Estimation of Total Maize Yields in Nigeria

Ikughur A. J.[1,*], Imande M. T.[2], Ohioma E. O. B.[3]

[1]Department of Mathematics/Statistics/ Computer Science, University of Agriculture, Makurdi, Nigeria
[2]Department of Mathematics/Computer Science, Benue State University Makurdi, Nigeria
[3]Department of Real Sector and Household Statistics, National Bureau of Statistics, Abuja, Nigeria

Abstract The problem of food security is of great concern to all stakeholders for national development. Food shortage in an economy inflicts undesirable costs on citizens as food is a basic human need. Of statistical concern is the estimation of crop yield in a single period called survey period. When data is available over different periods of time, it becomes plausible that the trend of the estimated yields be also investigated. Thus, this study utilized design and model based ratio and regression estimators to provide estimates of total national yields of maize and then fit linear trend to determine the growth in crop yield over time in Nigeria. It further investigated the performance of the estimator using samples of sizes 20 and 30 drawn using simple random sampling without replacement strategy from the 37 survey plots (36 states and FCT, Abuja) and the variables namely, maize yield (Y) and land area (X) were observed for estimation. Similarly, the asymptotic properties of the estimator are also investigated. Results revealed that there was increase in the year by year estimate of total yield of maize in each of the sample size cases except in 2006 where a sudden fall was experienced. It further showed that regression estimator performed better than the ratio estimator when the Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Coefficient of Variation (CV) are utilized for all sample sizes considered for both design and model based estimators.

Keywords Food security, Estimation, Auxiliary variable, Relative efficiency

1. Introduction

In most developing countries, agriculture is both the main traditional pursuit and the key to sustained growth of the modern economy. Economic growth has gone hand in hand with agricultural progress. On the other hand, stagnation in agriculture is the principal explanation for poor economic performance, while rising agricultural productivity has been the most important concomitant of successful industrialization (Abayomi, 1997). In Nigeria, the trend in the share of agriculture in the gross domestic product (GDP) shows a substantial variation and long-term decline from 60% in the early 1960s through 48.8% in the 1970s and 22.2% in the 1980s (Abayomi, 1997). A slight increase in the GDP was witnessed in 2010, 23.89% (NBS, 2011) and 23.33% in 2013 (NBS, 2014).

It has been predicted that there would be a catastrophic widespread food famine, the type the world has never experienced before the year 2020 (Food Security and National Development, 2014) even as Wood (2011) observed that, the earth's population would grow so great as to overcome the earth's ability to   provide enough food and other infrastructures necessary for human survival.

Food Security is achieved when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life (FAO, 1996). Thus, the main goal of food security is for individuals to be able to obtain adequate food needed at all times, and to be able to utilize the food to meet the body's needs. It is therefore of national interest to devise strategies to guard against food insecurity and hence the need to provide statistical information for good national planning and execution of sustainable agricultural development programmes and policies that will ensure stable food supply.

One useful means for planning agricultural activities for the future is the use of agricultural statistics that are collected either by censuses, which require enumeration of the total population of interest, or by samples requiring enumeration of only a small part of the population and estimating population characteristics. The data so collected are put in the form of models to estimate or predict the characteristics of interest. There are several models or estimators of crop yield that are found in the works by Jain *et al.,* (1984), Martis *et al.,* (1988), Brus and Gruijter (1997), Bornn and Zidek (2012) among others.

In this study, we employ the design based ratio estimator to provide estimates of total maize yields in Nigeria under the assumption that each administrative state in Nigeria forms a survey plot under which the yields reported are assumed to be realization of a crop-cutting experiment. The

aim of the study, therefore, is to provide estimates of total crop yields (for maize) in Nigeria and the trend of maize yield over time. Specifically, we utilize the regression and design based ratio estimator to provide estimates of total crop yields for samples of sizes 20 and 30, and to study the efficiency of the estimators using design based efficiency criteria.

2. Literature Review

Agricultural productivity refers to the output produced by a given level of input(s) in a given economy (Fulginiti and Perrin, 1998). More formally, it can be defined as "the ratio of the value of total farm outputs to the value of total inputs used in farm production" (Olayide and Heady, 1982). Crop Yield on the other hand, is the measurement of the amount of a crop that was harvested per unit of land area (also referred to as "actual seed generation from the plant" or "agricultural output").

Mathematically, we define crop yield as crop yield = (amount of harvested product) / (crop area).This ratio expresses in kilograms (kg) or metric tonnes (t) of product per hectare (ha). In essence, the estimation of crop yield thus involves both estimation of the crop area and estimation of the quantity of product obtained from that area which Michael (2007) portrayed as the product of two components namely: area (to be) harvested and (expected) yield per unit area.

In survey sampling studies, there are several estimators that utilize auxiliary information. These include ratio estimators, regression estimators, probability proportional to size with and without replacement strategies (Chaudhuri, 2010 and Raj, 1965) among others and they provide foundational estimators of population characteristics when auxiliary information are available.

Anneke (2011), revealed that pioneers in sampling and survey design in the late 1940s developed a method (crop-cut methods) for estimating crop yields based on sampling of small subplots within cultivated fields in India. One of such early contribution include works by Cochran (1940). In the 1950s, crop-cut methods were adopted as the standard method recommended by the Food and Agriculture Organization of the United Nations (FAO) to measure crop production (FAO, 1982; Murphy, Casley, and Curry, 1991).

In India, the sampling design of the crop surveys is a stratified two-stage random sampling with villages in the circles as the primary sampling units, fields growing the crop in the selected villages as the second stage units and standard plots usually of size 5 m x 5 m in the selected fields as the ultimate sampling units while in Nigeria, agricultural data is collected by National Bureau of Statistics (NBS) through the two data collection infrastructure; National Integrated Survey of Household (NISH) and National Integrated Survey of Establishment (NISE). A two-phase sampling scheme has been adopted by NBS for traditional/peasant sector of the sample survey. Cochran (1940) made the first attempt to investigate the problem of estimation of population mean when auxiliary variables are present and he proposed the ratio estimator of population mean making use of information related to the supplementary characteristics having positive correlation with the variable under consideration.

Bellow (2007) revealed that the National Agricultural Statistics Service (NASS) in United States Department of Agriculture used ratio estimation method to derive county level yields. The simple ratio estimator was computed as the sum of Quarterly Agricultural Surveys (QAS) reported crop production divided by the corresponding sum of reported harvested acreage. It was deduced that the estimator produced unreliable yields due to fluctuations in harvested area from year to year.

Arunachalam and Balakrishnan (2012) investigated the trends in area, production and productivity of wheat crop grown during 1950 - 1951 to 2009 - 2010 in India using different non-linear models for data. They found that the Sinusoidal model was suitable to fit the trend in production of wheat. Similarly, Makwoski and Micheal (2013) used dynamic linear model for predicting crop yield trends in foresight studies on food security. They observed that the world's population is fast growing to pass 9 billion in 2050 and this raises increasing concern about the capacity of agriculture to feed the world. They presented a foresight study on food security by estimating crop yields using statistical methods such as Time series, linear and non-linear regression, and moving average.

Although many models have been used to estimate crop yields at local level, none has estimated crop yields at national level in Nigeria while considering land use as auxiliary variable using the design and model based ratio estimation techniques along with the study of trend of the estimates. Therefore, this study focused on the use of design and model based ratio estimators to estimate maize yields in Nigeria. Trend analysis was also carried out to determine the growth of crop yields.

3. Methods

In this section, method of data collection and the estimator of crop (maize yield) using the design and model based estimation perspectives are presented. Data reported by National Bureau of Statistics (NBS) in her annual Agricultural Surveys from 1994 to 2009 is considered and it consists of the study variable, Y (maize yields) and auxiliary variable, X (land area) for the 36 states of Nigeria and the Federal Capital Territory (FCT). Using this as sampling frame with N=37, we select a samples of size n = 20 and n = 30 states using simple random sampling without replacement (SRSWOR) strategy. The choice of n=20 and n=30 is to permit the comparison of the ratio estimator of crop yields. A uni-cluster design with varying elements taken at one stage only is assumed in this study.

The Ratio Estimator

Consider a population of size N from which a sample of size $n < N$ is drawn without replacement. Suppose that a SRSWR scheme is implemented, $Y = \{y_1, \ldots, y_n\}$ and $X = \{x_1, \ldots, x_n\}$ are drawn from the sample. Let $\hat{Y}$ and $\hat{X}$ be unbiased estimators of the population totals Y and X of the study and auxiliary variable respectively. Then, the design based ratio estimator of the population total is defined as

$$\hat{Y}_R = \frac{\hat{Y}}{\hat{X}} X \qquad (1)$$

Under probability sampling design, the ratio estimator is a biased estimator of population total as $E_P[\hat{Y}_R] \neq Y$.

Bias and Mean Squared Error of the Ratio Estimator

The approximate bias of the ratio estimator denoted as B $(\hat{Y}_R)$ is given as

$$B\ (\hat{Y}_R) = Y\left\{\left[\frac{V(\hat{X})}{X^2}\right] - \left[\frac{cov(\hat{X},\hat{Y})}{XY}\right]\right\} \qquad (2)$$

while the Mean Square Error of the ratio estimator, denoted by MSE $(\hat{Y}_R)$, is given as

$$MSE\ (\hat{Y}_R) = Y^2\left\{\left[\frac{V(\hat{Y})}{Y^2}\right] + \frac{V(\hat{X})}{X^2} - 2\left[\frac{cov(\hat{X},\hat{Y})}{XY}\right]\right\} \qquad (3)$$

Since the samples are selected using the SRSWOR procedure, the following results for estimator, it's bias and MSE that conform with the SRSWOR scheme are given below:

$$\hat{Y}_R = \frac{\sum_{i\in s} Y_i}{\sum_{i\in s} X_i} X \qquad (4)$$

$$B\ (\hat{Y}_R) = \frac{N^2(N-n)}{Nn} Y\left\{\frac{S_x^2}{X^2} - \frac{S_{xy}}{XY}\right\} \qquad (5)$$

$$MSE(\hat{Y}_R) = \frac{N^2(N-n)}{Nn} Y^2\left\{\frac{S_Y^2}{Y^2} + \frac{S_x^2}{X^2} - 2\frac{S_{xy}}{XY}\right\} \qquad (6)$$

The corresponding sample means, variances for Y and X and the covariance are respectively given below;

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (7)$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (8)$$

$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}[x_i - \bar{x}]^2 \qquad (9)$$

$$S_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}[x_i - \bar{y}]^2; \qquad (10)$$

and

$$S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}[x_i - \bar{x}][y_i - \bar{y}] \qquad (11)$$

When samples have been drawn, the observations are used to estimate the population mean square error and the associated inference based on the sample so selected. The estimated MSE of the population total for ratio estimator under SRSWOR is given by;

$$\widehat{MSE(\hat{Y}_R)} = \frac{N^2(N-n)}{Nn} \frac{1}{n-1}\sum_{i\in s}[Y_i - \hat{R}X_i]^2 \qquad (12)$$

where

$$\hat{R} = \frac{\sum_{i\in s} Y_i}{\sum_{i\in s} X_i} \qquad (13)$$

Super-population models

Consider the super population model

$$y_i = \alpha + \beta x_i + e_i \qquad (14)$$

With the following assumptions namely, $\xi(e_i|x_i) = 0$, $\xi Var(e_i|x_i) = \sigma^2 x_i$ and $\xi(e_i e_j|x_i x_j) = 0$, defined as model based expectation, variance and covariance respectively. Under super population model, noting that $\alpha = 0$ we have

$$E(\hat{Y}_R) = \beta E(x) + X\frac{\xi e_i}{x} = \beta X \qquad (15)$$

which confirms that ratio estimator is a special case of regression estimator.

Model Based Estimator of MSE and its stability

The model based estimator of population MSE for ratio estimator is

$$\xi[MSE(\hat{Y}_R)] = K_1 K_2 [\nabla_1\sum_{i\in s} x_i^2 - \nabla_2\sum_{i\in s} x_i] \qquad (16)$$

where $K_1 = \frac{N^2(N-n)}{Nn}$ and $K_2 = \frac{1}{n-1}$. In (16) above, there are two components namely

$$\nabla_1\sum_{i\in s} x_i^2 = (\beta^2 + \hat{R}^2 - 2\hat{R}\beta)\sum_{i\in s} x_i^2 \qquad (17)$$

and

$$\nabla_2\sum_{i\in s} x_i = \sigma^2\sum_{i\in s} x_i \qquad (18)$$

Thus (17) and (18) when summed together satisfy the non-negativity of variance condition and the stability of the variance of the estimator could be investigated using the criterion namely

$$\beta^2 + \hat{R}^2 - 2\hat{R}\beta \geq 0 \qquad (20)$$

for estimated values of R and β.

Investigating Asymptotic Property of the Estimator

The Chebychev's Inequality is used here to investigate the asymptotic property of the ratio estimator. In particular, if $g(X) = (X - \mu_x)^2$ and we set $K = \varepsilon$ so that $E(g(X)) = E(X - \mu_x)^2 = \sigma_x^2$. Thus, by Chebychev, $P_r\{|X - \mu_x| \geq \varepsilon\} \leq \frac{\sigma_x^2}{\varepsilon^2}$ or $P_r\{|X - \mu_x| < \varepsilon\} > 1 - \frac{\sigma_x^2}{\varepsilon^2}$ and further substituting $\varepsilon = n\sigma_x$, we get $P_r\{|X - \mu_x| \geq n\sigma_x\} \leq \frac{1}{n^2}$.

For the estimator of maize yield $\hat{Y}_R$, Let the bias be $B_p(\hat{Y}_R) = E_p(\hat{Y} - Y)$ and the MSE given as $M_p(\hat{Y}_R) = E_p(\hat{Y}_R - Y)^2$, then $M_p(\hat{Y}_R) \geq K^2 \text{Prob}[|\hat{Y}_R - Y| \geq K]$ (see Chaudhuri, 2010)

$$\Rightarrow K^2 \text{Prob}[(\hat{Y}_R - Y) \geq K] \leq \frac{M_p(\hat{Y}_R)}{K^2} \qquad (21)$$

$$\text{Or } P_r[|\hat{Y}_R - K| \leq Y \leq \hat{Y}_R + K] \geq 1 - \frac{M_p(\hat{Y}_R)}{K^2} \qquad (22)$$

Neyman (1934) demonstrated this version of Chebychev's inequality in the context of survey sampling by taking K = $\alpha\sigma_p(\hat{Y}_R)$ where α is a positive constant. It follows that

$$P_r\left[[\hat{Y}_R - \alpha\sigma_p(\hat{Y}_R), \hat{Y}_R + \alpha\sigma_p(\hat{Y}_R)]\right]$$

$$\geq \left(1 - \frac{1}{\alpha^2}\right) - \left(\frac{1}{\alpha^2}\left(\frac{|B_p(\hat{Y}_R)|}{\sigma_p(\hat{Y}_R)}\right)\right) \qquad (23)$$

where $MSE_p(\hat{Y}_R)$ is used as proxy for $\sigma_p(\hat{Y}_R)$.

Regression Estimator of Population Total

The linear regression estimator like others is designed to increase precision by the use of an auxiliary variate x_i that is correlated with y_i. We suppose that y_i on x_i are obtained for every unit in the sample and that the population mean $\bar{X}$ of the x_i is known. The linear regression estimate of $\bar{Y}$, the population mean of the y_i, is

$$\bar{y}_{lr} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x}) \qquad (24)$$

Where *lr* denotes linear regression and $\hat{\beta}$ is estimate of the change in y when x is increased by unity where $\hat{\beta} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$. Therefore, the corresponding estimate of the population total Y is given as

$$\hat{Y}_{lr} = N\bar{y}_{lr} \qquad (25)$$

The sample estimate of $V(\bar{y}_{lr})$ is given as

$$v(\bar{y}_{lr}) = \frac{1-f}{n(n-2)}\sum_{i}^{n}[(y_i - \bar{y}) + \beta(x_i - \bar{x})]^2 \qquad (26)$$

The corresponding estimates of population total is

$$\hat{Y} = N\bar{y}_{lr} \qquad (27)$$

Where the estimated variance can also be written as

$$Var(\bar{y}_{lr}) = \frac{N(N-n)}{n}S_y^2(1 - \rho^2) \qquad (28)$$

Comparison of Regression and Ratio Estimators

Two comparable variances for the estimated population mean $\bar{Y}$ for regression and ratio estimators are $V(\bar{y}_{lr}) = \frac{N-n}{n}S_y^2(1 - \rho^2)$ and $V(\bar{y}_R) = \frac{N-n}{Nn}(S_y^2 + R^2 S_x^2 - 2R\rho S_y S_x)$ respectively.

Regression estimator is more efficient than the Ratio estimator if and only if $V(\bar{y}_{lr}) < V(\bar{y}_R)$

Which is equivalent to $\frac{N-n}{n}S_y^2(1 - \rho^2) - \frac{N-n}{Nn}(S_y^2 + R^2 S_x^2 - 2R\rho S_y S_x) < 0$. Assuming that $\frac{N-n}{n}$ and also $\frac{N-n}{Nn}$ are negligible, we realize $S_y^2\rho^2 + R^2 S_x^2 - 2R\rho S_y S_x > 0$ and by factorization noting that, $\beta = \rho \frac{S_y}{S_x} \Rightarrow \beta S_x = \rho S_y$, the condition becomes $(\rho S_y - R S_x)^2 > 0$ or $(\beta - R)^2 > 0$. If $\frac{N-n}{Nn}$ and hence $\frac{N-n}{n}$ are not negligible, then $S_y^2(1 - \rho^2) - \frac{1}{N}S_y^2 - \frac{R^2 S_x^2}{N} + \frac{2R\rho S_y S_x}{N} < 0$ or $S_y^2\left(1 - \rho^2 - \frac{1}{N}\right) - R^2 S_x^2\left(\frac{1}{\sqrt{N}}\right)^2 + 2R\rho N^{-1}S_y S_x < 0$. When $N \to \infty, \frac{1}{N} \to 0$ so that $S_y^2(1 - \rho^2) < 0$ or $S_y^2 - \rho^2 S_y^2 < 0$. Thus $S_y^2 < \rho^2 S_y^2$ is the necessary condition for regression estimator to be more efficient than ratio estimator.

Measures of Efficiencies

In this study, we utilize some measures of efficiency for our inference namely, coefficient of variation and Relative efficiency computed by

$$CV(\hat{Y}_R) = \frac{RMSE}{\hat{Y}_R}X\,100 \qquad (29)$$

and

$$CV(\hat{Y}_{lr}) = \frac{RMSE}{\hat{Y}_{lr}}X\,100 \qquad (30)$$

where the Root Mean Square Error is given as

$$RMSE(\hat{Y}_{lr}) = \sqrt{MSE(\hat{Y}_{lr})} \qquad (31)$$

or in terms of Relative Efficiency

$$RE\left(\frac{\hat{Y}_R\backslash n}{\hat{Y}_{lr}\backslash n}\right) = \frac{RMSE(\hat{Y}_{lr}\backslash n) - RMSE(\hat{Y}_R\backslash n)}{RMSE(\hat{Y}_{lr}\backslash n)} < 0 \qquad (32)$$

4. Results and Discussion

The results for the Design Based Estimation of total maize yields using ratio and regression estimators, their Mean Squared Error (MSE) and Root Mean Squared Error for sample of size 20 are shown on Tables 1, 2 and 3 as well as Figures 1, 2 and 3 respectively. Tables 4, 5 and 6 compare estimates of total maize yields using ratio and regression estimators, its Mean Squared Error (MSE) and Root Mean Squared Error for samples of size 20 for both ratio and regression estimators. The total maize yield for the samples of size n= 20 and n=30 were compared using ratio and regression estimators. The results shown in Figures 1, 2 and 3 reveal that regression estimator is better when the sample size is 20 and the two estimators are approximately equal for samples of size 30.

The total maize yield for the samples of size n= 20 and n=30 were estimated using ratio and regression estimators. Similarly the efficiency of the estimator determined by the relative efficiency (RE) as well as coefficient of variation (CV) is used as optimality criteria. Results showed that regression estimator is better when the sample size is 20 but when the sample size is 30, the two estimators tend to be approximately the same efficiencies at most points in time. This is in agreement with theoretical results of laws of large numbers (Neyman, 1934, Chaudhuri, 2010) that as n→ ∞, the anticipated variance tends to 0. Thus, throughout the period of estimation, CV (30) <CV (20) always. Similarly MSE (30)<MSE (20). In terms of Relative Efficiency (RE) and asymptotic value of the variance for n=20 and n=30, we discovered that the RE of the estimators for n = 30 is far less than 100% indicating that the variation for n = 30 is less than that of n= 20.

The results in Table 3 showed the expected Mean Squared error (RξMSE) and hence, the stability of the estimator. Thus, the model based RMSE of our estimator for n=20 and n=30 showed that the model's root mean squared error for n=30 is consistently better than that for n=20. In terms of stability of the estimator we have shown that, the estimator is stable under the condition of $\beta^2 + \hat{R}^2 \geq 2\hat{R}\beta$ or $\beta^2 + \hat{R}^2 - 2\hat{R}\beta \geq 0$.

Table 1. Estimate of Total Maize Yield using Ratio Estimator for samples of size n=20 and n=30 with the optimality Criteria utilized
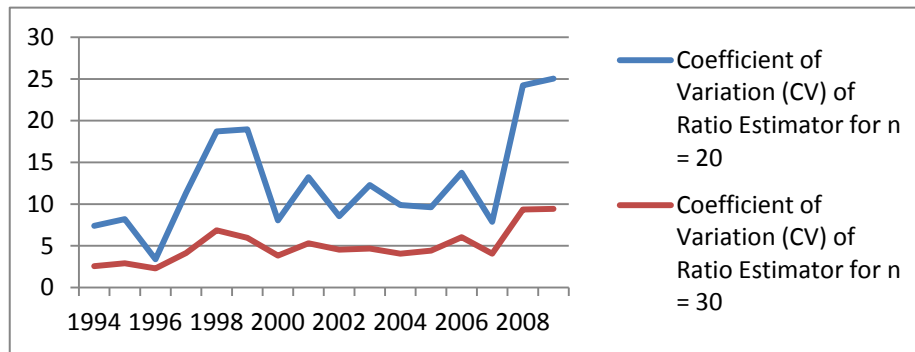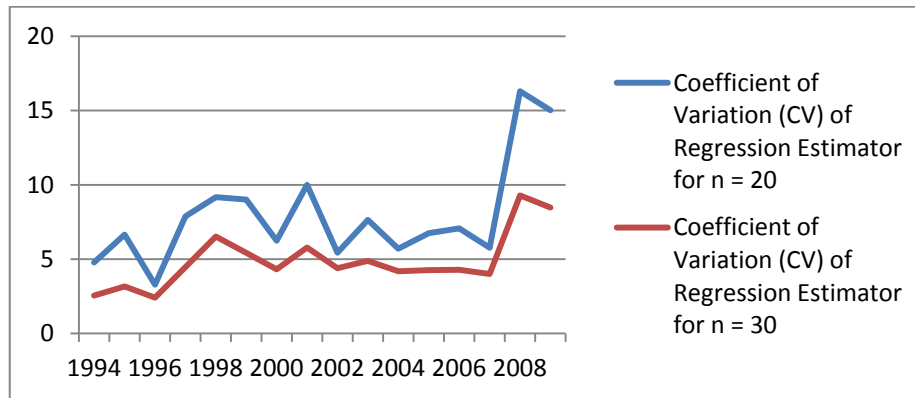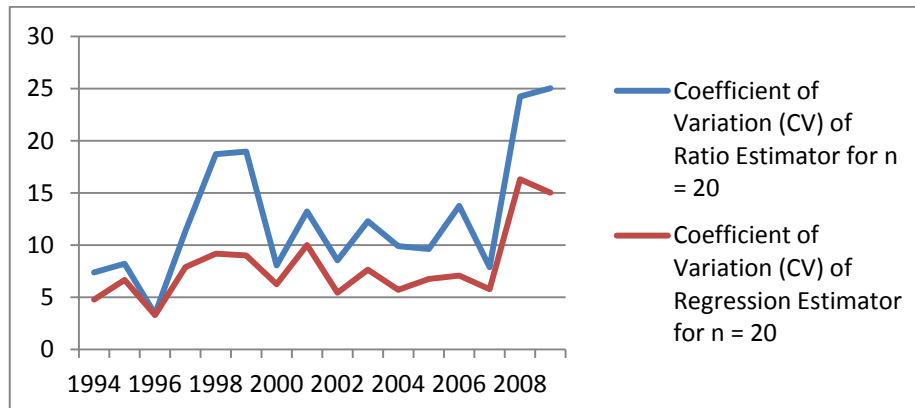
Year	Estimate of total maize yield for n = 20	Estimate of total maize yield for n = 30	Root Mean Error of Ratio Estimator for n = 20	Root Mean Error of Ratio Estimator for n = 30	Coefficient of Variation (CV) of Ratio Estimator for n = 20	Coefficient of Variation (CV) of Ratio Estimator for n = 30	Relative Efficiency of Ratio Estimator at n= 20 and n=30	Asymptotic Value of the Variance of n=20	Asymptotic Value of the Variance of n=30
1994	4781.176	5283.7	353.26	135.78	7.389	2.57	53.14	311.982	20.483
1995	4870.462	5453.6	399.6	159.08	8.205	2.917	54.58	399.21	28.118
1996	5159.583	4982.3	174.82	114.19	3.388	2.292	44.34	76.407	14.488
1997	5206.313	5720	590.67	236.3	11.345	4.131	53.99	872.224	62.042
1998	5257.723	5901.7	983.76	404.98	18.711	6.862	54.92	2419.444	182.23
1999	5395.698	6436.1	1023.02	383.52	18.96	5.959	55.36	2616.416	163.434
2000	5287.187	5438.4	425.14	208.61	8.041	3.836	52.72	451.857	48.354
2001	5621.424	6432.3	743.21	341.13	13.221	5.303	72.39	1380.899	129.301
2002	5198.033	5674.1	443.86	257.6	8.539	4.54	40.27	492.532	73.732
2003	5010.417	5577.2	615.8	261.13	12.29	4.682	53.71	948.021	75.765
2004	5217.042	5702.7	516.18	231.7	9.894	4.063	54.46	666.092	59.65
2005	5459.539	5672.6	525.37	251.48	9.623	4.433	53.79	690.037	70.272
2006	4772.711	5053.6	656.75	305.2	13.761	6.039	53.98	1078.314	103.496
2007	5741.597	5838.1	452.28	236.67	7.877	4.054	53.66	511.402	62.239
2008	7037.407	7695.7	1706.15	719.52	24.244	9.35	55.94	7277.351	575.226
2009	7084.103	7801.9	1773.46	736.36	25.034	9.438	97.21	7862.907	602.474

Table 2. Estimate of Total Maize Yield using regression Estimator for samples of size n=20 and n=30 with the optimality Criteria utilized

Year	Estimate of total maize yield for n = 20	Estimate of total maize yield for n = 30	Root Mean Error of Regression Estimator for n = 20	Root Mean Error of Regression Estimator for n=30	Coefficient of Variation (CV) of Regression Estimator for n = 20	Coefficient of Variation (CV) of Regression Estimator for n = 30	Relative Efficiency of Regression Estimator at n= 20 and n=30
1994	5631.967	5277.1	269.02	134.59	4.777	2.55	49.97
1995	5228.593	4996.8	348.05	158.21	6.657	3.166	54.54
1996	5077.971	4591.5	167.03	110.75	3.289	2.412	33.69
1997	5401.078	4909.1	425.9	219.15	7.886	4.464	48.54
1998	6996.629	5813.9	641.76	379.43	9.172	6.526	40.88
1999	7375.347	6422.5	664.62	348.61	9.011	5.428	47.55
2000	5449.787	4818.5	340.42	208.07	6.246	4.318	38.88
2001	6725.934	5789.4	672.49	334.58	9.999	5.779	50.25
2002	6220.836	5803.2	337.95	254.97	5.433	4.394	24.55
2003	5906.903	5253.9	451.41	256.9	7.642	4.89	43.09
2004	6164.659	5496.6	352	230.23	5.71	4.189	34.59
2005	6589.589	5888.2	444.85	251.03	6.751	4.263	43.57
2006	7920.847	7112	560.01	304.88	7.07	4.287	45.56
2007	6641.342	5894.6	383.52	236.38	5.775	4.01	38.37
2008	8312.778	7288.6	1355.25	676.87	16.303	9.287	50.06
2009	9162.115	8078.4	1376.18	684.9	15.02	8.478	50.23

Table 3.   Model Based Mean Squared Error, Stability and Root Expected Mean Squared Error for the Estimator at n=20

Year	ξMSE	Stability	RξMSE	Year	ξMSE	Stability	RξMSE
1994	19108.73	0.004963	138.2343	2002	336358.9	1.068603	579.9646
1995	372545.9	1.294633	610.3654	2003	293388.4	0.948104	541.6534
1996	369813.1	1.402675	608.1226	2004	284615.4	0.860547	533.4936
1997	246489.9	0.962422	496.4775	2005	342594.9	0.886086	585.316
1998	199977.3	0.602674	447.1882	2006	1018399	1.339957	1009.158
1999	170333.1	0.508756	412.7142	2007	312434.8	0.761183	558.9587
2000	266842.7	1.033927	516.5682	2008	32321.61	0.040752	179.7821
2001	96318.41	0.340505	310.3521	2009	47810.4	0.074423	218.6559

Figure 1.   Coefficient of Variation (CV) of Ratio Estimator for n= 20 and n=30

Figure 2.   Coefficient of Variation (CV) of Regression Estimator for n=20 and n=30

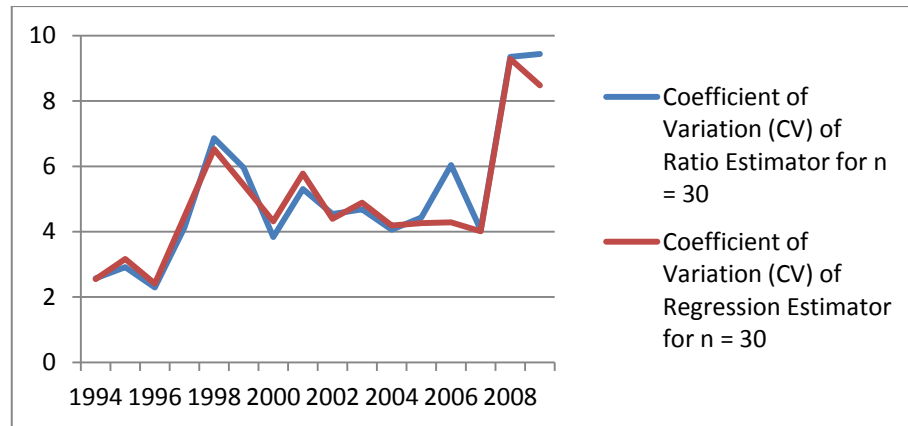**Figure 3.**   Comparison of CV of Ratio and Regression Estimators for n=20

Figure 4. Comparison of CV of Ratio and Regression estimators for n=30

5. Concluding Remark

The study utilized Design based Ratio and regression Estimators to provide estimates of total maize yields in Nigeria from 1994 to 2009. The study investigated the performance of the estimator using MSE and CV (and hence, the corresponding model based estimates of MSE and CV) criteria for samples of sizes 20 and 30. The Mean Squared Error and Coefficient of Variation obtained showed that the regression estimator performed better than ratio estimator for sample of size 20. However, for n = 30, the two estimators appear to be converging with each other. The model based MSE also is a pointer to the fact that large sample size is preferable. Importantly, the estimators have succeeded in estimating total maize yields in Nigeria whose trend is increasing with time.

We conclude that both the ratio and regression estimators are useful in providing estimates of crops yields so as to provide needed information to evolve policy that will ensure Food Security in Nigeria especially, when data at state levels are available and large sample size is considered. Similarly, the trend shows an increasing yield over time. However, when small sample size is of interest, regression estimator provides better estimates of population characteristics.

REFERENCES

[1]   Abayomi, Y.O. (1997). *The Agricultural Sector in Nigeria, the Way Forward,* 21(3), Local Government Finance Office, Research Department, Central Bank of Nigeria.

[2]   Anneke, F. and Todd B. (2011). *Estimating Yield of food Crops Grown by Smallholders Farmers,* A Review in the Uganda Context. IFPRI Discussion Paper 01097.

[3]   Arunachalam and Balakrishnan (2012). Statistical Models for Wheat Production, International *Journal of Statistics and Application*, 2(4); 40-46.

[4]   Bellow, M., and Lahiri, P. (2012) "*Evaluation of Methods for County Level Estimation of Crop Harvested Area that Employ Mixed Models"*, Proceedings of the DC-AAPOR /WSS Summer Conference, American Statistical Association, Bethesda Maryland.

[5]   Bornn, L. and Zidek, J.V. (2012). Efficient Stabilization of crop yield production in the *Canadian Prairies Agricultural and Forestry Methodology* 152: 223-232.

[6]   Brus, D.J. and de Gruijter, J.J. (1997). *Random Sampling on Geo-statistical Modelling; focusing between design based and model-based sampling Strategies for Soils* Geoderma 80: 1-44.

[7]   Chauldhuri, A. (2010). *Essential of Survey Sampling:* PHI Learning Limited, New Delhi-1100 Indian.

[8]   Cochran, W.G. (1940). The Estimation of the Yields of Cereal Experiments by Sampling for the Ratio Gain to Total Produce. *Jour. Agric. Soc.* 30: 262-275.

[9]   Cochran, W.G. (1977). *Sampling Techniques,* Wiley Eastern Limited, New York: John and Sons.

[10] David, I. P. (1978). "*Non-sampling Errors in Agricultural Surveys Review, Current Findings and Suggestions for Future Research.*" Paper presented at the Philippine Statistical Association Annual Conference, June 19, Manila, Philippines.

[11] FAO (Food and Agriculture Organization of the United Nations). (1982). *The Estimation of Crop Areas and Yields in Agricultural Statistics. Economic and Social Development Paper No. 22.* Rome, Italy.

[12] Fulginiti, L.E. and PerriR.K (1998). Agricultural Productivity in Developing Countries. *Agricultural Eonomics.* 19 (1-2): 45-51.

[13] Makwoski D. and Micheal L (2013). *Use of dynamic linear model for predicting crop yield trends in foresight studies on food security.* ICAS VI, 6th International Conference on Agricultural Statistics.

[14] Martis, J.H, Gaito, T, Grant, W.E, Irwing, W.C and Ritchie, J.C (1988). A Markov Chain Approach to Crop Yield Forecasting. *Agricultural Systems*, 18: 171-187.

[15] Michael, E. B. (2007). *Improved County-Level Estimation of Crop Yield Using Mode l- Based Methodology* with a Spatial Component.

[16] Murphy, J., CasleyD. J., and CurryJ. J. (1991). *Farmers' Estimations as a Source of Production Data. World Bank Technical Paper 132.* Washington, DC: World Bank.

[17] National Bureau of Statistics, (NBS, 2007). *Filling Data Gaps in Agricultural Statistics.* NBS (2014) Gross Domestic Product Rebasing Publication.

[18] Okafor, F.O. (2002). *Sample Survey Theory with Application.* 1st Edition, University of Nigeria, Nsukka, Afro-Orbis Publication Ltd.

[19] Olayide, S.O and Heady, E. (1982). *Introduction to Agricultural Production Economics.* Ibadan University Press, Nigeria.

[20] Raj, D. (1965). *On a method of using multi-auxiliary information in sample Surveys,* Journal of the American Statistical Association, 60: 154–165.

[21] Rao, J.N.K., Hartley, H.O and Cochram, W.G. (1962). *On a simple procedure of Unequal Probability Sampling Without replacement. Jour. Roy. Stat. Soc. Ser* B, 24: 422-491.

[22] Wood, J. (2011). Mathius, *Famine Disease and World Population* Retrieved 24 Jan 2015. Available from http//www.malthus.famine.disease and world.Population/htm/.

[23] Yates, F. (1960). *Sampling Methods for Censuses and Surveys*. Charles Griffin & Company Limited, London.