Pakistan Journal of Social Sciences 4 (2): 252-260, 2007

Asymmetry in Agricultural Data Sources: Implication for Policy Formulation in the Last Three Decades in Nigeria

[1]Udoh Edet Joshua and [2]Omonona Bolarin Titus
[1]Department of Agricultural Economics and Extension University of Uyo, Uyo, Nigeria
[2]Department of Agricultural Economics, University of Ibadan, Ibadan, Nigeria

Abstract: This study discusses the statistical inconsistency in agricultural data sources and its implications for policy formulation in the last three decades. Time series data of some crops output from FOS, FAO, USDA and CBN are statistically analyzed to test the likelihood of data asymmetry among the different sources. The results of the analysis recognize the existence of data asymmetry; among the four data sources for the period under review. This, therefore, makes the estimations based on these different data sources to result in conflicting and contrary estimates and indices as shown in the constant growth rate analysis. The study then concludes by recognizing that analysis on rather inconsistent data series not only systematically fail to predict outcomes correctly but also becomes increasingly obvious that the predictive failure will lead to wrongful policy statements. Therefore, the study recognizes the need for the observed asymmetry to be removed. It calls for actions to be taken to harmonize and unify the different data sources. Hence, there is an urgent need for conducting agricultural census in the country in the new millennium.

Key words: Asymmetry, agricultiral, imlication, crops, policy formulation

INTRODUCTION

It is a well-established fact that a general framework of policy identification and formulation is based on analysis of economic data. In other words, the synthesis of policy necessary for development efforts precedes analysis of data. Modeling and simulation are based on the nature and type of data available. Inconsistency in data could have direct transmission into faulty modeling and simulation with resultant faulty policy initiation. No matter how robust an analytical technique is, the reliability of its parameter estimates or indices depends on the type of data used. Too frequently, poor econometric techniques result in inappropriate inferences being drawn from poor data (Adams, 1992). The popular saying that, "garbage in, garbage out" holds sway in any empirical study when the data used are faulty.

From empirical point, the literature is awash with various quantitative studies that are similar in design, scope and analytical framework, but have different parameter estimates and indices that are statistically different from one another. The culprit for these differences is partly the source and type of data used. The existence of essential distinct sources of data impart great problem on empirical studies.

Policy analysts, econometricians, agricultural economists and statisticians have tried, through various means to `fine tune' data to obey economic theories, econometric and statistical criteria and expectations. The adverse impact of data asymmetry is further compounded by lack of consensus and unified data source. While some statistical bodies such as the Federal Office of Statistics (FOS) collect, collate, analyze and publish different data on the same economic variable at the same time others such as the Central Bank of Nigeria (CBN) only publish these data after some refinements. These different data, though on the same variables, inherently have peculiar econometric and statistical problems and within the context of policy analysis, different policy options are developed as occasioned by available data series. The implication of asymmetric data source for policy framework is the development of wrong premise for policy formulation. The problem perhaps could be one of the major reasons for non-workability of well-articulated policies in the real world.

On the point that some statistical bodies such as CBN do not produce but report data collected by the other agencies such as FOS, one expects that there should not be any difference between what is reported by the agency that only report what some others collect. But this is not true, as significant differences exist, for example, between the crop output of the FOS and other agencies that are reporters of such data. This, therefore, shows the need to further test for asymmetry in crop output data from different sources in Nigeria.

Corresponding Author: Udoh Edet Joshua, Department of Agricultural Economics and Extension, University of Uyo, Uyo, Nigeria

In Nigeria, policy analysts and different governments formulate policies and develop programs/projects based on the available information, which in most cases have been unworkable and undesirable (Idachaba, 1997 and 2000). In the past decades and even now, economic researchers depend on official sources for economic data. These are FOS, FAO (Food and Agriculture Organization), CBN, USDA (United States Department for Agriculture), IMF (International Monetary Fund), World Bank, etc. These official sources are however distinct in nature and ownership. Thus, each has a way of organizing and collecting macro-economic data. This uniqueness in organization and collection of data by each body has resulted in different sets of data for particular variables. This perhaps is a violation of the principle of sampling survey (Cochran, 1977). As such, it is possible to have different results and inferences for a particular variable by using different data sources. Therefore, within the domain of time-series econometrics, it is intuitively clear that estimation(s) based on these `unparallel' data would lead to wrong result(s) that are inconsistent and non-convergent. Perhaps, this could explain the existence of non-stationarity in macro economic time-series (income, consumption, output, money and trade, etc) data (Adams, 1992). Capturing the salient features of data from different sources that provide different estimates to economic relationship could be a wasteful exercise. Specifically, within the context of data admissibility and dynamic specification, this could present unacceptable result (Hendry and Richad, 1982).

Not minding that there are differences in the technique of data collection, sample sizes and approach to data collection by these agencies/organizations, it is expected that the differences in the output of a crop, as published by different organisations, to be insignificant for Nigeria in any particular year. This is because; these data form the basis for policy formulation and analysis. As long as this is one of the reasons why crop output data are published by these organisations, then a comparison of these data is not out of place. In addition, since these crops output data were measured in metric tonnes, there seems to be no basis to doubt the comparison of the different data sources. If data on a particular crop as reported by an organization is different from the other, then any two independent researchers having access to that particular crop output data but from different sources are likely to make different conclusions (in form of policy statements), which do not augur well for the development of the country.

According to Hendry and Richard (1982), Spanos (1986), building econometric models on rather inconsistent data not only systematically fail to predict outcomes, but becomes increasingly obvious that the predictive failure would lead to wrong policy statements. The issue of interest, therefore, centers squarely on how reliable the various data sources available to economic analysts in Nigeria are. Could official data (especially government data) be taken as suggestive of a possible line of inquiry and not as absolutely accurate? The essence of the problem is with the presence of uncorrelated random walks and statistically different diagnostic statistics between and among time-series data from different official sources available in Nigeria. It should be noted that the future of economic research and policy development depends on the type of data available for analysis and data generating process.

The main thesis of this study is that if two sets of data for a particular economic variable are drawn from the same population they should be closely related (Johnson and Dinardo, 1997). The crucial question is whether there are inconsistencies between the various data sources or whether they are correlated to the level that each can be used interchangeably. This study then seeks to consider cross correlation among different official data sources to provide conclusive evidence for the existence of a single data pool or other wise. The analysis is therefore based on output levels of some selected agricultural crops as reported by FOS, FAO, USDA and CBN in their various bulletins and publications. The choice of agricultural data is based on the inherent difficulties encountered by researchers in studying and obtaining data from heterogeneously small-scale farmers who are unevenly scattered in the country (Olayemi, 1980). In essence, there exist strong likelihood for data asymmetry for different sources of survey procedures adopted by each of the sources.

Sellitz *et al.* (1976) have identified several major sources of measurement errors in data but Ariyo (1996) summarized them as four. These are those occurring through the respondents, situational errors, the measurer as an error and the error due to the measurement instrument employed. Ariyo and Adeniran (1998) however, came up with three causes of divergence among macro-economic data in African countries. These are the problems of differences in data processing and management procedures, especially the non-response by member countries and non-conformance with the reporting format of the data-publishing agency. The second category relates to the level of cooperation from the respondent organisations in terms of accuracy and timeliness of responses or rendition of required returns. The third relates to the absence of generally accepted formats for reporting data. In addition, the process of converting and reprocessing returns to the required format also generates its own errors. This is probably one of the problems of the data reporting agencies like CBN.

THEORETICAL AND ANALYTICAL FRAMEWORK

Theoretical framework: The axiom of this study is that if two different data sources for a particular variable (crop output) are positively and highly correlated, then the results of any analysis based on each data set will be plausibly similar. But if they are positively uncorrelated, the results obtained from any analysis will be different.

Suppose two series of data from sources A and B are known to be correlated and generated by the following known processes:

$$A_t = A_{t-1} + U_t;\ U_t \sim \text{IId}\,(0, \sigma u^2) \qquad (1)$$

$$B_t = B_{t-1} + V_t;\ U_t \sim \text{IId}\,(0, \sigma v^2) \qquad (2)$$

then,

$E(U_t, V_s) = 1\ \forall_t$,s and
$E(U_t, V_{t-k}) = E(V_t, V_{t-k})$ and $O\forall K$; or otherwise.

The above restrictions ensure that data sources A ad B are correlated random variables and strongly reinforce data symmetry, therefore supporting the thesis of a single a data pool. Consequently we would expect the correlation coefficient (r_{AB}) to be equal to one in case of symmetry and less than unity in the case of asymmetry of data. Besides, the sign of r_{AB} for symmetrical data should be positive and otherwise for asymmetric data.
The collorary of the above thesis is

$\theta_{IA} \approx \theta_{Ib} \ldots\ldots\ldots\ldots\ldots\ldots \theta_{iN}$

where θ_{IA} represents output data of ith crop from source A and so on. The identity in data sources (represented by a correlation coefficient of 1) shows the existence of close substitution and consistency among the data sets.

Analytical framework: The study relied on time-series data of output levels of maize, millet, sorghum, rice, yam, cassava and wheat as reported in FOS, FAO, USDA and CBN statistical bulletins, collected between 1960 and 1999 inclusive.

The study, therefore, employed the Analysis of Variance (ANOVA) and correlation analysis to give statistical nature and behavior of the data sources with regard to the level of association among them. The pair-wise correlation coefficient for the data set is expressed as

$$r = \frac{n\sum_{i=1}^{m}\sum_{j=1}^{p} n_{ij}X_iY_j - (\sum_{i=1}^{m} n_iX_i)(\sum_{j=1}^{p} n_iY_j}{\sqrt{n\sum_{i=1}^{m} n_iX_i^2 - (\sum_{i=1}^{m} n_iX_i)^2}\sqrt{n\sum_{j=1}^{p} n_jY_j^2 - (\sum_{j=1}^{p} n_iY_j)^2}} \qquad (3)$$

where X_i and Y_j represent data series from sources A and B; and n represents the number of observations.

To further test for the level of consistency of each data source, a growth rate model is also employed. Based on the nature of data considered for the study, we assume an exponential growth model as

$$Y_t = Y_o e^{\beta t} \qquad (4)$$

This is linearised into

$$LnY_t = \alpha + \beta t \qquad (5)$$

Where $\alpha = \ln Y_o$
The resultant slope coefficient gives a compound growth rate as

$$g_i = e^b - 1 \qquad (6)$$

when $g_{iA} = g_{iB} = g_{iC} = g_{in}$, then the hypothesis of data consistency and close substitution is accepted for FOS, FAO, USDA and CBN data series. As such, prediction and policy based on either of the data will give consistent and identical result.

RESULTS AND DISCUSSION

Data series: The output data from the Federal Office of Statistics, (FOS), Food and Agricultural Organization (FAO), United States Department of Agriculture (USDA) and the Central Bank of Nigeria (CBN) are presented for maize, millet, rice, sorghum, yam, cassava and wheat in that order in Table 1a to 1g (See the appendix).

A cursory look at the data shows that the FOS and FAO report crop output data for all the years under consideration. While the USDA crop output were not reported as from the early 1980s for all the crops in addition to that of wheat that was not reported for the 1960s. As for that of CBN, crop output data were not available for the 1960s but thereafter were available. For all data series, it is noticeable that variations exist among the different sources. When the series is rising in one source it is declining in other source. Statistical and regression analyses presented in the subsequent sections give clearer view of the level of inconsistency among the sources.

Table 1: Test of Inconsistency in Output Data of Crops

Crops	F-Statistics	Probability	Decision
Maize	13.7648	0.0187	$\sigma_i^2 = \alpha_{ii}^2$ at $\alpha = 0.05$
Millet	16.4687	0.0097	$\sigma_i^2 = \alpha_{ii}^2$ at $\alpha = 0.01$
Sorghum	10.9659	0.2145	$\sigma_i^2 \neq \alpha_{ii}^2$ N.S
Rice	14.334	0.068	$\sigma_i^2 = \alpha_{ii}^2$ at $\alpha = 0.05$
Yam	0.9624	0.4144	$\sigma_i^2 \neq \alpha_{ii}^2$ N.S
Cassava	9.1858	0.12	$\sigma_i^2 = \alpha_{ii}^2$ at $\alpha = 0.01$
Wheat	10.6466	0.0001	

Source: ANOVA Result of Crop output at from different sources.

Analysis of Variance (ANOVA): This is based on reliability test to statistically test the null hypothesis of no significant difference among the various data sources for each of the crop output. The F statistics for each set of crops are presented in Table 1.

The reliability test shows that there exist consistency and similarity among the various data sources for some crops and inconsistency for some crops. The symmetry in data sources exists for maize, millet, rice and wheat, but not for sorghum, yam and cassava. It therefore follows that for the period considered, output levels of maize, millet, rice and wheat as reported in FOS, FAO, USDA and CBN statistical bulletin were similar, but different for sorghum, yam and cassava.

Correlation analysis: The two-way ANOVA shows there exists data asymmetry among the various data sources for some crops output. But the correlation matrices presented in Table 2 explain the extent to which the various data sources associated among themselves.

The correlation coefficients among all the data sources indicate different degrees of association, ranging from weak to strong and negative to positive. There is no case of perfect association among the data sources indicating therefore that there exist differences among the various data sources. The strongest degree of association is observed among the crop output data obtained from the FOS, FAO and CBN. But USDA data series have inverse association with other sources. On crop-by-crop basis, association is high in maize, millet, rice and wheat but low in sorghum, yam and cassava. This is however, consistent with the result of the reliability test. By inference, therefore, while some data sources reported seemingly similar output figures, the data series from these sources do not tally. Specifically, the data from USDA for the output of all the crops differed significantly from others.

Constant growth rate analysis: The results of reliability test and correlation analysis strongly reinforce data asymmetry and reject the thesis of single data pool for the crops considered. As such there is a strong likelihood that estimations based on these data from different sources will result in conflicting and contrary estimates and indices, thus different policy relevance and implications.

Table 2: Crop Output Correlation Matrix

	FOS	FAO	USDA	CBN
Correlation matrix for Maize				
FOS	1			
FAO	0.982		1	
USDA	-0.772	-0.721	1	
CBN	0.945	0.948	-0.817	1
Correlation matrix for Millet				
	FOS	FAO	USDA	CBN
FOS	1			
FAO	0.914	1		
USDA	-0.717	-0.741	1	
CBN	0.792	0.832	-0.651	1
Correlation matrix for Sorghum				
	FOS	FAO	USDA	CBN
FOS	1			
FAO	0.972	1		
USDA	-0.572	-0.413	1	
CBN	0.996	0.972	-0.574	1
Correlation matrix for Rice				
	FOS	FAO	USDA	CBN
FOS	1			
FAO	0.759	1		
USDA	-0.529	-0.7	1	
CBN	0.634	0.776	-0.787	1
Correlation matrix for Yam				
	FOS	FAO	USDA	CBN
FOS	1			
FAO	0.432	1		
USDA	-0.243	-0.178	1	
CBN	0.503	0.425	-0.503	1
Correlation matrix for Cassava				
	FOS	FAO	USDA	CBN
FOS	1			
FAO	0.863	1		
USDA	-0.633	-0.346	1	
CBN	0.993	0.839	-0.679	1
Correlation matrix for wheat				
	FOS	FAO	USDA	CBN
FOS	1			
FAO	0.999	1		
USDA	-0.3	-0.299	1	
CBN	0.999	0.997	-0.284	1

Source: Correlation Analysis Results

The above premise is verified and confirmed with the estimation of constant growth rate for each source o f the crop data. The results are presented as follow in Table 3.

DW is Durbin Watson statistics, r^2 is the coefficient of determination and g is the constant growth rate. Figures in parentheses are standard errors and *** denotes significance at 1% level.

The results of trend analysis for all the crops confirm the thesis of data asymmetry. Estimation involving the different crop data sources results in different parameter estimates that are to an extent, statistically different. In all the cases, the intercepts and responsive coefficients are different in terms of signs, magnitudes and level of significance. But specifically, the difference is much noticeable in the USDA data and a bit in CBN data.

Models as those estimated above, should be consistent with the data sets irrespective of the source of

Table 3: The growth rate for each source of the crop data.

Growth model for Maize				
$lnY_{(FOS)}$	= 6.245 +	0.0611t	r^2 = 0.747	DW = 0.667
	(0.203)***	(0.009)***	$g_{(FOS)}$ = 6.3%	
$lnY_{(FAO)}$	= 6.358 +	0.0624t	r^2 = 0.840	DW = 0.667
	(0.150)***	(0.007)***	$g_{(FAO)}$ = 6.3%	
$lnY_{(USDA)}$	= 9.378 +	0.269t	r^2 = 0.84	DW = 0.387
	(0.659)	(0.029)***	$g_{(USDA)}$ = -23%	
$lnY_{(CBN)}$	= 1.188 +	0.248t	r^2 = 0.840	DW = 0.490
	(0.596)***	(0.027)***	$g_{(CBN)}$ = 28%	
Growth model for Millet				
$lnY_{(FOS)}$	= 7.666 +	0.0234t	r^2 = 0.786	DW = 1.041
	(0.069)***	(0.003)***	$g_{(FOS)}$ = 2.3%	
$lnY_{(FAO)}$	= 7.665 +	0.0248t	r^2 = 0.797	DW = 1.588
	(0.047)***	(0.002)***	$g_{(FAO)}$ = 2.5%	
$lnY_{(USDA)}$	= 10.487 +	0.303t	r^2 = 0.848	DW = 0.396
	(0.705)	(0.032)***	$g_{(USDA)}$ = -26%	
$lnY_{(CBN)}$	= 2.017 +	0.229t	r^2 = 0.747	DW = 0.338
	(0.758)***	(0.034)***	$g_{(CBN)}$ = 26%	
Growth model for Rice				
$lnY_{(FOS)}$	= 4.783 +	0.0794t	r^2 = 0.820	DW = 0.249
	(0.236)***	(0.011)***	$g_{(FOS)}$ = 8.3%	
$lnY_{(FAO)}$	= 5.005 +	0.0794t	r^2 = 0.808	DW = 0.628
	(0.215)***	(0.011)***	$g_{(FAO)}$ = 8.2%	
$lnY_{(USDA)}$	= 7.9964 +	0.227t	r^2 = 0.809	DW = 0.487
	(0.614)	(0.027)***	$g_{(USDA)}$ = -20.3%	
$lnY_{(CBN)}$	= 0.387 +	0.235t	r^2 = 0.867	DW = 0.422
	(0.502)***	(0.022)***	$g_{(CBN)}$ = 26.5%	
Growth model for Sorghum				
$lnY_{(FOS)}$	= 8.039 +	0.10149t	r^2 = 0.518	DW = 0.653
	(0.092)***	(0.004)***	$g_{(FOS)}$ = 1.5%	
$lnY_{(FAO)}$	= 8.019 +	0.021t	r^2 = 0.777	DW = 0.727
	(0.061)***	(0.003)***	$g_{(FAO)}$ = 2.1%	
$lnY_{(USDA)}$	= 10.810+	0.322t	r^2 = 0.863	DW = 0.414
	(0.702)	(0.031)***	$g_{(USDA)}$ = -27%	
$lnY_{(CBN)}$	= 1.420 +	0.256t	r^2 = 0.783	DW = 0.362
	(0.758)***	(0.034)***	$g_{(CBN)}$ =29%	
Growth model for Yam				
$lnY_{(FOS)}$	= 8.235 +	0.0623t	r^2 = 0.626	DW = 1.0519
	(0.256)***	(0.015)***	$g_{(FOS)}$ = 6.4%	
$lnY_{(FAO)}$	= 8.1587+	0.0754t	r^2 = 0.534	DW = 1.509
	(0.395)***	(0.023)***	$g_{(FAO)}$ = 7.8%	
$lnY_{(USDA)}$	= 11.8827 +	0.4999t	r^2 = 0.860	DW = 0.526
	(0.979)***	(0.057)***	$g_{(USDA)}$ = -39.3%	
$lnY_{(CBN)}$	= 8.345 +	0.0507t	r^2 = 0.671	DW = 0.163
	(0.185)***	(0.011)***	$g_{(CBN)}$ = 5.2%	
Growth model for Cassava				
$lnY_{(FOS)}$	= 6.928 +	0.1226t	r^2 = 0.749	DW = 0.4947
	(0.358)***	(0.021)***	$g_{(FOS)}$ = 13.04%	
$lnY_{(FAO)}$	8.607 +	0.0488t	r^2 = 0.468	DW = 1.049
	(0.305)***	(0.018)***	$g_{(FAO)}$ = 5.0%	
$lnY_{(USDA)}$	= 10.2028 +	0.4459t	r^2 = 0.826	DW = 0.494
	(1.004)***	(0.058)***	$g_{(USDA)}$ = -35.9%	
$lnY_{(CBN)}$	= 7.234 +	0.1165t	r^2 = 0.875	DW = 0.483
	(0.213)***	(0.012)***	$g_{(CBN)}$ = 12.36%	
Growth model for Wheat				
$lnY_{(FOS)}$	= 2.432 +	0.064t	r^2 = 0.625	DW = 0.399
	(0.302)***	(0.013)***	$g_{(FOS)}$ = 6.7%	
$lnY_{(FAO)}$	1.822 +	0.0831t	r^2 = 0.679	DW = 0.389
	(0.335)***	(0.015)***	$g_{(FAO)}$ = 8.7%	
$lnY_{(USDA)}$	= 1.334 +	0.0232t	r^2 = 0.236	DW = 0.330
	(0.356)***	(0.016)***	$g_{(USDA)}$ = -2.3%	
$lnY_{(CBN)}$	= 0.119 +	0.147t	r^2 = 0.820	DW = 0.419
	(0.383)***	(0.017)***	$g_{(CBN)}$ = 15.8%	

data and should inform our understanding of economic phenomenon/structure and/or provide the basis of testing implications of economic theory. Therefore, the validity of models as conditioned by the available data set is based on the precision, efficiency and consistency of estimates. In as much as each source of data generates a unique parameter estimate, a researcher may be wrong in choosing BLU (best linear and unbiased) estimates upon which to base policy options on. This is the case where we have different annual constant growth rates for each crop output from the different data sources considered. Though, a bit of similarity in growth rates for all crops from FOS and FAO data is noticed, much differences are observed for USDA and CBN data sources. Quite interestingly, USDA source produces negative growth rates in all the cases. Since the growth rate for output of each crop from different data source is not identical, the thesis of data sources consistency and close substitution is rejected.

Therefore, there is bound to be conflicting results as macro-economic researchers use different data sources for a particular variable. Thus, there exists no such thing as a `true' model (least of all a prior) for a particular economic variable if data available are asymmetric as observed in the case of output levels of some crops in Nigeria.

Since in empirical research, data sets generally consist of a relatively small sample of observation on a large number of variable, statistical manageability and economic clarity required consistency in data set so that inference made from one (sample) data set is similar to other (samples) data sets drawn from the same population. The high level of asymmetry noticeable among the various data sources imposes problem capable of destroying the validity of empirical results. Therefore, the policy framework based on these asymmetric data sources leave much to be desired.

APPENDIX

Table 1(a): Maize production (1961-1998)

Years	FOS	FAO	USDA	CBN
1961	1107	1000	1000	-
1962	990	900	900	-
1963	1167	1050	1050	-
1964	1118	1090	1090	-
1965	1158	1040	1040	-
1966	1131	1020	1020	-
1967	1098	1000	1000	-
1968	1056	950	950	-
1969	1566	1426	1426	1443
1970	427	1376	1310	1274
1971	1322	931	930	639
1972	830	1188	1182	808
1973	690	441	1287	528
1974	724	980	1350	1332
1975	769	1260	1400	1668
1976	1253	1293	1440	650
1977	943	1350	1550	658
1978	643	1480	1640	488
1979	608	1500	1670	612
1980	1550	1720	1703	720
1981	620	1580	1750	766
1982	628	1650	1775	594
1983	594	594	-	2121
1984	2058	2218	-	1250
1985	1190	1229	-	1342
1986	1336	1300	-	4721
1987	4612	4720	-	5342
1988	5268	5321	-	5067
1989	5008	5182	-	5882
1990	5768	5670	-	5826

Table 1(a): Continued

1991	5810	5821	-	5840
1992	5840	5832	-	5840
1995	6290	6352	-	6574
1994	6290	6961	-	6888
1995	7240	7440	-	7322
1996	6217	6553	-	6720
1997	6285	6325	-	6300
1998	6435	6545	-	6475
1999	6742	6895	-	6789

Table 1(b): Millet production (1961-1998)

Years	FOS	FAO	USDA	CBN
1961	2484	2484	2484	-
1962	2600	2600	2600	-
1963	2530	2530	2530	-
1964	2732	2732	2732	-
1965	2729	2729	2729	-
1966	1749	1747	1747	-
1967	2590	2590	2590	-
1968	2196	2196	2196	-
1969	3298	3298	3298	3106
1970	3077	3284	3284	2834
1971	2911	2688	2688	2391
1972	2524	3414	3048	3794
1973	2882	2350	2150	5554
1974	4322	3000	2800	2552
1975	4653	3000	2865	2893
1976	2653	2900	2865	2579
1977	2779	2950	2950	2386
1978	2521	3060	3100	2386
1979	2171	3130	3130	2354
1980	2756	3130	3130	2682
1981	2096	3230	3180	2666
1982	3034	3300	3240	2783
1983	2783	2783	-	3349
1984	3349	3527	-	3349
1985	3684	3782	-	3521
1986	4111	4221	-	4272
1987	3905	3850	-	3754
1988	5136	5363	-	5228
1989	4770	4850	-	7452
1990	5136	5500	-	5235
1991	4109	4572	-	4225
1992	4501	4770	-	4474
1995	4602	4774	-	4627
1994	4757	4785	-	4762
1995	4900	4981	-	4886
1996	5803	5532	-	5721
1997	5997	5768	-	5828
1998	6328	6473	-	6426
1999	6741	6789	-	6854

Table 1(c): Rice production (1961-1998)

Years	FOS	FAO	USDA	CBN
1961	133	119	354	-
1962	257	258	370	-
1963	195	196	304	-
1964	220	221	405	-
1965	231	232	355	-
1966	199	200	406	-
1967	385	386	391	-
1968	353	353	375	-
1969	325	325	386	-
1970	293	293	490	280
1971	279	580	462	279
1972	397	600	66	447
1973	459	342	514	487
1974	498	356	523	525
1975	522	575	586	504

Table 1(c): Continued

1976	426	524	611	218
1977	276	647	620	410
1978	372	842	826	280
1979	235	750	871	160
1980	245	1090	596	105
1981	257	1240	1165	158
1982	288	1400	1380	212
1983	145	155	-	145
1984	157	167	-	157
1985	196	198	-	196
1986	283	285	-	283
1987	808	800	-	716
1988	2081	2093	-	2072
1989	3303	3203	-	3303
1990	2500	2520	-	2503
1991	3226	3236	-	3196
1992	3260	3190	-	3240
1995	3065	3165	-	3075
1994	2427	2528	-	2467
1995	2920	2930	-	2924
1996	3122	3142	-	3125
1997	3230	3320	-	3244
1998	3486	3492	-	3480
1999	3698	3665	-	3645

Table 1(d): Sorghum production v(1961-1998)

Years	FOS	FAO	USDA	CBN
1961	3985	3966	3966	-
1962	4499	4506	4509	-
1963	4060	4069	4069	-
1964	4230	4239	4239	-
1965	4226	4235	4235	-
1966	3154	3160	3160	-
1967	3382	3389	3389	-
1968	2843	2821	2821	-
1969	3941	4080	4080	-
1970	4044	4080	4080	4053
1971	5265	3140	3140	3794
1972	3367	3988	3561	2298
1973	2546	2968	2968	3125
1974	3609	3500	3590	4738
1975	4204	3590	3680	2920
1976	2933	3680	3750	2950
1977	3051	3750	3785	3286
1978	3028	3770	3800	2409
1979	2719	3785	3750	2644
1980	3303	3800	3825	3364
1981	3475	3750	-	3364
1982	3626	3800	-	3740
1983	3292	3283	-	3294
1984	4608	4708	-	4628
1985	4911	4901	-	4922
1986	5432	5455	-	5445
1987	5465	5455	-	5455
1988	5182	5242	-	5193
1989	3303	7265	-	7265
1990	2500	6265	-	4185
1991	3226	5367	-	5367
1992	5909	5619	-	5902
1995	6051	6121	-	6224
1994	6197	6243	-	6321
1995	6377	6323	-	6287
1996	7514	7614	-	7454
1997	7954	7941	-	7952
1998	8401	8412	-	8406
1999	9874	9564	-	925

Table 1(e): Yam production (1970-1998)

Years	FOS	FAO	USDA	CBN
1970	8230	12033	16200	12033
1971	7990	9766	16404	9766
1972	7758	9600	16107	6900
1973	6911	9636	18800	6941
1974	9003	7160	17200	7120
1975	7598	8620	17600	8512
1976	7948	6470	18000	6512
1977	6342	6376	18000	6423
1978	6223	5866	18100	5823
1979	5588	5256	18100	5352
1980	5120	5248	18180	5316
1981	5120	5212	18200	5232
1982	5397	5385	18480	5271
1983	4047	4047	-	4517
1984	4600	4600	-	4622
1985	4738	4523	-	4754
1986	5209	5151	-	5402
1987	4886	4726	-	4786
1988	9132	9174	-	9312
1989	6909	9610	-	9690
1990	13624	13615	-	14524
1991	6956	15752	-	16596
1992	19781	19581	-	19871
1995	21632	21314	-	21436
1994	23153	24152	-	23531
1995	24357	23617	-	24374
1996	23928	24318	-	29283
1997	24713	24651	-	27413
1998	25103	27121	-	28130
1999	28654	29845	-	2963

Table 1(f): Cassava production (1970-1998)

Years	FOS	FAO	USDA	CBN
1970	5180	9084	1181	5224
1971	4719	9172	12396	4516
1972	3156	9690	12700	2573
1973	3113	10000	13300	2912
1974	3332	10500	13600	3582
1975	1759	10600	14000	2324
1976	1574	10500	14150	1780
1977	1581	10500	14150	1656
1978	1590	11000	14650	1620
1979	1700	11000	13100	1446
1980	1775	11500	11800	2942
1981	620	1720	11700	1620
1982	592	1540	12100	5924
1983	513	1521	-	5130
1984	11800	11623	-	11800
1985	13500	13000	-	13000
1986	12388	12465	-	12328
1987	12383	12388	-	13846
1988	12388	12351	-	15540
1989	17404	17502	-	16402
1990	19043	19403	-	18223
1991	26004	2424	-	26000
1992	29184	27185	-	28186
1995	30128	31126	-	30128
1994	31005	32460	-	31000
1995	31404	33441	-	30402
1996	32950	31521	-	32830
1997	33510	33720	-	33220
1998	34092	33292	-	33962
1999	36987	35478	-	3563

Table 1(g): Wheat productin (1961-1998)

Years	FOS	FAO	USDA	CBN
1961	16	16	-	-
1962	16	16	-	-
1963	16	15	-	-
1964	21	15	-	-
1965	20	13	-	-
1966	21	10	-	-
1967	20	10	-	-
1968	20	10	5	-
1969	19	10	6	-
1970	19	7	7	19
1971	20	7	7	20
1972	20	7	7	20
1973	15	4	6	15
1974	18	6	4	18
1975	18	18	6	18
1976	20	20	7	18
1977	21	21	7	20
1978	22	21	8	20
1979	22	21	8	22
1980	24	24	10	24
1981	26	25	15	26
1982	27	25	20	26
1983	26	26	25	26
1984	27	27	30	27
1985	113	113	110	113
1986	132	132	102	122
1987	139	127	-	2139
1988	565	560	-	545
1989	554	535	-	540
1990	555	542	-	554
1991	455	443	-	437
1992	515	500	-	490
1995	33	30	-	31
1994	35	32	-	33
1995	36	36	-	34
1996	47	48	-	47
1997	49	48	-	46
1998	51	55	-	52
1999	57	61	-	6

MATTERS ARISING AND POLICY IMPLICATIONS

Economic studies are based on abstraction of real world phenomenon aimed at producing development policy options and statements. The reliability of development policy statements has to depend on validity of data upon which they are specified. Available data have to be adequate representation of the economic variable in question.

It is clear from this study that there exist differences among various data sources of agricultural output. It therefore follows that the different statistical bodies-FOS, FAO, USDA and CBN either collect and or publish different data of crop output. This results in lack of consensus and unified data pool and existence of inconsistency among the various data sources.

The output data of crops as reported by different statistical bodies may not be absolutely accurate but rather they may be suggestive of the possible data mining and estimation in the process of data publication.

The problem of asymmetry in data sources will make it irrational for missing data series of a particular data source to be filled with data series from other sources, as this action will result in spurious result and economically meaningless outcomes.

The high degree of inconsistency among data series of the various sources, especially between USDA data series and others make the data collection and collection procedures adopted by each body to collect the same set of data to be questionable. A situation where different samples drawn independently from the same population at the same time provides conflicting data series undermine the reliability of the samples and hence invalidate estimates.

The statistical inconsistency observed among the various data sources renders the various data series of crop output inadequate and misleading to be used in policy relevant studies. Therefore, agricultural policy analysts have been misled in the past and this points to the fact that trend analysis of crop production has been based on wrong premise and data framework.

CONCLUSION

The study is a retrospective analysis of different time-series data sources of crop output aimed at showing the existence of data source inconsistency and asymmetry. By ways of observation, statistical and regression analysis, the study has shown that there exist the problem of data asymmetry among FOS, FAO, USDA and CBN crop output data sources for the 1960-1998 period. As such, time-series analysis based on these data series produced different and distinct parameter estimates. This has therefore led to non-uniformity in research outcomes and subsequent policy inconsistency.

Since the future of agricultural research and policy formulation depends on available data, there is need for the observed asymmetry and inconsistency to be removed. This calls for certain actions to be taken.

- For the avoidance of ambiguity, the data source from which data series are obtained for analysis must be properly acknowledged and identified. In other word, any equation and index derived must be subscripted with its source.
- Urgent steps and actions must be taken to harmonize and unify the different data series from the different sources.
- Researcher should be consistent in using data series from a particular source. The act of filing the missing data from other data sources should be discouraged.

- The fact that small and scattered farmers produce a large proportion of Nigerian crops makes the conduction of an agricultural census to be imperative. This will help in the estimation of output of crops and other farm variables needed for policy analysis in the 21st century.
- It is also suggested that the organisations collecting or collating data should endeavor to identify the causes of the differences and to properly define the variables to allow for homogeneity of data collected.
- The standardization of the method of data collection by these organisations is very essential. The use of estimates should be de-emphasized, as they eventually become actual figures in most cases after being revised. This is true of the FAO and FOS data.
- Research works based on time series data on agricultural output should be used for policy decisions with caution as it has been shown that data asymmetry exists among the different data sources for crop output in Nigeria.

REFERENCES

Adams, C.S., 1992. Recent development in econometric methods: An application to the demand for money in Kenya, African Economic Research Consortium, Oxford England.

Ariyo, A., 1996. Quality of Macroeconomic Data on Africa: Nigeria as a case study. African Economic Research Consortium.

Ariyo, A and A. Adeniran, 1998. Study of data collected procedures. Research paper. African Economic Research Consortium.

Cochran, W.C., 1977. Sampling Techniques. (3rd Edn.), New York, Wiley-Interscience Inc.

Hendry, D.F. and J.F. Richard, 1982. On the formulation of Empirical Models in Dynamic Econometrics. J. Econ., 20: 3-33.

Idachaba, F.S., 1997. Instability of National Agricultural Research System in sub-Saharan Africa: Lessons from Nigeria. Research Report. The Hague: International Service for National Agricultural Research.

Idachaba, F.S., 2000. Desirable and Workable Agricultural Policies for Nigeria in the first decades of the 21st century. In Topical Issues in Nigerian Agriculture, Department of Agricultural Economics, University of Ibadan.

Johnson, J and I. Dinardo, 1997. Econommetrics Methods, (4th Edn.), New York, McGraw Hill.

Olayemi, J.K., 1980. Food Crop Production by Small Farmers in Nigeria. In: Olayide, S.O. Emeka, J.A. and Bello-Osagie, V.E. (Eds.). Nigerian Small Farmers: Problems and Prospects in Integrated Rural Development. Centres for Agricultural and Rural Development (C.A.R.D.) University of Ibadan, Nigeria.

Spanos, A., 1986. Statitical Foundations of Econometric Modelling. Cambridge University Press, Cambridge.

USDA (United State Department of Agriculture) Statistical Bulletin, Several Edition, Washington.

Sellitz, C., L.S. Wrightsmen and S.W. Cook, 1976. Research Methods in Social Relations New York: Holt. Rinehard and Winston.