proposed;

$$\bar{y}_{pss} = \phi\bar{y}''_{2s} + (1-\phi)\bar{y}'_{12} \qquad 4.1$$

where ϕ is a constant,
$\bar{y}''_{2s}$ is the mean based on the post stratified n''_2 units $= (\sum_{i=1}^{k} W_i \bar{y}''_{2i})$,
$\bar{y}'_{12}$ is the mean based on the matched sample n'_2, which is termed as $\bar{y}'_{12} = \bar{y}'_2 + \beta_{21}(\bar{y}_1 - \bar{x}'_1)$,
where β_{21} the regression coefficient of the variate of the second occasion on the variate of the first occasion is assumed to be known. Then the proposed estimator becomes,

$$\bar{y}_{pss} = \phi\left(\sum_{i=1}^{k} W_i \bar{y}''_{2i}\right) + (1-\phi)\left[\bar{y}'_2 + \beta_{21}(\bar{y}_1 - \bar{x}'_1)\right] \qquad 4.2$$

The estimator $\bar{y}_{pss}$ is unbiased for $\bar{Y}_2$
The variance of the estimator is given to be

$$V(\bar{y}_{pss}) = \phi^2\left[\left(\frac{1}{n}-\frac{1}{N}\right)\sum_{i=1}^{k} W_i S_{2i}^2 + \frac{(N-n)}{(N-1)n^2}\sum_{i=1}^{k}(1-W_i)S_{2i}^2\right] + (1-\phi)^2\left[\left(\frac{1}{n'_2}-\frac{1}{n_1}\right)S_2^2(1-\rho_{2i}^2) + \left(\frac{1}{n_1}-\frac{1}{N}\right)S_2^2\right] - 2\phi(1-\phi)\frac{S_{2i}^2}{N}$$

With optimum value of ϕ to be

$$\phi_{opt} = \frac{Var(\bar{y}'_{12}) - Cov(\bar{y}''_{2s}, \bar{y}'_{12})}{Var(\bar{y}'_{12}) + Var(\bar{y}''_{2s}) - 2Cov(\bar{y}''_{2s}, \bar{y}'_{12})} \qquad 4.3$$

Then

$$V(\bar{y}_{pss})_{opt} = \frac{Var(\bar{y}''_{2s})Var(\bar{y}'_{12}) - [Cov(\bar{y}''_{2s}, \bar{y}'_{12})]^2}{Var(\bar{y}'_{12}) + Var(\bar{y}''_{2s}) - 2Cov(\bar{y}''_{2s}, \bar{y}'_{12})} \qquad 4.4$$

4.2. PROPOSED SCHEME (SEPARATE DOUBLE SAMPLING REGRESSION)

Motivated by Trivedi and Shukla ([TS08]), this work proposed a scheme, that instead of stratifying only the unmatched portion of the second occasion, we try to stratify both the matched and the unmatched portion of the second occasion. Based on this, an estimate of μ_2 is being proposed as:

4.2.1. Estimation Strategy

It is assumed the population of size N remains unchanged over both occasions population remains unchanged. Let a population be of size N that is sampled over two occasions. Assume that the size of the $n_1=n_2$. The first occasion sample of size n_1 using SRS without replacement is stratified into L strata. Let denote the number of units in the sub sample falling into stratum

$(h=1,2,\dots,l)$. $\sum_{h=1}^{L} n_{1h} = n_1 = \{n_{11}, n_{12}, \dots, n_{Lh}\}$

denoting the resulting partitions of the first occasion sample. In the second sampling, m_h of the units in the first occasion are selected from strata independently *(h=1,2,...l)*, and $\sum_{h=1}^{l} m_h = m$ (m for matched). The remaining u units (u for unmatched) are discarded and replaced by a new selection from the units not previously selected will now be post stratified as Trivedi and Shukla ([TS08]) proposed.

$\bar{y}_{2m}$ - The population mean on the matched sample on the second occasion. $\bar{y}_{mh} = \frac{1}{m_h}\sum_{i=1}^{m_h} y_{mhi}$ The sample mean based on n_{2m} units observed on the second occasion and common with the first occasion, $\bar{y}_{2u}$ is the mean based on the post stratified n_{2u} units ($\sum_{i=1}^{k} W_h \bar{y}_{2h}$), on the second occasion, $\bar{x}_{mh}$ is the sample mean based n_{1m} units common to both the occasions and observed on the first occasion, $n_1 = n_{1m} + n_{1u}$ is the sample units observed on the first occasion, $n_2 = n_{2m} + n_{2u}$ is the units observed on the second occasion and S_i^2 = The population mean square error for *the i-th* occasion, i=1, 2,…, n.

4.2.2. The estimator

For Estimating $\bar{Y}_2$ based on successive sampling using separate double sampling regression, an estimator was proposed;

$$\bar{y}_{dss} = \phi \bar{y}_{2u} + (1-\phi)\bar{y}_{2m} \qquad 4.5$$

where ϕ is a constant. Motivation of taking this constant is taken from Trivedi M. and D Shukla ([TS08]).

Theorem 4.3
The proposed estimator $\bar{y}_{dss}$ is unbiased for $\bar{Y}_2$

Proof: It is clearly seen that $\bar{y}_{dss}$ is an unbiased estimator of $\bar{Y}_2$, since $\bar{y}_{2m}$ is a double sampling estimator ([Coc77]) and is unbiased for the population mean. Also $\bar{y}_{2u}$ is SRSWOR estimator (Cochran, [Coc77]) and is also unbiased for the population mean.

Theorem 4.4
The variance of the estimator is

$$V(\bar{y}_{dss}) = \phi^2(V(\bar{y}_{2u})) + (1-\phi)^2 V(\bar{y}_{2m}) + 2\phi(1-\phi)COV(\bar{y}_{2m}, \bar{y}_{2u})$$

Proof: We have the estimator

$$\bar{y}_{dss} = \phi \bar{y}_{2u} + (1-\phi)\bar{y}_{2m}$$

Then

$$V(\bar{y}_{dss}) = \phi^2(V(\bar{y}_{2u})) + (1-\phi)^2 V(\bar{y}_{2m}) + 2\phi(1-\phi)COV(\bar{y}_{2m}, \bar{y}_{2u}) \qquad 4.6$$

The estimator $\bar{y}_{2m}$ is a separate double sampling regression mean then

$$\bar{y}_{2m} = \sum_{h=1}^{l} W_h\left(\bar{y}_{1mh} + \beta_{21h}\left(\bar{X}_h - \bar{x}_{1mh}\right)\right) \qquad 4.7$$

$$V(\bar{y}_{2m}) = \sum w_h'^2\left(\frac{1}{n_h} - \frac{1}{N_h}\right)S_{yh}^2 + \sum_{h=1}^{l} w_h^2\left(\frac{1}{m_h} - \frac{1}{n_h}\right)S_{yh}^2(1-\rho_h^2) \qquad 4.8$$

Proof: $\bar{y}_{2m}$ is a double sampling estimator, a separate double sampling regression is being proposed. Where the first phase is a stratified random sampling,

$$V(\bar{y}_{2m}) = E_1V_2(\bar{y}_{2m} / firstphase) + V_1E_2(\bar{y}_{2m} / firstphase) \qquad 4.9$$

$$V(\bar{y}_{2m}) = \sum w_h'^2\left(\frac{1}{n_h} - \frac{1}{N_h}\right)S_{yh}^2 + \sum_{h=1}^{l} w_h^2\left(\frac{1}{m_h} - \frac{1}{n_h}\right)S_{yh}^2(1-\rho_h^2)$$

Also the estimator $\bar{y}_{2u}$ is a post stratified mean then

$$V[\bar{y}_{2u}] = V\left[\sum_{h=1}^{l} W_h \bar{y}_{2h}\right] =$$

$$\left(\frac{1}{u} - \frac{1}{N}\right)\sum_{h=1}^{l} W_h S_{2h}^2 + \frac{(N-u)}{(N-1)u^2}\sum_{h=1}^{l}(1-W_h)S_{2h}^2$$

And

$$COV(\bar{y}_{2m}, \bar{y}_{2u}) = -\frac{S_2^2}{N}$$

Now putting the equations together we have

$$V(\bar{y}_{dss}) = \phi^2\left(\frac{1}{u}-\frac{1}{N}\right)\sum_{h=1}^{l} W_h S_{2h}^2 + \frac{(N-u)}{(N-1)u^2}\sum_{h=1}^{l}(1-W_h)S_{2h}^2$$
$$+(1-\phi)^2\left(\sum w_h'\left(\frac{1}{n_h}-\frac{1}{N_h}\right)S_{yh}^2 + \sum_{h=1}^{l} w_h\left(\frac{1}{m_h}-\frac{1}{n_h}\right).S_{yh}^2(1-\rho_h^2)\right) - 2\phi(1-\phi)\frac{S_2^2}{N} \quad 4.10$$

4.5. OPTIMUM CHOICE

Getting optimum value of ϕ we have to differentiate 3.19 with respect to ϕ and equating to zero, then we have

$$\phi_{opt} = \frac{Var(\bar{y}_m) - Cov(\bar{y}_{2u}, \bar{y}_m)}{Var(\bar{y}_{2m}) + Var(\bar{y}_{2u}) - 2Cov(\bar{y}_{2u}, \bar{y}_{2m})} \quad 4.11$$

Putting this value in the variance expression we have

$$V(\bar{y}_{dss})_{opt} = \frac{Var(\bar{y}_{2u})Var(\bar{y}_{2m}) - [Cov(\bar{y}_{2u}, \bar{y}_m)]^2}{Var(\bar{y}_{2m}) + Var(\bar{y}_{2u}) - 2Cov(\bar{y}_{2u}, \bar{y}_{2m})} \quad 4.12$$

5. DATA PRESENTATION AND ANALYSIS

5.1. DESCRIPTIVE ANALYSIS

In population one, the data set considered is the prices of milled rice in Nigeria. This was collected from five different markets in different parts of Nigeria. These markets include Giwa in Kaduna state, Bida in Niger state, Bodija in Oyo state, Monday market in Borno and Umuahia in Abia state. The average price was collated per week for a year for each market. The price for 2011 and 2012 successive years was considered in this study.
Descriptive Analysis of Population one:

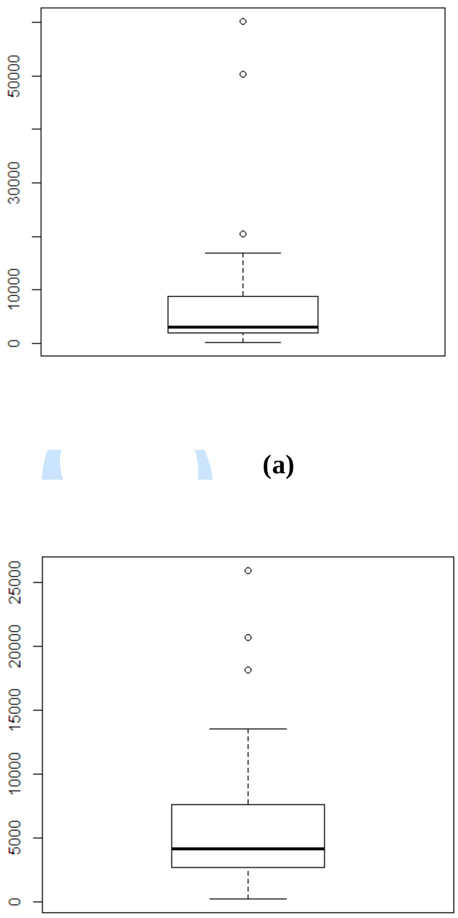
Location of markets	2011	2012
Giwa	13446.15	14328.85
Bida	11457.21	13542.79
Bodija	11177.89	12950.00
Monday	13057.69	15480.77
Umuahia	11501.92	12675.00

Fig. 1. Mean per market

In Population two, the data set considered is the number of immunized children against polio in Lagos state. The data is a secondary data collected from abstracts of local government statistics of Lagos state. The average number of immunized children for 2011 and 2012 successive years was considered in this study.

Sex	2011	2012
Male	7395.26	5771.91
Female	6817.30	6343.02

Fig. 2. Mean per sex

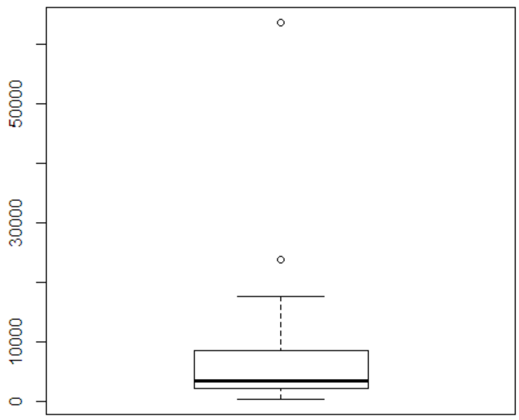


(a)

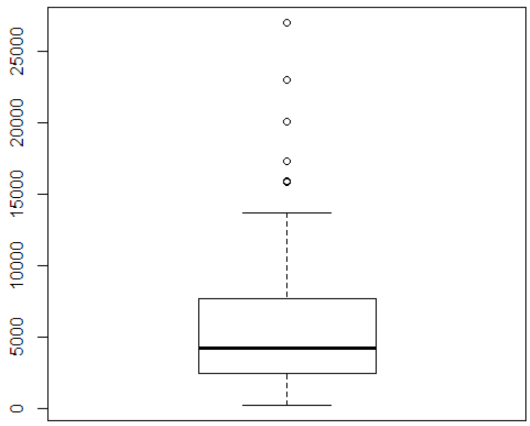
(b)

Fig. 3. Box plots for immunized male children for 2011 and 2012 respectively

From the box and whiskers plot, the outliers (dots) as seen in the figure above shows that there are some locations were more than normal numbers of male children were immunized against polio.
In the 2012 box plot it shows that there are four locations where more than normal numbers of female children were immunized, for 2011 box plot there are just two outliers with one extremely high, this high point in the number of immunized female children also correspond with the number extremely high male children in 2011.



(a)



(b)

Fig. 4. Box plots for immunized female children for 2011 and 2012 respectively

5.2. SUCCESSIVE SAMPLING ANALYSIS

5.2.1. Population One

The total population is 260 for both occasions, a simple random sampling of size 120 from the population on the first occasion, in selecting the second sample, 60(m for matched) of the units in the first sample are retained for matching. The remaining 60(u for unmatched) units are discarded and replaced by a new selection of 60 from the units not previously selected.

5.2.2. Trivedi and Shukla (2008) Scheme

The unmatched(u) is further stratified based on each market locations and the Travedi and Shukla ([TS08]) estimator is estimated based on the following characteristics gotten from the data.

5.2.3. The proposed Scheme

Based on the characteristics explained in 5.1.2.1, this work attempts to stratify both the matched and unmatched portion of the second occasion. The stratification is done based on market locations, thereby having five strata. In selecting the second sample, the whole of first occasion is stratified then 60(m for matched) of the units of the stratified first sample are retained and matching is done thereby succeeded in stratifying both the matched portion of the first and second occasion. The remaining 60(u for unmatched) units are discarded and replaced by a new selection of 60 from the units not previously selected. The unmatched (u) is further stratified based on each market locations and the proposed estimator is estimated based on the following characteristics gotten from the data.

5.2.4. Summary of parameters

$N=260,\ n=120,\ \mu=\lambda=0.5,\ m=n\lambda=60,\ u=n\mu=60,$
$W_1=W_2=W_3=W_4=W_5=0.2,\ N_1=N_2=N_3=N_4=N_5=52,$
$n_1'=23,\ n_2'=28,\ n_3'=20,\ n_4'=21,\ n_5'=28,$
$n_1=12,\ n_2=14,\ n_3=10,\ n_4=10,\ n_5=14$
$w_1'=0.19,\ w_2'=0.233,\ w_3'=0.167,$
$w_4'=0.175,\ w_5'=0.233$

Table 1

Parameters	Existing	Proposed
ϕ_{opt}	0.56	0.25
μ_{2u}	14044.75	14044.75
μ_{2m}	13567.21	14121.39
μ_2	13834.63	14092.94
$V(\mu_{2u})$	34945.25	34945.25
$V(\mu_{2m})$	40914.2	26198.11
$V(\mu_2)$	25572.38	21513.36

The table above shows the parameters of the proposed estimators and Trivedi and Shukla ([TS08]), and it can be seen that the proposed ϕ_{opt}, $V(\mu_{2m})$ is lesser than the earlier proposed by Trivedi and Shukla ([TS08]). Also the proposed $V(\mu_2)$ is lesser than the estimator proposed by Trivedi and Shukla ([TS08]). It can be said that the proposed estimator is more efficient than Trivedi and Shukla ([TS08]), the relative efficiency is

1.19>1. It should be noted $\bar{y}_{2m} = \bar{y}'_2 = \mu_{2m}$ which is the mean of the matched portion of the second occasion and $V(\mu_{2m}) = V(\bar{y}'_2) = V(\bar{y}_{2m})$ which is also the variance of the matched portion of the second occasion, also $\bar{y}_{2u} = y''_{2s} = \mu_{2u}$ is the mean of the unmatched portion of the second occasion, $V(\mu_{2u}) = V(\bar{y}''_{2s}) = V(\bar{y}_{2u})$ is the variance of the unmatched portion of the second occasion. Lastly $\mu_2 = \bar{y}_{pss} = \bar{y}_{dss}$ is the mean of the second occasion $V(\mu_2) = V(\bar{y}_{pss}) = V(\bar{y}_{dss})$ is the optimum variance of the second occasion obtained after inserting ϕ_{opt} into the variance of the second occasion.

5.3. POPULATION TWO

The total population is 93 for both occasions, a simple random sampling of size 40 from the population on the first occasion, in selecting the second sample, 20(m for matched) of the units in the first sample are retained for matching. The remaining 20(u for unmatched) units are discarded and replaced by a new selection of 20 from the units not previously selected.

5.3.1. Trivedi and Shukla Scheme

The unmatched (u) is further stratified based on sex of child and the Trivedi and Shukla ([TS08]) estimator is estimated based on the following characteristics gotten from the data.

5.3.2. Proposed Scheme

Based on the characteristics explained in 5.1.2.1, this work attempts to stratify both the matched and unmatched portion of the second occasion. The stratification is done based on market locations, thereby having five strata. In selecting the second sample, the whole of first occasion is stratified then 20(m for matched) of the units of the stratified first sample are retained and matching is done thereby succeeded in stratifying both the matched portion of the first and second occasion. The remaining 20(u for unmatched) units are discarded and replaced by a new selection of 20 from the units not previously selected. The unmatched (u) is further stratified based on each market locations and the proposed estimator is estimated based on the following characteristics gotten from the data.

5.3.3. Summary of parameters

N=93, n=40, $\mu=\lambda$=0.5, $m=n\lambda$=20, $u=n\mu$=20, W_1=0.495, W_2=0.505, N_1=46, N_2=47, $n'_1 = 22$, $n'_2 = 18$, $n_1 = 11$, $n_2 = 9$, $w'_1 = 0.55$, $w'_2 = 0.45$

Table 2

Parameters	Existing	Proposed
ϕ_{opt}	0.80	0.53
μ_{2m}	6249.21	6922.02
μ_{2u}	3928.55	3928.55
μ_2	4391.81	4525.74
$V(\mu_{2u})$	751385.50	751385.50
$V(\mu_{2m})$	1870842.00	799064.40
$V(\mu_2)$	677564.40	577275.30

The table above shows the parameters of the proposed estimators and Trivedi and Shukla ([TS08]), and it can be seen that the proposed ϕ_{opt}, $V(\mu_{2m})$ is lesser than the earlier proposed by Trivedi and Shukla ([TS08]). Also the proposed $V(\mu_2)$ is lesser than the estimator proposed by Trivedi and Shukla ([TS08]). It can be said that the proposed estimator is more efficient than Trivedi and Shukla ([TS08]), the relative efficiency is 1.17>1. It should be noted $\bar{y}_{2m} = \bar{y}'_2 = \mu_{2m}$ which is the mean of the matched portion of the second occasion and $V(\mu_{2m}) = V(\bar{y}'_2) = V(\bar{y}_{2m})$ which is also the variance of the matched portion of the second occasion, also $\bar{y}_{2u} = y''_{2s} = \mu_{2u}$ is the mean of the unmatched portion of the second occasion, $V(\mu_{2u}) = V(\bar{y}''_{2s}) = V(\bar{y}_{2u})$ is the variance of the unmatched portion of the second occasion. Lastly $\mu_2 = \bar{y}_{pss} = \bar{y}_{dss}$ is the mean of the second occasion $V(\mu_2) = V(\bar{y}_{pss})_{opt} = V(\bar{y}_{dss})_{opt}$ is the optimum variance of the second occasion obtained after inserting ϕ_{opt} into the variance of the second occasion.

CONCLUSION

From the analysis, the variance of the matched part $V(\mu_{2m})$ from both data sets was lesser than the one proposed by Trivedi and Shukla ([TS08]), it can be concluded that stratification of the matched part of the second occasion gives a smaller variance, for the purpose of practical application it is recommended the matched part of the second sample should be stratified. Since the variance of the matched part is lesser, it follows consequently that the overall variance of the second occasion will be smaller.

CONTRIBUTION TO KNOWLEDGE

1. The study derived a new estimator in double sampling by stratifying initial sample n' and subsample n thereby giving a separate double sample regression estimate.
2. This study further extended this derived estimator to successive sampling by stratifying all the units of the second occasion.
3. This estimator derived was compared with other existing estimator and found to be more efficient.

REFERENCES

[Ama09] **G. N. Amahia** - Unpublished University of Ibadan, Nigeria lecture notes on STA 764. *Design and analysis of sample survey*, 2009.

[AE13] **A. E. Anieting, V. O. Ezugwu** - *Estimation of the Current Occasion Parameters Using Successive Sampling Approach*. American Journal of Engineering Research (AJER) e-ISSN : 2320-0847 p-ISSN : 2320-0936 Volume-02, Issue-09, pp-204-207, 2013.

[AG01] **E. Artes, A. V. Garcia** - *Estimating the current mean in successive sampling using a product estimate*. Conference on agricultural and environmental statistical application in Rome, XLIII-1- XLIII-2, 2001.

[ARA98] **E. Artes, M. Reuda, A. Arcos** - *Successive sampling using a product estimate, applied sciences and the environment*. Computational mechanics publication, 85 – 90, 1998.

[ARA05] **E. Artes, M. Rueda, A. Arcos** - *Successive sampling using a product estimate when the auxiliary variables are negatively correlated*. Applied science and environment, page 6, 2005.

[Coc77] **W. G. Cochran** - *Sampling techniques*, John Weily and Sons Inc. New York, 1977.

[E+11] **S. O. Eze, G. N. Amahia, O. M. Olayiwola, A. A. Adewara** - *Efficiency of Second Occasion over First Occasion on Successive Sampling For Regression Estimation*. Journal of Mathematics Research vol. 3, No. 2, 2011.

[GA04] **Garcia L., Artes A.V.** - *On Estimation of population in the sampling variance on several occasions*, survey methodology, volume 6, No 1, 9-22, 2004.

[GK77] **S. C. Gupta, V. K. Kapoor** - *Fundamentals of Applied Statistics*. Sultan Shand and Sons, New Delhi, 1977.

[HS79] **D. Holt, T. M. F. Smith** - *Post-stratification*, J.R. Stat. Soc., A, 142, 33-36, 1979.

[H+07] **P. S. Housila, T. Ritech, S. Sarjinder, K. Jong-Min** - *Quantile estimation in successive sampling*. Journal of the Korean statistical society, 36(4) 543-556, 2007.

[H+10] **P. S. Housila, T. Ritech, S. Sarjinder, K. Jong-Min** - *Estimation of Population variance in successive sampling*, cda.mrs.Umn.edu/jongmink/research, 2010.

[Jag86] **P. Jagers** - *Post-stratification against bias in sampling*, Int. Stat. Rev., 54, 159-167, 1986.

[Jes42] **R. J. Jessen** - *Statistical investigation of a sample survey for obtaining farm facts*. Lowa agriculture experiment statistical research bulletin, 304, 1942.

[JOT85] **P. Jagers, A. Oden, L. Trulsson** - *Post-stratification and ratio estimator*, Int. Stat. Rev., 53, 221-238, 1985.

[Kau05] **A. Kaustav** - *Some aspects of Successive Sampling*. I.A.S.R.I, Library Avenue, New Delhi-110012, 2005.

[Kow203] **B. Kowalczyk** - *Estimation of the Population Total on the Current Occasion Under Second Stage Unit Rotation Pattern*. Statistics In Transition. Vol. 6, No. 4, pp 503-513, 2003.

[Kul63] **G. Kullduf** - *Some problem of optimum allocation for sampling on two occasions*. Review International statistical institute, 31, 24-57, 1963.

[Oka87] **F. C. Okafor** - *Comparison of estimators of population totals in two-stage successive sampling using auxiliary information*. Survey method. 13 (1), 109-121, 1987.

[O+12] **O. M. Olayiwola, S. A. Agboluaje, O. O. Salawu, J. O. Fantola** - *On double sampling approach for comparing estimates of Students' enrolment in Oyo state Public Secondary Schools*. Elixir Statistics 43, 6994-7003, 2012.

[RMA07] **M. Rueda, J. F. Munoz, A. Arcos** - *Successive sampling in estimation quartile with p-auxiliary variables*. Spring link journal (quality and quantity) pdf (264.4kb), 2007.

[R+06] **M. Rueda, J. F. Munoz, S. Gonzalez, A. Arcos** - *Estimating quartiles under sampling on two occasions with arbitrary sample design*. Computational statistics and data analysis. Vol. 51, issue 12. pp 6596-6613, 2006.

[Sen71] **A. R. Sen** - *Successive sampling with two auxiliary variables*. Sankya B33, 371-378, 1971.

[Sen72] **A. R. Sen** - *Successive sampling p-auxiliary variables*. Ann. Mathematical Statistics. 43, 203-204, 1972.

[Smi91] **T. M. F. Smith** - *Post stratification*, The statistician, 40, 315-323, 1991.

[SC84] **D. Singh, F. S. Chaudhary** - *Theory and analysis of sample survey design: New Age International* (P) Lon (formally) Wiley Eastern Ltd. New Delhi, 1984.

[SP10] **G. N. Singh, K. Priyanka** - *Estimation of population mean at current occasion in presence of several varying auxiliary variates in two-occasion successive sampling*. Statistics In Transition-new series, Vol. 11, No. 1, pp. 105-126, 2010.

[SS73] **R. Singh, B. V. Sukhatme** - *Optimum stratification with ratio and regression methods of estimation*. Ann. Inst. Stat. Math, 25, 627-633, 1973.

[S+84] **P. V. Sukatme, B. V. Sukatme, S. Sukatme, C. Asok** - *Sampling Theory of Survey with Applications*, Iowa State University Press, Indian Society of Agricultural Statistics, New Delhi, 1984.

[TS08] **M. Trivedi, D. Shukla** - *Efficient Estimator in Successive Sampling using Post-Stratification*, 2008.

[Yat49] **F. Yates** - *Sampling methods for censuses and surveys*, Charles Griffin and Co., London, 1949.