Available online at www.elixirpublishers.com (Elixir International Journal)

Statistics

Elixir Statistics 67 (2014) 21836-21845

ISSN: 2229-712X

Optimum Allocation in Stratified and Post Stratified Sampling Designs

Apantaku F.S, Asiribo O.E, Adelakun A.A, Amahia G.M and Olayiwola O.M
Department of Statistics, Federal University of Agriculture, PMB 2240, Abeokuta, Nigeria.

ARTICLE INFO

Article history:
Received: 4 December 2013;
Received in revised form:
28 January 2014;
Accepted: 19 February 2014;

Keywords
Stratification,
Post stratification,
Optimum allocation.

ABSTRACT

One of the main problems in sampling survey is the optimal allocation of resources. The solution of this problem is rather arbitrary due to the fact that no best allocation is defined. In this model, the allocation problem was to find the allocation of a sample to strata which minimizes cost of investigation. The idea of optimal allocation under a multivariate stratified sampling based on an alternative approach earlier worked on by Diaz- Garcia and Ranos–Quiroga was applied. The matrix of the variance-covariances of the vector of the stratified variables was obtained. An emperical data from a household survey conducted in Ogun state was used. The frame consisted 880,970 households in the twenty local government areas (LGA) of Ogun state. Each of the 20 LGAs that made up the state was seen as a stratified cluster. Post stratification in this study ensured that some variables that are suitable for stratification was achieved after selection of the sample. The four characteristics of interest were occupation, income, household size and educational level. The study transformed all variables in the stratified sampling plan by using dummy variables ranging from 1-3. The estimates used in the computation were calculated using statistical software Splus. The post stratification estimator $\bar{Y}_{st}$ does not have the same variance as the stratified sample mean. Stratification produced a gain in precision in the estimation of characteristics of the household survey. The results of the estimates in the study revealed that proportionate allocation can lead to smaller sampling variances than for simple random samples of the same size.

Introduction

Stratification is one of the most widely used techniques in sample survey design serving the dual purposes of providing samples that are representative of major sub-groups of the population and of improving the precision of estimators (Holt and Smith, 1979). The essence of stratification is the classification of the population into sub-populations, or strata, based on some supplementary information, and then the selection of separate samples from each of the strata. The benefits of stratification derive from the fact that the sample sizes in the strata are controlled by the sampler, rather than being randomly determined by the sampling process. One of the main problems in sampling survey is the optimal allocation of resources. Usually, the solution of this problem is rather arbitrary due to the fact that no best allocation is defined. In terms of this model, the allocation problem is to find the allocation of a sample to strata which minimizes cost of investigation subject to a given condition about the sampling error. Stratified random sampling technique is used to improve the accuracy of survey results, or to lower the cost of a survey without losing accuracy (Fuller, 1993).

Successful stratification seeks to maximise homogeneity within strata and heterogeneity between strata. Such grouping has no effect upon the real dispersion that exists in the population of interest. Stratification is merely concerned with the division of overall dispersion into variability within strata and variability from stratum to stratum. When the population of interest is made up of individuals, it is quite common to stratify by such measures as income or age

The method of using stratification is to increase the precision of the sample mean as in contrast to proportionate sampling. It involves the deliberate use of widely different sampling rates for the various strata. Optimum stratification is used when the standard deviations of the population strata are known to differ substantially. This technique is a method of allocating larger size samples to those strata with larger standard deviations. The designation optimum allocation to disproportionate sampling refers to the aim of assigning sampling rates to the strata in such a way as to achieve the least variance for the overall mean per unit of cost (Diaz-Garcia and Cortez, 2006).

Tele:
E-mail addresses: fsapantaku@yahoo.com, samapantaku@yahoo.com

When stratified sampling is used, an allocation that is optimum for one character is generally not so for others (Pirzada and Maqbool, 2003). In classical sampling theory textbooks, only fixed populations are considered (Kish, 1965; Cochran, 1977; Sukhatme *et al*., 1984). The allocation of resources in a stratified sampling design aimed at estimating a population parameter is usually carried out according to one of the two alternatives,

(1) To achieve maximum precision for a given total cost of survey or

(2) To achieve a given precision at a minimum cost.

Post-stratification may be used on the subclasses even if a proportionate sample of the entire population has been selected. Post stratification is an example of improving the estimator by the proper utilization of ancillary sources of information. Post-stratification is a weighting procedure applied mainly, on many surveys. Post-stratification uses case weights to assure that sample totals equal some external total based on the target population. Post-stratification is necessary since stratification variables are unknown prior to sampling and estimates are required. Post-stratification can reduce the variance of estimates if the post-stratification variables are effective.

Arthanari and Dodge (1981), Sukhatme *et al*. (1984) and Diaz-Garcia and Cortez (2008) among many others, proposed the problem of optimum allocation in multivariate stratified random sampling as a deterministic multi-objective mathematical programming problem, by considering an objective function as a cost function subject to restrictions on certain functions of variances or vice versa, i.e., considering the functions of variances as objective and subject to restrictions on costs. Noting that, for the case when the function of costs is taken as the objective function, the problem of optimum allocation in multivariate stratified random sampling is reduced to a classical uni-objective mathematical programming problem.

The objectives of the study are:

1. To compare optimum allocation in stratified sampling with optimum allocation in post stratified sampling.
2. To compare using empirical data the relative efficiency in stratified sampling with post stratified sampling.

Methodology

An effective sampling technique within a population represents an appropriate extraction of useful data which provides meaningful knowledge of the important aspects of the population. However in real surveys, more than one variable are considered and the survey designer is often faced with the choice between several potential interest variables, while each variable separately may be well suited for certain purposes. Stratified sampling is one of the classical methods for obtaining such information.

Survey design

Once stratified sampling has been chosen, it is necessary to determine how to divide the population into strata, and how to allocate the sample to those strata. One decision that must be made is the choice of a variable or variables on which to stratify. Since it is rare to conduct a survey with only one item of interest, the stratification variable or variables are chosen (or constructed) to have a strong correlation with as many items of interest as possible. Methods for construction of optimal stratum boundaries (with the goal of improving the precision of estimates) have been proposed by Dalenius and Hodges (1959), Singh (1971), Lavallée and Hidiroglou (1988), and Sweet and Sigman (1995). Once stratum boundaries have been defined, and a maximum sample size or total cost determined, it is straightforward to determine the number of sample units to allocate to each stratum if the allocation is done on a single variable (Cochran, 1977).

Notation

When a sample is obtained, a typical problem that arises is that of estimating various characteristics of the population. This is usually complicated by the fact that the different characteristics may have different variances, which means that the sample sizes for each characteristic may vary. For a formal expression of the problem of optimum allocation in stratified sampling, consider the following notation (Cochran, 1977; Sukhatme *et al*., 1984; and Thompson, 1997).

The sub index $h=1,2,......,H$ denotes the stratum, and $i=1,2,........,N_h$ the unit within stratum $h.$ Moreover:

N_h Total number of units within stratum h

n_h Number of units from the sample in stratum h

y_{hi} Value obtained for the $i-th$ unit in stratum h

$n=(n_1, n_2, ..., n_H)'$ Vector of the number of units in the sample

$W_h = \frac{N_h}{N}$ Relative size of stratum h

$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h}$ Population mean in stratum h

$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$ Sample mean in stratum h

$S_h^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1}$ Population variance in stratum h

c_h Cost per sampling unit in stratum h

$\bar{y}_{st} = \sum_{h=1}^{H} W_h \bar{y}_h$ Estimator of the population mean in the stratified sampling.

$V(\bar{y}_{st})$ Variance of $\bar{y}_{st}$.

Optimum Stratification

Optimum stratification is used when the standard deviation of the population strata are known to differ substantially. By taking larger samples in the strata with higher variability and smaller samples in the strata with lower variability, the sample size may be considered as combating variability. To make use of optimum stratification methods, one must have knowledge or an estimate of the standard deviation in each of the individual strata. This knowledge or estimate of the standard deviation is necessary before the sample is selected as it serves as a basis for allocation of the sample size among the individual strata.

By applying the optimal expression to stratified random sampling, the variance function is

$$\begin{aligned} \mathrm{var}(\bar{y}_{st}) &= (1-f_h)\sum W_h^2 \frac{S_h^2}{n_h} \\ &= \left(1-\frac{n_h}{N_h}\right)\sum W_h^2 \frac{S_h^2}{n_h} \\ &= \frac{\sum W_h^2 S_h^2}{n_h} - \frac{\sum W_h^2 S_h^2}{N_h} \end{aligned} \quad (2.3.1)$$

and the cost function as

$$C = \sum C_h n_h + K_v$$

While the direct optimal allocation is

$$n_h = \frac{K' W_h S_h}{\sqrt{C_h}} \quad (2.3.2)$$

and if

$$W_h = \frac{N_h}{N}$$

Then

$$n_h = \frac{K'}{N} \cdot \frac{N_h S_h}{\sqrt{C_h}} \qquad (2.3.3)$$

With fixed $C_0 = \sum C_h n_h = Cost - K_v$

We obtain

$$K' = \frac{\sum C_h n_h}{\sum W_h S_h \sqrt{C_h}} \qquad (2.3.4)$$

and

$$V_{\min}^2 = \frac{\left(\sum W_h S_h \sqrt{C_h}\right)^2}{\sum C_h n_h} \qquad (2.3.5)$$

With fixed

$$V_0^2 = \frac{\sum W_h^2 S_h^2}{n_h}$$

$$= Var(\bar{y}_{st}) + \frac{\sum W_h^2 S_h^2}{N_h} \qquad (2.3.6)$$

We obtain

$$K' = \frac{\sum W_h S_h \sqrt{C_h}}{V_0^2} \qquad (2.3.7)$$

and

$$C_{\min} = \frac{\left(\sum W_h S_h \sqrt{C_h}\right)^2}{V_0^2} \qquad (2.3.8)$$

When h takes on only two values, 1 and 2, the expressions for optimum allocation is

$$Optimum\left(\frac{n_1}{n_2}\right) = \frac{KV_1 / \sqrt{C_1}}{KV_2 / \sqrt{C_2}}$$

$$= \frac{V_1}{V_2}\left(\frac{C_2}{C_1}\right)^{1/2}$$

$$= \frac{V_1}{V_2}\sqrt{K} \qquad (2.3.9)$$

Here $K = C_1 / C_2$ in the cost function,

$$C = C_1 n_1 + C_2 n_2 = C_1 n_1\left(1 + \frac{K n_2}{n_1}\right) \quad (2.3.10)$$

Then with fixed $C_0 = C_1 n_1 + C_2 n_2$, we obtain

$$V_{\min}^2 = \frac{\left(V_1\sqrt{C_1} + V_2\sqrt{C_2}\right)^2}{C_1 n_1 + C_2 n_2} = \frac{C_1\left(V_1 + V_2\sqrt{K}\right)^2}{C_0} = \frac{C_1\left(V_1 + V_2\sqrt{K}\right)^2}{n_1\left(1 + K\, n_2/n_1\right)} \quad (2.3.11)$$

On the other hand, for fixed V_0^2,

$$C_{\min} = \frac{C_1\left(V_1 + V_2\sqrt{K}\right)^2}{V_0^2} \quad (2.3.12)$$

Optimum Allocation via Multi-objective Optimization

The estimator of the population mean in multivariate stratified sampling for the j-th characteristic is defined as

$$\bar{y}_{st}^{j} = \sum_{h=1}^{H} W_h \bar{y}_h^{j} \quad (2.4.1)$$

Where $\bar{y}_h^{j} = \frac{1}{n_h}\sum_{h=1}^{n_h} y_{hi}^{j}$ is the sample mean in stratum h of the j-th characteristic, and y_{hi}^{j} is the value obtained for the i-th unit in stratum h of the j-th characteristic. The Var $(\bar{y}_{st}^{j})$ is defined using the population variances $S_h^2, h = 1,2,......., H,$ which are usually unknown, and therefore these are substituted by the sample variances $s_h^2, h = 1,2,....., H,$ defined as

$$s_h^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(y_{hi} - \bar{y}_h)^2 \quad (2.4.2)$$

And thus $\hat{\text{V}}\text{ar}(\bar{y}_{st}^{j})$ is substituted by the estimated variance $\hat{\text{V}}\text{ar}(\bar{y}_{st}^{j})$, which is given by

$$\hat{\text{V}}\text{ar}(\bar{y}_{st}^{j}) = \sum_{h=1}^{H}\frac{W_h^2 s_{hj}^2}{n_h} - \sum_{h=1}^{H}\frac{W_h s_{hj}^2}{N}$$

Alternative Approach

The linear programming problem assumed that the covariances between the different characteristics are zero, which under the assumption of normality is equivalent to assuming the characteristics to be stochastically independent. This idea leads to proposal of optimum allocation under multivariate stratified sampling as

$$\min_n \theta$$
$$\text{subject to}$$
$$\sum_{h=1}^{H} c_h n_h + c_0 = C \qquad (2.5.1)$$
$$2 \le n_h \le N_h, \quad h = 1,2,........., H$$
$$n_h \in N.$$

Where $\theta = Cov(\bar{y}_{st})$ is the matrix of variances-covariances of the vector $\bar{y}_{st} = (\bar{y}_{st}^1,........., \bar{y}_{st}^G)'$. The difficulty of problem (2.5.1) lies in defining the meaning of the minimum of a matrix. To arrive at the minimum of matrix, Diaz-Garcia and Cortez (2006) used the value function method. The programme (2.5.1) using the value function technique is given as

$$\min_n v(\theta)$$
$$\text{subject to}$$
$$\sum_{h=1}^{H} c_h n_h + c_0 = C \qquad (2.5.2)$$
$$2 \le n_h \le N_h, \quad h = 1,2,........., H$$
$$n_h \in N.$$

If in particular $v(\theta) = \mathrm{tr}(\theta),$ and all the characteristics are given the same weighting such that $\sum_{j=1}^{G} \lambda_j = 1,\ \lambda_j \ge 0\ \forall\ j = 1,2,....., G,$ then the value function solution is obtained. If one simply define $v(\theta) = \det(\theta) = |\theta|,$ the following programme under approach (2.5.1) is obtained as

$$\min_n |\theta|$$
$$\text{subject to}$$
$$\sum_{h=1}^{H} c_h n_h + c_0 = C \qquad (2.5.3)$$
$$2 \le n_h \le N_h, \quad h = 1,2,........., H$$
$$n_h \in N.$$

The determinant and trajectory are not the only value functions that can be used. Alternative definitions of the value function that have been used in other statistical contexts include

(1)The sum of all the elements of the matrix $\theta = (\theta_{kl}); v(\theta) = \sum_{k,l=1}^{G} \theta_{kl}.$

(2) $v(\theta) = \lambda_1(\theta),$ where λ_1 is the maximum eigenvalues of the matrix of covariances θ.

(3) $v(\theta) = \lambda_G(\theta),$ where λ_G is the minimum eigenvalues of the matrix of covariances $\theta.$

In order to improve the precision of estimates in a sample survey and to correct for a bias caused by problems in the sample survey process, Bethlehem and Keller (1983) suggested for some kind of weighting. Generally, weights are computed such that the weighted distribution of the observed values of the auxiliary variables equals the distribution of the values in the population. Bethlehem and Keller (1983) presented a general framework for weighting based on an estimator constructed from a linear model which relates the target variables of the survey to the auxiliary variables. Their approach presents an alternative in situations where classical weighting by post-stratification produces disturbances.

Optimal Design for a Multivariate Stratified Sampling Adopted in the Study

The idea of optimal allocation under a multivariate stratified sampling in the study is based on an alternative approach as in Diaz-Garcia and Ramos-Quiroga (2011).

The linear programming problem is assumed to be

$$\min_{n} \theta$$

Subject to

$$\sum_{h=1}^{H} C_h n_h + C_0 = C$$

$$2 \le n_h \le N_h \qquad h = 1 \qquad (2.6.1)$$

Where $\theta = Cov\,(\bar{y}_{st})$. This is the matrix of variance covariances of the vector $\bar{y}_{st} = (\bar{y}_{st}^{1},\ \bar{y}_{st}^{G})'$.

The subindex $h = 1,2,.....,H$ denotes the stratum $i = 1,2,......,N_h$ or n_h within stratum h and $j = 1,2,.....,G$. denotes the characteristic (variable).

The covariance matrix of ỹ st denoted as cov (ỹst) is defined in matrix

$$Cov(\bar{y}_{st}) = \begin{bmatrix} \mathrm{var}(\bar{y}_{st}^{1}) & \mathrm{cov}(\bar{y}_{st}^{1,}, \bar{y}_{st}^{2}) & ..... & \mathrm{cov}(\bar{y}_{st}^{1}, \bar{y}_{st}^{G}) \\ \mathrm{cov}(\bar{y}_{st}^{2}, \bar{y}_{st}^{1}) & \mathrm{var}(\bar{y}_{st}^{2}) & .... & \mathrm{cov}(\bar{y}_{st}^{2}, \bar{y}_{st}^{G}) \\ : & : & : & : \\ \mathrm{cov}(\bar{y}_{st}^{G}, \bar{y}_{st}^{1}) & \mathrm{cov}(\bar{y}_{st}^{G}, \bar{y}_{st}^{1}) & .... & \mathrm{var}(\bar{y}_{st}^{G}) \end{bmatrix} \qquad (2.6.2)$$

and the estimated covariance of $\bar{y}_{st}^{i}$ and $\bar{y}_{st}^{j}$ as $Cov(\bar{y}_{st}^{i}, \bar{y}_{st}^{j})$. Thus $Cov(\bar{y}_{st}^{i}, \bar{y}_{st}^{j})$ is

$$Cov(\bar{y}_{st}^{i}, \bar{y}_{st}^{j}) = \sum_{h=1}^{H} \frac{W_h^2 S_h ij}{n_h} - \sum_{h=1}^{H} \frac{W_h S_h ij}{N} \qquad (2.6.3)$$

and

$$Cov(\bar{y}_{st}^{i}, \bar{y}_{st}^{j}) = \sum_{h=1}^{H} \frac{W_h^2 S_h i}{n_h} - \sum_{h=1}^{H} \frac{W_h S_h i}{N} \qquad (2.6.4)$$

and c_h is the cost per G -dimensional sampling unit in stratum h and its vector $c = (c_1,.............,c_G)'$.

The Case Study

To demonstrate the feasibility of the survey design adopted in this study, an empirical data from a household survey conducted in Ogun State was used. The frame consists of 880,970 households in the twenty local government areas (LGAs) of Ogun State. Each of the 20 local government areas that made up Ogun State was seen as a stratified cluster. Two local government areas with a sample size of 200 households each were randomly selected using simple random sampling technique. These local government areas are Abeokuta South and Ijebu North. The household data comprised of different characteristics out of which four socioeconomic characteristics of the 400 head of households were investigated. The four characteristics of interest were occupation, income, household size (dependant size) and educational level.

Post- stratified estimators are commonly used in sample surveys to improve the efficiency of estimators and to ensure calibration to known post-strata. Post-stratification in this study ensured that some variables that are suitable for stratification was achieved after selection of the sample.

The stratification technique in this study divided up the population into sub-population or strata. The study transformed all variables in the stratified sampling plan by using dummy variables ranging from 1 to 3.

Results

The estimates used in the computation were calculated using statistical analysis software Splus. The stratified data for the four characteristics is as shown in table 4.1.

Table 4.1: Stratified Data on Occupation, Income, Dependant Size and Educational Level of Heads of Households in Abeokuta South and Ijebu North

Item No.	Name	Stratum No.	Stratum Name	Size of Stratum: Abeokuta South	Size of Stratum: Ijebu-North
1	Occupation	1	Unemployed	10	2
		2	Paid employment	47	54
		3	Self employment	143	144
2	Income (in N'000)	1	0-10	42	28
		2	10-20	73	91
		3	20+	85	81
3	Dependant Size	1	Small (1-3)	138	140
		2	Moderate (4-7)	58	55
		3	Large (7+)	4	5
4	Educational Level	1	Primary	53	44
		2	Secondary	74	85
		3	Tertiary	73	71

With stratified sampling, the sample estimate of population mean (proportion) and the standard error of estimate associated with this statistic is determined by appropriately weighting the individual strata results. Optimum stratification is used when the standard deviation of the population strata is known to differ substantially.

By taking larger samples in the strata with higher variability and smaller samples in the strata with lower variability, the sample size is considered as combating variability. When applied to the household data from Abeokuta South and Ijebu North, the resultant sample allocation is as in Tables 4.2 and 4.3.

Table 4.2: Optimum Allocation of Sample Size to Occupation, Income, Dependant Size and Educational Level in Abeokuta South

Item No.	**Name**	**Stratum No.**	**Stratum Name**	**No. In Sample** N_h	**Estimated Standard Deviation of Stratum** $\hat{S}_h$	$N_h\hat{S}_h$	$\frac{N_h\hat{S}_h}{\sum_{h=1}^{k} N_h\hat{S}_h}$
1	Occupation	1	Unemployed	10	0.025	0.250	0.004
		2	Paid Employment	47	0.118	5.546	0.097
		3	Self Employment	143	0.358	51.190	0.899
2	Income (in N'000)	1	0-10	42	0.161	6.762	0.123
		2	10-20	73	0.281	20.513	0.372
		3	20+	85	0.327	27.795	0.505
3	Dependant Size	1	Small (1-3)	138	0.353	48.714	0.850
		2	Moderate (4-7)	58	0.148	8.584	0.150
		3	Large (7+)	4	0.010	0.040	0.001
4	Educational Level	1	Primary	53	0.210	11.130	0.206
		2	Secondary	74	0.293	21.682	0.402
		3	Tertiary	73	0.289	21.097	0.391

Table 4.3: Optimum Allocation of Sample Size to Occupation, Income, Dependant Size and Educational Level in Ijebu North

Item No.	**Name**	**Stratum No.**	**Stratum Name**	**No. In Sample** N_h	**Estimated Standard Deviation of Stratum** $\hat{S}_h$	$N_h\hat{S}_h$	$\frac{N_h\hat{S}_h}{\sum_{h=1}^{k} N_h\hat{S}_h}$
1	Occupation	1	Unemployed	2	0.005	0.010	0.002
		2	Paid Employment	54	0.125	6.750	0.123
		3	Self Employment	144	0.333	47.952	0.876
2	Income (in N'000)	1	0-10	28	0.097	2.716	0.050
		2	10-20	91	0.314	28.574	0.530
		3	20+	81	0.279	22.599	0.419
3	Dependant Size	1	Small (1-3)	140	0.364	50.960	0.865
		2	Moderate (4-7)	55	0.143	7.865	0.134
		3	Large (7+)	5	0.013	0.065	0.001
4	Educational Level	1	Primary	44	0.171	7.524	0.136
		2	Secondary	85	0.331	28.135	0.509
		3	Tertiary	71	0.276	19.596	0.355

The design effect (d^2) equired that these stratified random samples are compared with the simple random sampling variance for a sample of the same size as estimated. Hence the design effect is estimated and is as shown in Tables 4.4 and 4.5.

Table 4.4: Design effect of the Mean for Abeokuta South Samples

Design Effects d^2	**Occupation**	**Income**	**Dependant Size**	**Educational Level**
Post-stratified	0.450	0.341	0.581	0.690
Optimum Allocation	0.338	0.211	0.371	0.322

Table 4.5: Design effect of the Mean for Ijebu North Samples

Design Effects d^2	**Occupation**	**Income**	**Dependant Size**	**Educational Level**
Post-stratified	0.595	0.325	0.875	0.789
Optimum Allocation	0.165	0.185	0.688	0.632

The post-stratification estimator $\bar{y}_{st}$ does not have the same variance as the stratified sample mean because stratification was not designed in the sampling scheme. Any stratified random sampling gives a smaller variance than a simple random sample. To buttress further on this, a comparison is made between simple random sampling and stratified random sampling with proportional and optimum allocation and the estimates are in Tables 4.6 and 4.7.

Table 4.6: Summary Estimates of Abeokuta South Sample Statistics

	Occupation	**Income**	**Dependant Size**	**Educational Level**
Mean $(\bar{y}_{st})$	1.833	2.067	1.3000	1.700
$V_{srs}(\bar{y})$	0.008	0.0214	0.0062	0.018
Var(post)	0.0036	0.0073	0.0036	0.0069
$V_{mod}(\bar{y}_{st})$	0.0027	0.0045	0.0023	0.0058

Table 4.7: Summary Estimates of Ijebu North Sample Statistics

	Occupation	**Income**	**Dependant Size**	**Educational Level**
Mean $(\bar{y}_{st})$	1.833	2.033	1.333	2.033
$V_{srs}(\bar{y})$	0.0079	0.0206	0.0016	0.0019
Var(post)	0.0047	0.0067	0.0014	0.0015
$V_{mod}(\bar{y}_{st})$	0.0013	0.0038	0.0011	0.0012

Conclusion

In this study, Post Stratified and Stratified Sampling approach have been presented as an appropriate research design and data collection instruments. Stratification by making use of existing knowledge concerning the population provided an effective means of reducing the measured variability of estimates. Stratification produced a gain in precision in the estimation of characteristics of the household survey of Abeokuta South and Ijebu North population. The sampling variance was indeed one of the key indicators of quality in sample surveys and estimation. The results of the estimates in the study revealed that proportionate allocation can lead to smaller sampling variances than for simple random samples of the same size.

Since the standard deviations of the sample strata computed differed substantially, optimum stratification is used as a method of allocating larger size samples to those strata with larger standard deviations.

References

Arthanari, T.S and Dodge, Y. 1981. Mathematical programming on statistics. A Wiley-Interscience, Publication, John Wiley & Sons Inc.

Bethlehem, J. C. and Keller, W. J. 1983 Linear models for weighting sample survey Data. International Statistical institute. **2**: 615 – 619.

Cochran, W. G. 1977. Sampling Techniques (3rd Edition), New York, Wiley.

Dalenius , T and Hodges, J.L. 1959 Minimum Variance Stratification. *Journal of the American statistical Association*, vol. 54, PP 88-101.

Diaz-Garcia, J.A and Cortez, L.U. 2008. Multi-objective optimisation for optimum allocation in multivariate stratified sampling. *Survey Methodology*, Vol. 34, No 2, 215-222.

Diaz-Garcia, J.A and Cortez, L.U. 2006. Optimum allocation in multivariate stratified sampling: Multi-objective programming.

Diaz-Garcia, J.A and Ramos-Quiroga, R. 2011. Multivariate Stratified Sampling by Stochastic Multi Objective Optimization. Xiv: 1106.0773VI. *Statistical Methodology. XIV*, 1116-1123.

Fuller, S. J. 1993. Selection of a stratified Random Sample. Informedix Marketing Research, Inc.

Harvey, A.C. and Shephard, N. 1993. Structural time series model. In Maddala, G.S. (Ed) *Handbook of Statistics,* Vol. 11. Barking: Elsevier Science Publishers.

Holt, D. and Smith, T. M. F. 1979. Post stratification. *Journal of Royal Statistical Society,* 142 (1) 33 – 46.

Kish, L. 1965. Survey sampling. New York, Wiley.

Lavallee, P., and Hidiroglou, M.A 1988. On the stratification of skewed population. *Survey Methodology*, Vol. 14, 33-43

Pirzada, S. and Maqbool, S. 2003. Optimal Allocation in Multivariate sampling Through Chebyshev's Approximation. *Bulletin of the Malaysian Mathematical Science Society*, 2, (26), 221 – 230.

Scott, A.J. and Smith, T.M.F. 1974. Analysis of repeated surveys using time series methods. *Journal of American Statistical Association.* Vol. 69, 674-678.

Singh, R. 1971. Approximately optimum stratification on the Auxiliary Variable, *Journal of the American Statistical Associations*, Vol.. 66, PP 829-833.

Sukhatme, P.V, Sukhatme, B.V, Sukhatme, S., and Asok, C. 1984. Sampling Theory of Survey with Applications. 3rd Edition. Ames, Iowa: Iowa State University Press.

Sweet, E.M and Sigman, R.S 1995. Evaluation of Model-Assisted Procedures for stratifying skewed population using Auxiliary Data. In proceedings of the section on survey research methods. American Statistical Association, 491-496.

Thompson, M.E.1997. Theory of Sample Survey. Chapman and Hall Publishers.