Relationship between Inflation and Stock Market Returns: Evidence from Nigeria

Douglason G. Omotor[1]

The linkage between stock prices and inflation has been subjected to extensive research in the past decades and has arouse the interests of academics, researchers, practitioners and policy makers globally, particularly since the 1990s. The issue has been the apparent anomaly of the negative relationship between inflation and stock market returns as most studies in the industrialized economies have shown. This paper investigates this relationship using monthly and quarterly data of Nigeria for the period 1985 to 2008. The findings of this paper seem to suggest that stock market returns may provide an effective hedge against inflation in Nigeria.

Keywords: Inflation, stock market, Fisher effect, Fama's proxy hypothesis, Nigeria.

JEL Classification: E31, G11

1. Introduction

The last two decades have been a tranquil for the Nigerian economy. Inflation rate for example, rose markedly in the fourth quarter of 2008 reaching a 3-year high of 15.1 per cent in December from its single digit level of 7.8 per cent at end of March, 2008. Precisely, the inflation rate was 6.5 per cent in December 2007. The inflationary pressure which continued into 2009 as some sources have it (notably the Central Bank of Nigeria, 2009), may have been attributed to rising food prices, inefficient and poor transport services, port congestion, depreciation of the naira and the rush to spend budgetary allocations by government agencies before fiscal year end (Sampson, 2009). During the same periods, the Nigerian capital market experienced a bullish trend when it started the year 2008 at 58,580 (with a market capitalization of N10.284 trillion), and went on to achieve its highest value ever of 66,371 on March 5, 2008,with a market capitalization of about N12.640 trillion (Aluko, 2008). The capital market has since the March 5 to October, 2008 lost about N3.38 trillion, over 26.7 percent; as market capitalization stood at N9.11 trillion. Nigeria equally faced a major decline in portfolio equity flows perceived to be correlated with the sharp fall in stock market. For instance, foreign portfolio investors withdrew $15 billion from the Nigerian capital market in January 2009 (Ajakaiye and Fakiyesi, 2009). The All Share Index (ASI) consequently shred a total share of 67 per cent from March 2008 to March 2009.

In attempt to find some reprieves for the continuous bearish trend in the market, the Central Bank of Nigeria took over the management team of 8 commercial banks effective from August 14, 2009 as the illiquidity in the capital market dove-tailed into the money market. The action described as a hybrid attempt to restructure these banks as a result of their debt exposure to the capital market is beginning to have its toll on the average general price level as analysts speculate precautionary cash balances. One puzzle left to be answered is if the sharp movements in general prices (inflationary) during these years have any linkage with the bullish/bearish capital market dominated activities before and after the 2008 crash.

This paper investigates the relationship between inflation and stock market returns using Nigerian data. Specifically, we effect the analysis by exploring the distinct impacts of inflation on the stock market returns at different time horizons, and also test the Fisher hypothesis by examining the relationship between, (b) contemporaneous inflation and stock market returns, and (c) between inflation and money on the one hand, and between inflation and real activity on the other. The outcomes of the analyses are expected to be of immense importance to investors particularly, in reaching rational decisions on asset allocation and advancement of the literature on financial economies.

The rest of the paper is organized as follows: the next sections briefly review some related literature and presents the historical perspective and performance of Nigerian capital market. Section 4 presents the model, data sources and measurements. Section 5 discusses the results. Section 6 explains the role money and economic activity played in the inflationary process, while the last section concludes the paper.

[1] Department of Economics, Delta State University, Abraka, Nigeria. yomotor@yahoo.com; +2348034083112

2. Review of Some Related Literature

The linkage between stock market returns and inflation if any has drawn the attention of researchers and practitioners alike particularly since the twentieth century. The foundation of the discourse is the Fisher (1930) equity stocks proclamation. According to the generalized Fisher (1930) hypothesis, equity stocks represent claims against real assets of a business; and as such, may serve as a hedge against inflation. If this holds, then investors could sell their financial assets in exchange for real assets when expected inflation is pronounced. In such a situation, stock prices in nominal terms should fully reflect expected inflation and the relationship between these two variables should be positively correlated *ex ante* (Ioannides, et.al., 2005:910). This argument of stock market serving as a hedge against inflation may also imply that investors are fully compensated for the rise in the general price level through corresponding increases in nominal stock market returns and thus, the real returns remain unaltered.

Further extension of the hedge hypothesis posits that since equities are claims as current and future earnings, then it is expected that in the long run as well, the stock market should equally serves as a hedge against inflation. Fama (1981) however, put up a proxy hypothesis when he argued the relationship between high rates of inflation and future real economic growth rates as negative. Views that rationalize the negative co-movements between inflation rates and real stocks returns however differ.

The inflation illusion hypothesis of Modigliani and Cohn (1970) point's out, that the real effect of inflation is caused by money illusion. According to Bekaert and Engstrom (2007:1), inflation illusion suggest that when expected inflation rises, bond yields duly increase, but because equity investors incorrectly discount real cash flows using nominal rates, the increase in nominal yields leads to equity under-pricing and vice versa.

Feldstein's (1980) variant of the inflation and stock market returns theoretical nexus, suggests that inflation erodes real stock returns due to imbalance tax treatment of inventory and depreciation resulting to a fall in real after-tax profit. Feldstein further observed that the failure of share prices to rise during substantial inflation was because of the nominal capital gains from tax laws particularly, historic depreciation cost (Friend and Hasbrouck, 1981). In Fama's (1981) hypothesis, which is based on money demand theory; correlation between inflation and stock market returns is not a causal one; rather, it is a spurious relationship of dual effect. Yeh and Chi (2009:168) in explaining the Fama's hypothesis observed that the reason for the revised correlation is because when inflation is negatively related to real economic activity, and there is a positive association between real activity and stock returns, the negative relationship and stock returns holds. This flow of relationship according to them is not direct.

Hoguet (2008), explanation of stock-inflation neutrality is anchored on two stances as outlined from Giammarino (1999); 1) that companies can pass on one-for-one costs; and 2) that the real interest rate which investors use to discount real cash flows does not rise when inflation rises and in addition, inflation has no long-term negative impact on growth.

The appropriate direction of the relationship or the neutrality between inflation and stock market returns relationship have equally generated a large body of evidence in the empirical literature. Earlier studies by Bodie (1976), Nelson (1976), and Fama and Schwert (1977) were aroused by the rising inflation of the 1970s in the US. According to Alagidede and Panagiotidis (2006), these studies compared the inflation hedge properties of common stocks with those of other financial and real variables for the US. They found that common stock acted as poor hedge against unexpected and expected inflation. In another development, Firth (1979) and Gultekin (1983) found reverse evidence using UK data. Jaffe and Mandelker (1976) also report a negative relation between annual stock returns and concurrent rates of inflation over short sample periods but a positive relation over the much longer period 1875-1970. In another vein, Marshall (1992) argued that the negative relationship between stock returns and inflation will be less pronounced during periods when inflation is generated primarily by monetary fluctuations. Studies that have agreed with this proposition are Graham (1996), who found a positive relationship between common stocks and inflation in the USA (1976-1982) during the period money rather than real activity was the cause of the inflation. Spyrou (2004) study of ten emerging economies further provide evidences that may suggest equity providing an effective hedge against inflation and that the inflation could be explained by a significant relationship between money and consumer prices in the emerging markets.

Rapach (2002) employed data of 16 OECD countries to determine the direction of the correlates. He observed that long-run inflation neutrality exists in the stock markets of the countries. Following the methodology of King and Watson (1997) in the establishment of time series properties, Rapach explained that the long-run Fisher effects exists if the long-run real stock returns do not respond to a permanent inflation shock (Yeh and Chi, 2009: 169).
Studies on the inflation-stock return maxim for the Nigerian economy as the scan on the literature revealed are however relatively sparse. The available few from our search equally have their limitations. Subair and Salihu (n.d.)using an error correction model to investigate the effects of exchange rate volatility on the Nigerian stock market though found exchange rate volatility to exert strong negative impact on the Nigerian stock market, the rate of inflation did not have any long run relationship with stock market capitalization. The reason for no long run relationship as adduced by the authors is the overbearing participation of the government in the market. First, the cointegration result which authors claimed to underscore this reason was not reported. Second, which market (stock exchange or foreign exchange) government participation is overbearing is not explicitly defined. However, in either of the two markets, government participation over the years has been eroded. Consequently, Subair and Salihu findings may be misleading.

Daferighe and Aje (2009) using annual data analyzed the impact of real gross domestic product, inflation and interest rates on stock prices of quoted companies in Nigeria from 1997-2006. The results among others showed that low inflation rate resulted in increased stock prices of quoted firms in Nigeria. Daferighe and Aje study suffers from misspecification drawbacks and spurious relationship. A high R^{-2} with suspected highly autocorrelated residuals signify that the conventional significant tests are biased. The integrated process of the variables was not analyzed, neither are the individual test of the series for random walks checked. The short data span of only ten points using a multiple regression technique is inappropriate.

Table 1. Summary of Some Previous Studies

Author(s)	Sign
Kessel (1956)	Positive
Nelson (1976)	Negative
Jaffe and Mandelker (1977)	Negative
Fama and Schwert	Negative
Firth (1979)	Positive
Fama	Negative
Gultekin (1983)	Negative
Pearce and Ripley (1988)	Neutral
Lee (1992)	Neutral
Amidhud (1996)	Neutral
Samarakoon (1996)	Positive
Anari and Kolari (2001)	Neutral
Crosby (2002)	Neutral
Spyrou (2001)	Negative
Mark	Neutral
Ioannides et.al. (2004)	Positive
Akmal (2007)	Positive
Yeh and Chi (2009)	Negative
Baekaert and Engstrom (2009)	Positive

Source: Author’s compilation

Yaya and Shittu (2010) examined the predictive power of inflation and exchange rate on Nigeria’s stock volatility. The QGARCH model shows a significant relationship of inflation and exchange rate to conditional stock market volatility. This study however did not test whether equities are a good hedge against inflation. This further creates the impetus for our study which sets out to determine direction of relationship between equities and inflation on the one hand; and if stock market returns provided an effective hedge against inflation in Nigeria on the second hand. However, some other previous studies (not on Nigeria) which attempt to empirically establish the direction of relationships between inflation and stock returns are summarized in Table 1. The progeny is however still inconclusive as the puzzle rears.

3. The Historical Perspective and Performance of Common Stocks

The historical monthly behavior of the nominal (and real) stock prices along with the general price index for the periods of 1985(1)-2008(12) are presented in Figures 1 to 4. The two series as shown in Figure 3 look related to each other. In real terms, economic units experienced the highest spike in 2005. This may not be unconnected with the new political order of return to democratic governance in the country in late 2003. The stock index which stood at a value of 58,579.77 on 2nd January 2008 with market capitalization of N10.284 trillion), attained its peak value of 66,371.2 on 5th March 2008 (market capitalization, N12.640 trillion). Since this unprecedented height, the stock index has been exhibiting secular bearish gyration. The index declined to 50,393.88 on 23rd July, 2008 along a capitalization of N10.091 trillion with a continued decline to 33,754.11 and a market capitalization of N7.405 trillion. A noticeable rise to 38,018.44 (market capitalization of N8.390 trillion) was experienced on 17th November, 2008. The rise was short-lived as the market weaned to 28,028.01 with capitalization record of N6.213 trillion and further decline to 20, 827.17 on 31 December, 2009 (market capitalization was N4.989 trillion). Since the start of the bearish market, the lowest threshold of 20,618.71 and a market value of N4.904 were recorded on 14 December 2009. Also from March 5, 2008 to 14 December 2009, the capital market lost well over N7.736 trillion, or about 61.2 per cent.

4. Model Specification, Data Sources and Measurements

4.1. Model specification

In this study, we apply a simple model in the estimation of the relationship between stock returns and inflation, following the lead by Spyrou (2001) as:

$$STK_t = \lambda_1 + \lambda_2 CPI_t + \varepsilon_t \qquad (1)$$

where, STK_t is return on the stock portfolio for Nigeria and CPI_t is the rate of inflation. λ_1 is a constant, and λ_2 is the slope coefficient that captures the sensitivity of the stock returns to inflation level. ε_t is the stochastic term which assumes the properties $\sim N(0, \delta^2)$. Economic theory as implied in the Fisher effect supports the existence of a linear relationship of the above system. Other studies that have previously estimated this form of linear relation include Jaffe and Mandelker (1976), Choudhry (2001), and Alagidede and Panagiotidis (2006) among others. Graham (1996) has argued elsewhere that although this simple model of analysis does not distinguish between expected and unexpected components of inflation, the resulting quantitative evidence are not different. λ_2 is *a priori* expected to be positive. The reason is because for emerging economies unlike industrialized economies as previous empirical studies have shown, inflation are primarily caused by money rather than real activity and the effect may appear less pronounced (Marshall, 1992). Should the estimated results of Equation (1)[1] follow this pattern, we shall then investigate whether indeed this proposition holds for Nigeria. Such a relation following Spyrou (2004) is functionally stated as:

$$LCPI = f\,(LM2, LRGDP) \qquad (2)$$

The objective here is to examine whether consumer prices (LCPI) are related more to money supply (LM2) and /or economic activity (LRGDP).

Figure 1. Stock Price Index

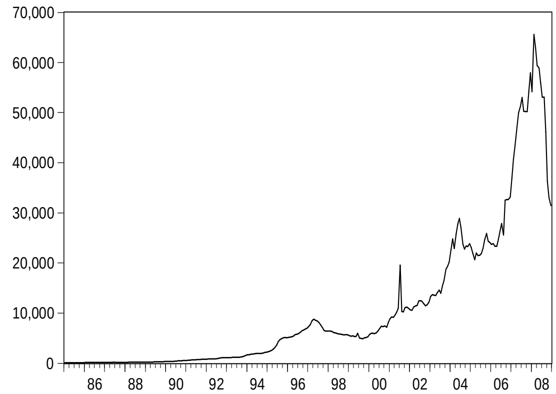


Figure2. Consumer Price Index

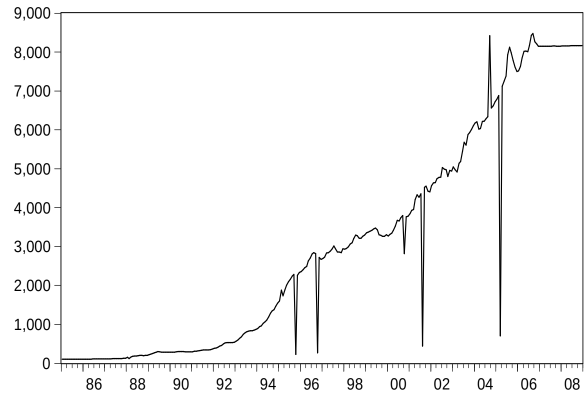


Figure3. Stock Price Index and the Consumer Price Index

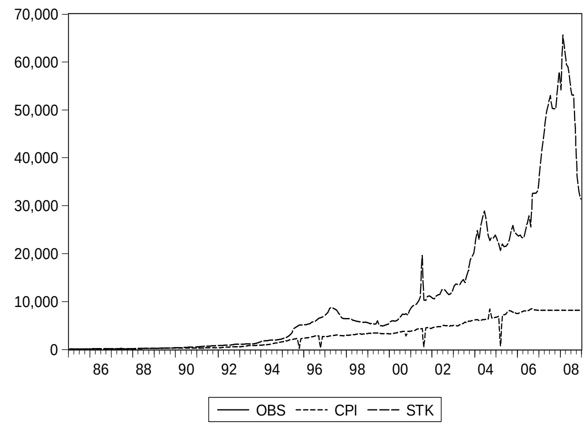
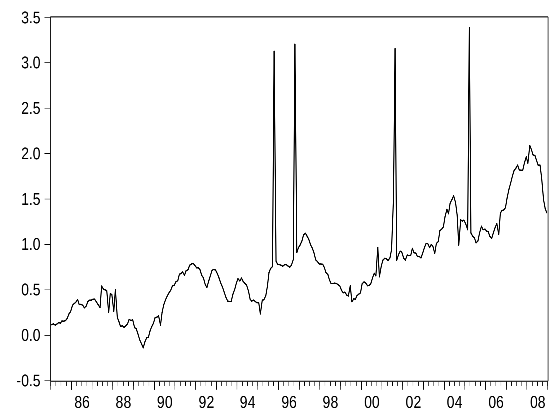


Figure 4. Real Stock Prices



The time series characteristics of the variables in Equations (1 and 2) will be investigated to determine their levels of integration or presence of unit root (stationarity). The level of integration of the variables or the order of autoregressive process (AR1) of the variables is considered by applying the augmented Dickey-Fuller (ADF) tests. The objective is to determine whether the underlying stochastic process that generated the series can be assumed to be invariant with respect to time (Pyndyck and Rubinfied, 1998:493). The ADF is specified when Σ_t is autoregressive to eliminate serial correlation of errors and it takes the form:

$$\Delta Y_t = \alpha + \beta_t + \delta Y_{t-1} + \sum_{j=1}^{p} \lambda_t \, \Delta Y_{t-1} + \varsigma_t \qquad (3)$$

If all the variables are found to be I(1), that is should the ADF unit-root tests show that the variables reject the null for their first differences, then we shall test for the cointegration of the variables. According to Granger representation theories (Granger, 1987), if two variables are non-stationary that is I(1), and the series have cointegrating relationship among them, then the dynamic function can be represented as an error correction mechanism (Engle and Granger, 1987). The error correction mechanism (ECM) according to Qayyum (2005) is popularized by David Handry through a number of studies (Hendry, Pagan and Sargan, 1984; Hendry and Ericsson, 1991; Hendry and Mizon, 1993).

In order to impose the cointegrating vectors on the error correction model, should the variables have cointegration relationships, Equation (1) will be transformed by linearilization and incorporation of a differenced operator (Δ) and lagged error term as in Equation (4):

$$\Delta STK_t = \lambda_o + \lambda_1 \Delta CPI_t + \mathrm{ECM}_{t-1} + \nu_1 \qquad (4)$$

4.2 Data sources and measurements

The data set consists of monthly stock from January 1985 to December 2008. The data were obtained from the Central Bank of Nigeria (CBN) *Statistical Bulletin, 50 Years Special Anniversary Edition* (2009). Data include monthly observations on Stock Price Index (STK) measured as the Nigerian Exchange's All Share Index simply ASI and consumer price index (CPI). The monthly STK is used as a proxy for stock returns (also known as equity returns). The growth rates of the series are defined as the first difference of the logarithmic price levels. For purpose of examining the stability of the estimate, the sample size of the data is split into two equal sub-periods: February 1985-January 1997, and January 1997-December, 2008. Further justification for this choice of exogenous break date is the fact that the Nigerian capital market experienced its first fundamental and unprecedented growth at end of 1996. For instance, turnover value of the exchange changed by 284 per cent to N7.063billion at end of 1996 from N1.83billion at end of 1995; while Foreign Investment Portfolio Transactions ($US million) increased from $1.137million to $32.99million during these respective periods thus setting a new platform of structural shifts of dealings on the market from January 1997 (See Table in Appendix for more information).

We use the log of broad money supply (LM2) measurement to represent money in the analysis. The use of M2 also finds favour in the argument of Hafer and Jansen (1991) and Laidler (1993) that the boundaries of narrow money shift over time to accommodate new financial instruments, thus making it plausible to apply M2 in money supply related analyses. Real activity (LRA) is proxied with the change in log of real gross domestic product Δlog (LRGDP). We use LRGDP as proxy because of unavailability in obtaining quarterly series for industrial production. The data covered the period 1985(1) to 2008(4).

5. Empirical Results

The descriptive statistics on the rates of the stock indices and the correlations with the rates of inflation (rates of change in consumer prices) are presented in Table 2. It can be observed that the sub-period mean values for both stock market returns and inflation were highest during the period 1985(2) -1997(1). This sub-period however recorded the lowest total risk (standard deviation), while the period 1997(1)-2008(12) returned the highest total risk. This is not however surprising as the worst market crash in the history of the Nigerian Capital Market occurred during this period.

Table 2. Descriptive statistics

Panel A. Monthly stock market returns

Sample period	Mean	Standard deviation	Skewness	Kurtosis	Sample size
1997(1)-2008(12)	0.010	0.096	-0.728	24.644	144
1985(1)-2008(12)	0.020	0.076	-0.910	36.872	287
1985(2)-1997(1)	0.029	0.046	0.427	14.420	144

Panel B. Monthly inflation rates

Sample period	Mean	Standard deviation	Skewness	Kurtosis	Sample size
1997(1)-2008(12)	0.008	0.389	0.076	34.817	144
1985(1)-2008(12)	0.015	0.390	-0.093	34760	287
1985(2)-1997(1)	0.023	0.392	-0.259	35.014	144

As regards the correlation coefficients, the results are mixed. The sub-period 1985(2)-1997(1) indicate a negative relationship between the stock market returns and inflation. This sub-period experienced the highest rate of inflation. The correlation coefficients of the full period of analysis [1985(1) 2008(12)] and the sub-period 1997(1) 2008(12) are positive. Activities in the market during the later sub-period may have influenced the behavior of the entire period, given that the growth during this sub-period account for over 93 per cent of the total period growth. Thus, the relatively higher risk during this period may not be amazing. Should we be guided preliminary by the signs and significance of the correlation coefficients, we may well conclude that the relationship between stock returns and inflation is positive. However, this should not be extended to imply causality.

Table 3. Correlation between changes in stock indices (ΔSTK) and inflation (ΔCPI)

Statistic	1997(1)-2008(12)	1985(1)-2008(12)	1985(2)-1997(1)
Correlation coefficient	0.323 (4.065)*, $\rho=0.000$	0.196 (3.365)*, $\rho=0.000$	-0.033 (0.404), $\rho=0.697$
observations	144	287	144

Note: *Denotes significance at 1%. *t*-statistics reported in parentheses. ρ =*probability*

The results from estimation of Equation (1) as presented in Table 4 suggest that for the empirical relationship between stock returns and inflation (for the whole period) is positive and statistically significant. However, the relationship is negative for the first sub-period and statistically not significant except the constant term, which may imply better role for other factors (e.g. Treasury bill rate as hedge) than inflation. Interestingly, the relationship is positive in the second sub-period (1997 1 2008 12) and statistically significant. These results in their various forms are not different from previous ones for emerging economies, but differ substantially from the documented negative relationship in more advanced North American economies (Spyrou, 2004). A possible explanation for the difference in behavior of the Nigerian Stock Market as an emerging one (like other emerging economies, in the 1990s) as it relates to the stock return-inflation nexus from those of the typical industrialized economies, particularly in the late 1990s and early 2000s, as earlier noted could be as a result money rather than real economic activity being the more significant determinant of inflation in Nigeria. The less pronounced effect may have accounted for the low levels of the coefficients of determination. The empirical verification of the money, economic activity and the stock returns-inflation nexus is addressed in section 5.

Table 4. Relationship between stock returns and inflation ($STK_t = \delta_1 + \delta_2 CPI_t$)

Period	δ_1	δ_2	R^2	Durbin-Watson Stat.
1985(2)-1997(1)	0.029 (7.521)*	-0.004 (0.404)	0.005	2.001
1997(1)-2008(12)	0.010 (1.292)	0.080 (4.065)*	0.098	2.060
1985(2)-2008(12)	0.019 (4.330)*	0.038 (3.365)*	0.035	2.010

Notes *Denotes significance at 1%. *t*- statistics reported in parentheses

Table 5. Augmented Dickey-Fuller (ADF) Unit Root Tests.

Variables	Intercept	Intercept and Trend	Variables	Intercept	Intercept and Trend
CPI	-2.09	-2.08	ΔCPI	-14.02	-9.89
STK	-1.42	-0.97	ΔSTK	-18.68	-18.72

Critical Values: 1% = -3.99; 5% = -3.43 and 10% = -3.14.

Next, we address the concern of one of the anonymous referees on the need to establish the integration order of each variable in Equation (1) even though they may be stationary given that they are rates of change. This query is underscored by the fact that well into the 1980s; empirical researches in macroeconomics were based on the assumption that the variables in such models are stationary. Problem of validity arises if statistical inferences associated with presumably stationary processes are indeed nonstationary. Clive Granger (1981) is credited with this change of realization and equally contributed enormously to the testing hypotheses of stationarity and other time series properties. In this regard, we conduct the tests for stationarity and detection of cointegration of the series in Equation (1).

5.1. Long-run relationship between inflation and stock market returns.

Stationarity test and cointegration

To test whether the two time series are nonstationary, the ADF unit root test is employed. Table 5 presents the results of these tests for levels as well as first differences of the variables. The null hypothesis is that the series are non-stationary (that is, presence of a unit root), and the alternative hypothesis is that they are stationary (that is, absence of a unit root). The test statistics suggest the levels of the series are not stationary but the first differences of the series are stationary; thus accepting the null hypothesis of an I(1) process. This implies that the series are integrated of order one and can be tested for cointegration in the Johansen sense.

The results of the Johansen (1991, 1995) system of maximum likelihood approach to cointegration analysis are presented in Table 6. The Johansen's trace test aimed at determining whether a long-term relation exists between the two series starts with the null hypothesis that there is no cointegrating relation, and if this hypothesis cannot be accepted, we test the hypothesis that there is at most one cointegrating equation. Since there are only two variables in the model, we test whether the number of cointegrating equations is zero, one, or two (Anari and Kolari, 2001). The results of the trace test suggest the existence of one cointegrating equation (or long-run relation). The maximum eigenvalue test equally report the existence of one cointegrating equation between stock returns and inflation. Cointegration also implies that causality exists between the variables in at least one direction but does not indicate the direction of causal relationship (Erdal, et.al. 2008).

Table 6. Trace and ʎ–max Test Statistics

Null, Ho	Alternative	Trace	5% Cri. Val.	Null, Ho	Alternative	ʎ–max	5% Cri. Val.
$r = 0$	$r \geq 1$	21.24	15.49	$r = 0$	$r = 1$	18.52	14.26
$r \leq 1$	$r \geq 2$	2.72	3.84	$r \leq 1$	$r = 2$	2.72	3.84

Trace test and Max-eigenvalue test all indicate 1 cointegrating equation at the 0.05 level.

Test of causality in the spirit of Granger when series co integrates, tantamount to estimating an error correction model (ECM) of Equation (1) using first differences of the variables. The results of the long-run relations based on the Least Squares technique are reported in Table 7 while the Granger causality results are presented thereafter In Table 8 respectively.

As shown in Table 7, the estimated Fisher coefficient (δ_2) is positive, very low (less than 1) but statistically significant at 1 percent and not serially correlated. The low level of the Fisher coefficient provides a conservative estimate of how inflation in the long-run affects stock market returns in a typically emerging market. Low levels of Fisher coefficient have also been reported by other researchers for emerging markets (some include Nigeria) using different periods, empirical methods and data series (see Alagidede and Panagiotidis, 2006; Spyrou, 2010 etc.). It may be argued as well that the low Fisher coefficient may result from failure of the market to include information contained in inflation and thus likely to offer only a partial hedge to investors against rising inflation.

We also report the estimate of the speed-of-adjustment coefficient (ϑ_3) in Table 7. The value of -0.005 means that stock market returns takes a longer time to return to their long-run equilibrium following movements in the Nigerian goods market. This finding of long time for inflation to be fully reflected in stock market returns is consistent with Fisher. As noted by Anari and Kolari (2001:598), the long time effect influenced Fisher (1930) invents of distributed lag models and consequent analysis of interest and inflation rate series for the United Kingdom and United States. The outcome is that interest rates follow price changes with long distributed lags of about fifteen to thirty years.

Table7. Long-Run Relation between Inflation and Stock Market Returns

$$\left(\Delta \mathrm{STK}_t = \vartheta_1 + \vartheta_2 \Delta \mathrm{CPI}_t + \vartheta_3 \mathrm{ECM}_{t-1}\right)$$

Period	ϑ_1	ϑ_2	ϑ_3	Prob.	R^2	Durbin-Watson Stat.
1985(1)-2008(12)	0.019 (4.372)	0.067 (5.341)	-0.005 (-4.672)	0.000	0.107	2.074

5.2. Granger causality

As earlier noted, cointegration imply that causality exists between the variables in at least one direction but does not indicate the direction of causal relationship. To avoid miss-specifying the model, we include the one period lagged error correction term following Chontanawat *et. al*. (2006) in the estimating the Granger causality test. The empirical result as presented in Table 8 suggests long-run uni-directional causal relationship from inflation to stock market returns and not the other way round.

Table 8. Pairwise Granger causality tests

Null Hypothesis	Obs.	F-Statistic	Probability
ΔSTK does not Granger cause ΔCPI ΔCPI does not Granger cause ΔSTK	286	29.765 2.862	2.E-12 0.053

5.3. Stability analyses

Stability tests were conducted over the sample period by applying both the Chow breakpoint and Quandt-Andrews tests. The results could not reject the null hypothesis of no breaks at specified breakpoints. Thus, there were periods at which significant drift in the relationship between inflation and stock market returns blipped. This is more noticeable in 1995 and after 2006. The results of the Chow breakpoint are contained in Table 9.

Table 9. Chow breakpoint Tests

F-Statistic	6.367	Prob. F(3,281) 0.0003
Log likelihood ratio	18.875	Prob. Chi-square 0.0003
Wald statistic	19.102	Prob. Chi-square 0.0003

6. Money, Economic Activity and the Stock Returns-Inflation Nexus

This section explains the role money and economic activity could have played in inducing the positive relation between stock returns and inflation nexus of Nigeria particularly in the late 1990s up till 2008; as eulogized by Marshall (1992) among others. To achieve this, the validity of the long term equilibrium among the variables (consumer prices, money and real activity) is examined using the variant of the Johansen technique detailed in Johansen and Juselius (1990).

We begin by first considering whether each series is integrated (the order of difference before stationarity is achieved) of the same order. To do this, the standard Augmented Dickey-Fuller (ADF) test and Schwarz Information Criterion (SIC) as indicator for lag selection were first determined. The ADF results are presented in Table 10 and in no case can the hypothesis that the series contain a unit be rejected. The first differences are, however, stationary and thus the series are I(1) and candidates for cointegration. All the ADF regressions include an intercept and trend, while the asymptotic critical values are from MacKinnon (1999) provided by the econometric software (EViews version 7).

The cointegration rank is then conducted with the maximum eigenvalue and trace test. The model lag selection is based on the lowest SIC. We however discussed only the results of the cointegration tests for the sample periods 1997(1)-2008(4) in Table 11 for two reasons; first, for purpose of space and second but most important, the robustness of its results when we analyzed the long-run relationships of the three sample periods. Consequent upon this, further analyses and discussions shall be restricted to this sub-period.

Table 10. ADF Unit Root Test on Variables

Series	Levels	First Difference	Order of Integration
LCPI	-2.051	-12.542*	I(1)
LM2	0.943	-10.689*	I(1)
LRGDP	1.725	-9.306*	I(1)

Note: * denotes significance at 1% critical values from MacKinnon (1999)

The trace statistic and the maximum eigenvalue statistic suggest one cointegration vector at 5 percent significance level for the three sample periods. Given the evidence of one cointegrating vector among the three variables ($r = 1$), we normalize the cointegrating vector on the natural log of consumer prices (LCPI) of Table 11. This also means that the hypothesis that $r = 0$ is rejected against the rule $r = 1$, but the hypothesis that $r = 1$ cannot be rejected against r = 2, and so on. The implication is that there is a long-run relationship between consumer prices (inflation index), money and real activity in the Nigerian capital market.

Table 11. Cointegration Results

Panel A: Trace Statistic

Period	Null	Alternative	Statistic	0.05 Critical Values	Prob. *
1997 1 2008 4	$r = 0$	$r \geq 1$	80.16687	29.79707	0.0000
	$r \leq 1$	$r \geq 2$	6.427749	15.49471	0.6450
	$r \leq 2$	$r = 3$	0.845752	3.841466	0.3578

Note: Trace test indicates 1 cointegrating equation at the 0.05 level.
*MacKinnon-Haug-Michelis (1999) p-values

Panel B. Max-Eigen Statistic

Period	Null	Alternative	Statistic	0.05 Critical Values	Prob. *
1997 1 2008 4	$r = 0$	$r \geq 1$	73.73912	29.79707	0.0000
	$r \leq 1$	$r \geq 2$	5.581997	15.49471	0.6450
	$r \leq 2$	$r = 3$	0.845752	3.841466	0.3578

Note: Max-eigenvalue test indicates 1 cointegrating equation at the 0.05 level.
*MacKinnon-Haug-Michelis (1999) p-values

We extracted the estimates of the normalized cointegrating coefficients below the cointegration vector with their standard errors reported in parentheses. The normalized cointegrating results reported on Table 12 connote that money supply is positively related to consumer price index, while the sign of the economic activity indicates a negative relationship with the consumer price index. The coefficients are equally statistically significant at 5 percent level. The implication of this shall be discussed alongside the tests of restrictions.

The sequential tests of restrictions were carried out by imposing over-identifying restriction on a re-estimated cointegrating relation; first, that money supply equals zero [LM2 = 0], and then the measure of economic activity [LRGDP = 0]. The results reported in Table 12 show that the restrictions are rejected, that is, the log-likelihood ratio statistic for testing the restriction based on the probability values for both restrictions are statistically significant. Thus the restrictions are rejected at the 5 percent level of significance. This further confirms the normalized cointegrating results that consumer prices are related to both money and real activity and that the money also matters in the determination of inflation in Nigeria. The implication of the results is that though money matter in the relation, equities are a good hedge against inflation in Nigeria (during the period of review) as economic theory suggest. This may not be unconnected to the bullish market trend during a significant part of the period of analysis before the great global financial crash in the later part of the period under review which eventually had a backlash effect.

Table 12. Normalized Cointegrating Vector and LR Test of Restrictions (1997 1 2008 4)

Normalized cointegrating vector (standard error in parentheses)

$$LCPI = \underset{(0.058)}{0.551}\ LM2 \quad \underset{(1.911)}{-22.827}\ LRGDP$$

LR test of restrictions [probability]

$$LM2 = 0 \quad \chi^2(1) = 5.288\ [0.021]$$
$$LRGDP = 0 \quad \chi^2(1) = 67.455\ [0.000]$$

7. Conclusions.

Fama's 'proxy hypothesis' explains the apparent anomaly of the negative relationship between inflation and stock market returns as against economic theory suggestion that equities are a good hedge against inflation. The focal objective of the paper is to investigate this relationship using monthly and quarterly data of Nigeria for the period 1985 to 2008. The findings of this paper seem to suggest that stock market returns may provide an effective hedge against inflation in Nigeria. This is explained by the significant and positive relationship between inflation and stock prices as the Fisher (1930) hypothesis postulates. This also implies that investors in making good portfolio decisions should perhaps view equities as long-term holdings against inflation's erosion of purchasing power. This is with caution as recent developments in the Nigerian capital market may have suggested that equities may not necessarily be the best performing asset class over the short term.

Another implied finding of the paper interesting to be mentioned is that the monetary and real sectors of the economy may not be independent of each other, as money may also matter in explaining the behaviour of inflationary process in Nigeria. Thus policies geared at controlling inflation should take into cognizance the role of monetary and real variables especially as these will go a long way in further deepening of the stock market.

Endnote

1. The results of Equation (1) from a simple Ordinary Least Squares (OLS) technique are presented in Table 4.

References

Ajakaiye, O. and T. Fakiyesi (2009). Global Financial Crisis. *Discussion Series Paper 8*: Nigeria. London: Overseas Development Institute.

Akmal, M.S. (2007) Stock returns and inflation: An ARDL econometric investigation utilizing Pakistani data *Pakistan Economic and Social Review* Volume 45(1): 89-105.

Alagidede, P. and T. Panagiotidis (2006). Can Common Stocks Provide a Hedge a against Inflation? Evidence from African Countries. The Rimini Center for Economic Analysis, Italy,WP 10-06.

Aluko, E. M. (2008). The Global Financial Meltdown: Impact on Nigeria's Capital Market and Foreign Reserves. http://www.gamji.com/aluko/aluko209.htm[downloaded [5-6-2010]

Amihud, Y. (1996), Unexpected inflation and stock returns revised – Evidence from Israel. *Journal of Money Credit and Banking*, Volume 28, pp. 22-33.

Anari, A. and J. Kolari (2001). Stock Prices and Inflation. *Journal of Financial Research*. 24(4): 587-602.

Baekaert, G. and E.C. Engstrom (2009), Inflation and the Stock Market: Understanding the 'Fed Model'(June 2009). NBER Working Paper No. w15024. Available at SSRN: http://ssrn.com/abstract=1413588

Bodie, Zvi. (1976). Common Stocks as a Hedge against Inflation. *Journal of Finance* 31(2).

Boudhouch, J. and M. Richarson (1993), Stock returns and inflation: A long-horizon prospective. *American Economic Review*, pp.1346-1355.

Chontanawat, J., Hunt, L. C., & Pierse, R. (2006). Causality between Energy Consumption and GDP: Evidence from 30 OECD and 78 non-OECD Countries. *Surrey Energy Economics Discussion Papers, SEEDS No. 113*. Surrey Energy Economics Centre (SEEC), University of Surrey, www.seec.surrey.ac.uk/Research/SEEDS.htm

Choudhry, T. (2001) “Inflation and Rates of Return on Stocks: Evidence from High Inflation Countries,” *Journal of International Financial Markets, Institutions and Money,* 11: 75-96.

Crosby, M. (2002), Stock Returns and Inflation. Australian Economic Papers. 40(2): 156-165.

Daferighe, E.E. and S.O. Aje (2009). The Impact Analysis of Real Gross Domestic Product, Inflation and Interest Rates on Stock Prices of Quoted Companies in Nigeria. International Research Journal of Finance and Economics. Issue 25.

Engel, R. F. and C. W. Granger (1987). “Co-integration and error correction: Representation, estimation, and testing.” *Econometrica*. 55, 251–76.

Erdal, G., H. Erdal and K. Esengun (2008). The Causality between Energy Consumption and Economic Growth in Turkey. *Energy Policy*, 36: 3838-3842

Fama E. F. 1981.Stock returns real activity, inflation, and money. *American Economic Review*, 71: 545-565.

Fama E. F. and Schwert G. W. (1977) Asset returns and inflation. *Journal of Financial Economics*, 5(2): 115-46.

Feldstein, M. (1980) Inflation and the Stock Market. *American Economic Review*, 70(5): 839-947.

Firth, M. (1979). The relationship between stock market returns and rates of inflation. *Journal of Finance, 34*, 743–749.

Fisher I. (1930). *The Theory of Interest*. Macmillan. New York

Friend, I. and J. Hasbrouck (1980). *Effect of inflation on the Profitability and Valuation of U.S. Corporations*. University of Pennsylvania.

Giammarino, Ronald in Cohn, Richard A. and Lessard, Donald R. (1980). The Effect of Inflation on Stock Prices: International Evidence. *The Journal of Finance.* 36(2).

Graham, F.C. (1996). Inflation, real stock returns and monetary policy. *Applied Financial Economies*, 6: 29-35.

Granger, C. W. J. (1981). Some Properties of Time Series Data and their use in Econometric Model Specification. *Journal of Econometrics*. 16: 121—130.

Gultekin N. B. (1983) Stock market returns and inflation: evidence from other countries. *Journal of Finance*, 38: 49-65.

Hafer, R.W. and D.W. Jansen (1991). The Demand for Money in the United States: Evidence from Cointegration Tests. *Journal of Money, Credit and Banking*. 23:155-168.

Hendry, D. F., A. R. Pagan, and J. D. Sargan (1984) Dynamic Specification. In Z. Girliches and M. D. Interligator (ed.) Handbook of Econometrics. Vol. II. Elsevier.

Hendry, D. F., and G. E. Mizon (1993) Evaluating Dynamic Econometric Models by Encompassing the VAR. In P. C. B. Phillips (ed.) Models, Methods and Applications of Econometrics. Oxford: Basil Blackwell.

Hendry, D. F., and N. R. Ericsson (1991) Modeling the Demand for Narrow Money in the United Kingdom and the United States. European Economic Review 35, 833–886.

Hoguet, G.R. (2008) *Inflation and Stock Prices*. State Street Global Advisor. http://www.ssga.com/library/esps/Inflation_and_Stock_Prices_George_Hoguet_8.21.08rev3CC RI1221060800.pdf. [download 26-05-2010].

Ioannides, D., C. Katrakilidis and A. Lake (2005). The relationship between Stock Market Returns and Inflation: An econometric investigation using Greek data, http://conferences.telecom bretagne.eu/asmda2005/IMG/pdf/proceedings/910.pdf [download04/02/2010].

Jaffe, J. F., and G. Mandelker (1976). "The Fisher effect" for risky assets: An empirical investigation. *Journal of Finance.* 31, 447–458.

Johanson, S. (1988), Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, 12, pp. 231-254.

Johansen, S. (1991), Estimation and HypothesisTtesting of Cointegration Vectors in GaussianVvector Autoregressive Models, *Econometrica* 59, 1551-80.

Johansen, S. and K. Juselius (1990). Maximum likelihood estimation and inference on cointegration- with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52: 169-210.

Juselius, K., (I99I) Long-run relations in a well-defined statistical model for the data generating process: Cointegration analysis of the PPP and UIP relations between Denmark and Germany, in J. Gruber, ed.: *Econometric Decision Models:New' Methods of Modeling and Applications* (Springs Verlag, Berlin).

Kessel, R.A., 1956, Inflation caused wealth redistribution: A test of a hypothesis, *American Economic Review* 46, 128-141.

King, R.G., Watson, M.W., 1997. Testing Long-Run Neutrality. *Economic Quarterly*, Federal Reserve Bank of Richmond, 69-101, 83/3 summer.

Laidler, D. (1999). The Quantity of Money and Monetary Policy. *Working Paper* No. 99–5. Ottawa: Bank of Canada.

Lee, B.S, 1992, Causal Relationships Among Stock Returns, Interest Rates, Real Activity, and Inflation, *Journal of Finance, 47*, 1591-1603.

MacKinnon, J. (1991) 'Critical Values for Cointegration Tests', in Engle, R. and Granger, C. (eds) *Long run Economic Relationships*, Oxford: Oxford University Press.

MacKinnon, James G., Haug, Alfred A. and Michelis, Leo (1999), "Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration," *Journal of Applied Econometrics*, 14:563-577.

Marshall, D.A. (1992). Inflation and asset returns in a monetary Economy, *The Journal of Finance*, 47: 1315-1342.

Modigliani, F. and Cohn, R. A. (1979) Inflation, Rational Valuation, and the Market. *Financial Analysts Journal*, 35: 24-44.

Nelson C. R. 1976. Inflation and rates of return on common stocks. *Journal of Finance*, 31: 471-483.

Nigerian Stock Exchange Performance. http://www.mbendi.com/exch/16/p0005.htm [Downloaded 25-02-2011].

Pearce, D. and V. V. Roley (1988). Firm Characteristics, Unanticipated Inflation, and Stock Returns *The Journal of Finance*, Vol. 43(4): 965-981.

Pindyck, R. and D. Rubinfeld. 1998. Econometric Models and Economic Forecasts. Fourth Edition. New York: McGraw-Hill.

Qayyum, A. (2005), "Modelling the Demand for Money in Pakistan". *The Pakistan Development Review*. 44(3):233-252

Rapach, D.E. (2002). The long-run relationship between inflation and real stock prices. *Journal of Macroeconomics*. 24? 311-351

Samarakoon, L. P. (1996) Stock Market Returns and Inflation: Sri Lankan Evidence. *Sri Lankan Journal of Management*, 1(4): 293-311, 1996. Available SSRN:http://ssrn.com/abstract=1395015.

Sampson, E. (2009) Financial Meltdown: The Rate and Shape of Capital Markets. *Zenith Economic Quarterly* 4(1): 67-80.

Spyrou, S. I. (2001) Stock returns and inflation: evidence from an emerging market, *Applied Economic Letters*, 8: 447–50

Spyrou, S. I. (2004) Are Stocks a Good Hedge Against Inflation? Evidence from Emerging Markets. *Applied Economics*, 36: 41-48.

Subair, K. and O.M. Salihu (n.d.). Exchange Rate Volatility and the Stock Market: The Nigerian Experience. www.aabri.com/OC2010Manuscripts/OC10113.pdf. [downloaded 24-02-2011].

Yaya, O.S. and O.I Shittu (2010). On the Impact of Inflation and Exchange Rate on Conditional Stock Market Volatility: A Re-assessment. American Journal of Scientific and Industrial Research. DOI: 10.5251/ajsir.2010.1.2.115.117.

Yeh, C.C and C.F. Chi (2009)The Co-Movement and Long-Run Relationship between Inflation and Stock Returns: Evidence from 12 OECD Countries. *Journal of Economics and Management*. 5(2): 167-186.

Appendix

Nigerian Stock Exchange Performance

Indicator	**(End) 1995**	**(End) 1996**	**% Change**
All-share index	5092.15	6992.10	37.3
Market Capitalisation (N billion)	171.1	285.6	66.9
Turnover Volume (N million)	397	882	122.2
Turnover Value (N billion)	1.83	7.063	284.0
Number	31	36	16.1
Value (N billion)	7.063	21.500	202.800
Foreign Investment Portfolio Transactions ($US million)	1.137	32.99	
Average P/E Ratio	9.2	12.2	

Source: http://www.mbendi.com/exch/16/p0005.htm [Downlaod 25-02-2011].

Is the Stock Market a Leading Indicator of Economic Activity in Nigeria?

Alvan E. Ikoku[1]

In an effort to address the lacuna in leading indicator studies of African economies and Nigeria in particular, this paper examines the causal relationships among stock market prices, real GDP and the index of industrial production in Nigeria, using quarterly data from 1984Q1 to 2008Q4. Granger causality tests indicate bi-directional causality between stock prices and GDP but no causality between stock prices and industrial production or between GDP and industrial production. Stock prices and GDP are found to be cointegrated, leading to the estimation of vector error correction models. Out-of-sample forecasts constructed with AR(1), ARIMA, structural ARIMA, and VEC models indicate that stock prices contain information that can be used to improve the accuracy of GDP forecasts and enhance the conduct of macroeconomic policy in Nigeria.

Keywords: Leading indicators; stock index, Granger causality, cointegration, vector error correction models, forecasting GDP, Nigeria

JEL: E32, E37, G15, G17

1.0 Introduction

Policymakers in most advanced and several developing nations use economic indicators to predict the direction of aggregate economic activity. When these economic indicators can reliably signal changes in aggregate economic activity several months or quarters into the future, they facilitate the conduct of macroeconomic policies which must anticipate the future and take corrective action in order to keep the economy growing at, or close to, capacity with price stability. Because of their embodiment of expectations, financial market variables such as equity prices and the yield curve tend to perform well as leading indicators.

Our primary interest in this paper is to investigate whether or not stock prices are leading indicators of economic activity in Nigeria. Equity market prices reflect the expectations of investors and market operators regarding the performance of firms and the economy in general with respect to economic growth, profitability, the level of interest rates and inflation among other variables. To the extent that these expectations are largely correct, stock market prices could be used as an indicator of future economic activity. If stock prices can reliably predict GDP growth, then they can be used to create, along with other indicators, a composite index of leading economic activity. The improvement in forecasting accuracy to be derived from such a composite leading index will enhance the conduct of monetary and fiscal policies, moderate the vagaries of business cycles and significantly enhance economic welfare.

Leading indicators tend to perform better than benchmark autoregressive models in forecasting the future path of economic activity (Stock and Watson, 2003b). However, in order to perform well as leading indicators, Moolman and Jordaan (2005) claim that time series must have a stable relationship with the business cycle, need to be published in a timely manner, must be final data not subject to revisions and should be available on a monthly basis. Stock prices obviously meet three of the four requirements listed by Moolman and Jordan. However, we need to examine their relationship to the business cycle or aggregate economic activity in a rigorous manner in order to establish their suitability as leading indicators. Data constraints currently preclude an examination of the role of the term structure of interest rates as a leading indicator of economic activity in Nigeria.[2] However, few studies have been conducted on the role of stock prices as leading indicators in African countries[3], and this paper attempts to bridge this lacuna with respect to Nigeria.

The paper is organized as follows. Section 2 presents a review of the theoretical literature and empirical evidence on stock prices as leading indicators from advanced and emerging economies. In section 3, the data and methodology are discussed while the results of diagnostic tests, including unit root, Granger causality and Johansen

---

[1] Deputy Director & Head, Financial Policy Division, Monetary Policy Department, Central Bank of Nigeria.
*The views expressed in this paper are those of the author alone and do not necessarily represent those of the Central Bank of Nigeria or CBN policy. Useful contributions by seminar participants at the Central Bank of Nigeria are gratefully acknowledged. We would like to thank Dr. Sarah Alade, Deputy Governor, Economic Policy, and Dr. Okorie Uchendu, Director of Monetary Policy, for financial and moral support.

[2] A thorough analysis of the information content of the yield curve will become feasible over time as time series data on bond yields accumulate. Long-maturity instruments, such as the 10-year and 20-year bonds, were introduced in Nigeria as recently as August 2007 and March 2009, respectively.
[3] Jefferis, Okeahama and Matome (2001) and Mauro (2003) are among the few studies on African economies.

cointegration, are presented in section 4. In section 5, we present the results of out-of-sample forecasts conducted with and without stock prices as a structural variable. Section 6 concludes the paper.

2.0 Review of Theoretical and Empirical Literature

2.1 Theoretical Bases for Stock Prices as Leading Indicators

There are at least four theoretical bases for the role of stock prices as leading indicators of economic activity—stock prices as aggregators of expectations, the cost of raising equity capital, the financial accelerator and the wealth effect.

The standard valuation model recognizes the value of a share of common stock as the present value of the expected future dividends from owning stocks. The Gordon (1959) , or constant growth, model in equation (1) shows the now familiar relationship between expected dividends, D_1, the required return on equities, r, the anticipated growth rate of earnings, g, and the current price, P_0, of common stocks.

$$P_0 = \frac{D_1}{r - g} \tag{1}$$

This relationship holds even if an investor has a short time horizon. An investor with a one year horizon will receive D_1 plus P_1upon selling the stock. However, P_1 is a function of D_2 to D_∞. While computationally convenient, the Gordon model is valid only when r > g and when g, the growth rate, is constant (Brigham and Houston, 2007). More generally, the value of a stock today can be expressed at the present value of an infinite stream of dividends:

$$P_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1+r)^t} \tag{2}$$

If stock prices depend on expected dividends and dividends depend on the profitability of firms, then stock prices should embody expectations held by investors regarding the level of economic activity. This forward-looking property of stock markets suggests that stock prices would perform well as leading indicators, subject to the reliability of investors' forecasts of economic activity and corporate profits. Stock prices should decline if investors anticipate a slow-down in economic activity and rise if they expect an acceleration of economic activity. In short, stock and other asset prices are leading indicators of economic activity because they are forward-looking economic variables (Stock and Watson, 2003a). The behavior of global stock markets in the first three quarters of 2009 indicates that they anticipated the nascent economic recovery by one to two quarters.

Because the optimal capital structure usually involves a mix of debt and equity, the cost of equity capital is a significant portion of most firms' weighted average cost of capital, the hurdle rate for investment projects. Firms issuing equity in order to obtain investment funds must not only consider the required return on their equity but must also take flotation costs into account. Given the high cost of raising external equity, firms may be more willing to issue equity when stock prices are high in order to maximize the proceeds from selling ownership stakes. Even though some scholars [see for example, Ritter (1991), Baker and Wurgler (2000), and Hirshleifer (2001)] claim that firms knowingly sell overvalued equity to investors, thereby violating some of the tenets of the efficient market hypothesis, there is no doubt that higher stock prices are consistent with a lower cost of equity for firms. If a lower cost of equity reduces the weighted average cost of capital and makes more capital projects economically feasible, a positive relationship could develop between stock prices and subsequent economic activity.

The *financial accelerator* channel stems from that fact that rising stock prices lead to an improvement in the balance sheets of firms and households which, in turn, improves their creditworthiness [see Fazzari et al. (1988) and Bernanke *et al*. (1996)]. The increase in creditworthiness reduces borrowing costs and increases the borrowing capacity of firms and households, stimulating investment spending and current consumption. Predictably, the financial accelerator also operates in downturns. According to Bernanke et al., "the theory underlying the financial accelerator suggests that (1) borrowers facing relatively high agency costs in credit markets will bear the brunt of economic downturns (the flight to quality); and that (2) reduced spending, production and investment by high-agency-cost borrowers will exacerbate the effects of recessionary shocks" p. 14. The financial accelerator is similar to the cost of capital channel because both operate through the capital structure of firms and households. However, while the cost of capital channel is conventionally deemed to operate through the issuance of equity and the financial accelerator through the issuance of debt, both channels could conceivable operate through the issuance of both debt and equity.

The *wealth effect* operates via the consumption function, when households consume not only out of earned income but also as a result of perceived increases in the value of their assets, including real estate and equity. Increasing stock market wealth seems to improve consumer sentiment and raise expectations of higher incomes in the future (Otoo, 1999). Case *et al*., (2005) estimated the marginal propensity to consume out of housing wealth in the range of 11 – 17 percent and out of equity wealth of about 2 percent in 14 western nations. Using micro data for the U.S., Bostic et al. (2009) found "an important role for both financial wealth and housing wealth in the determination of household consumption patterns. The results suggest the estimation of significant coefficients in both cases; the implied elasticity with respect to total consumption is .02 percent for financial assets, and .04 percent for house values. House values were much more important for non-durable and food consumption and financial assets were much more important for durable consumption" p. 14. The operation of the wealth effect was palpable in the United States before the financial and economic crises, with households using home equity loans to tap the rising values of their homes to fund consumption spending. The consequent collapse of U.S. consumption expenditures following the decimation of asset prices suggests that the wealth effect operates with rising as well as falling asset values. We must keep in mind, however, that the importance of the wealth effect in determining the role of stock prices as leading indicators depends crucially on the extent of stock ownership in a country. There is more empirical evidence in favor of the wealth effect in the U.S. than in several European nations with lower stock-ownership rates [see Paiella (2007) and Simone (2009)]

2.2 Empirical Evidence on Stock Prices as Leading Indicators

Several studies of advanced economies have found stock prices to be a fairly reliable indicator of GDP growth. Because of its leading role in the use of leading indicators to predict business cycles, most of the studies of the advanced economies have been done on the U.S. economy. The Dow Jones composite index of stock prices was included in the index of leading indicators for the U.S. economy more than seventy years ago by Mitchell and Burns (1938). However, studies of other advanced economies are becoming more prevalent in the literature, as the leading indicator approach becomes more widely adopted. Table 1a summarizes the empirical evidence from the advanced economies.

Leading indicator studies of emerging markets are much less common than studies of the advanced economies. This paucity of studies may be partly due to data inadequacies, as quarterly GDP surveys have only recently begun for many less developed countries. Leading indicator studies of African economies are quite rare and usually part of group studies of several advanced and developing countries. Most studies on African stock markets focus on the role of stock market development, as measured by the ratio of market capitalization to GDP, in economic growth. For example, Osinubi (2004), Adebiyi (2005), and Nurudeen (2009) found that there was a positive link between stock market development and economic growth in Nigeria. However, Akinlo et al. (2009) found weak evidence of this relationship in Nigeria, even though they found that stock market development Granger-caused economic growth in Egypt and South Africa. The focus in leading indicator studies is on the information content of stock prices in terms of their ability to help predict the direction of economic activity in the near future, not on the long run relationship between financialization and economic growth. Table 1b summarizes the empirical evidence from the emerging economies.

The review of the literature indicates that stock prices have a sound theoretical basis for leading economic activity. The empirical evidence is mixed, but mostly supportive of this hypothesis. Among advanced countries, stock markets tend to be stronger leading indicators in countries with Anglo-Saxon backgrounds; this is perhaps due to the fact that stock markets tend to play larger roles in the economies of such nations. Among emerging economies, stock prices tend to become stronger leading indicators as the economy develops and financial markets become larger in relation to GDP. A rigorous investigation of the role of stock markets in predicting economic activity in Nigeria will enhance the body of knowledge in this area as well as provide policymakers with an additional tool with which to manage the Nigerian economy.

Table 1a – Empirical Evidence from Advanced Economies

Study	Nation(s)	Data Range	Periodicity	Findings
Fama (1981)	U.S.	1953 - 1987	Monthly, Quarterly, Annual	Stock prices led all real variables.
Pearce (1983)	Canada, France, Germany, U.K. & U.S.	1955 - 1983	Quarterly	Stock prices tend to rise midway through recession.
Huang and Kracaw (1984)	U.S.	1962 - 1978	Quarterly	Stock prices led GDP by four quarters.
Campbell (1989)	U.S.	1953 - 1989	Quarterly	Stock prices and Yield Curve led GDP.
Lee (1992)	U.S.	1947 - 1987	Monthly	Stock prices Granger-cause industrial production.
Comincioli (1996)	U.S.	1970 - 1984	Quarterly	Stock prices Granger-cause GDP with lags of one to three quarters.
Otoo (1999)	U.S.	1980 - 1999	Monthly	Stock prices are leading indicator.
Choi, Hauser and Kopecky (1999)	Canada, France, Germany, Italy, Japan, U.K. & U.S.	1957 - 1996	Monthly	Stock prices useful for forecasting only in U.S., Canada, U.K. & Japan.
Burgstaller	Austria, Japan & U.S.	1976 - 2002	Monthly	Stock prices had no predictive power; Stock prices weakly affect consumption.
Stock and Watson (2003a)	Canada, France, Germany, Italy, Japan, U.K. & U.S.	1959 - 1999	Monthly, Quarterly, Annual	Inconsistent results from review of sixty-six papers.
Stock and Watson (2003b)	U.S.	1986 - 2002	Quarterly	Stock prices and other leading indicators superior to benchmark AR model.
Gan, Lee, Yong and Zang (2006)	New Zealand	1990 - 2003	Monthly	Stock index caused by GDP (not leading indicator)
Foresti (2007)	U.S.	1959 - 1999	Quarterly	Stock prices had predictive power with lags of up to five quarters.

Table 1b – Empirical Evidence from Emerging Economies

Study	Nation(s)	Data Range	Periodicity	Findings
Leigh (1997)	Singapore	1975 - 1991	Quarterly	Stock prices Granger-cause GDP.
Christoffersen and Slok (2000)	Czech Republic, Hungary, Poland, Russia, Slovakia & Slovenia	1994 - 1999	Monthly	Stock prices led industrial production by one to six months.
Husain and Mahmood (2001)	Pakistan	1959 - 1999	Quarterly	Stock prices lagged GDP (not leading indicator).
Nishat and Shaheen (2004)	Pakistan	1973 - 2004	Quarterly	Stock prices led Industrial Production by one quarter.
Jefferis, Okeahalam and Matome (2001)	Botswana, South Africa & Zimbabwe	1985 - 1996	Quarterly	Stock prices cointegrated with GDP; leading indicator.
Mauro (2003	Argentina, Chile, Greece, India, Mexico, South Korea, Thailand & Zimbabwe	1971 - 1998	Quarterly, Annual	Stock prices in all nations except India led GDP by up to four quarters; signal stronger in nations with high market capitalization.
Amadja (2005)	Indonesia, Malaysia, the Philippines, Singapore & Thailand	1997 - 2003	Monthly	Stock prices Granger-caused GDP in Singapore and Thailand; no causality in Malaysia and the Philippines.
Mun, Siong & Thing (2008)	Malaysia	1977 - 2006	Annual	Stock prices Granger-caused GDP with a lag of up to two years.
Bahadur and Neupane (2006)	Nepal	1988 - 2005	Annual	Stock prices had no impact on GDP. However, market capitalization Granger-caused GDP with a four-year lag.
Pilinkus (2009)	Lithuania	1999 - 2008	Monthly	Stock prices Granger-caused GDP.

3.0 Methodology and Data

3.1 Methodology

Two basic methodological approaches are adopted to determine whether or not the stock market is a leading indicator of economic activity in Nigeria. The first approach is to conduct the familiar test proposed by Granger (1969) in order to determine whether or not changes in nominal or real stock prices precede changes in economic activity (as measured by GDP or IIP). The results of the Granger-causality test are crucial for the use of stock prices as a leading indicator, especially if the lead over economic activity is reliable and of sufficient length to give useful signals to policy makers. It is important to mention here that the Granger-causality test is actually a *test of precedence* and does not imply that changes in stock prices cause changes in economic activity in the conventional sense. In addition to Granger-causality tests, we utilize unit root tests, correlation analysis and cointegration tests to analyze the basic properties of the time series.

The second methodological approach is to determine the usefulness of stock prices in forecasting economic activity. An AR(1) is used as the baseline forecasting model, augmented by an optimized[4] autoregressive integrated moving average (ARIMA) model. Then we build four structural models—two ARIMA models and two Vector Error Correction models (VECMs) employing nominal and real values of the stock index, respectively, as structural variables. We seek to determine whether or not the structural models have superior forecasting ability, in terms of smaller forecast errors, compared to the baseline AR(1) and ARIMA models.

In order to simulate an actual forecasting environment, the 100-quarter sample period is divided into two sub periods—data from 1984Q1 to 2007Q2 (94 percent of the total) are used to estimate models while data from 2007Q3 to 2008Q4 (6 percent of the total) are used for forecast evaluation. As such, the out-of-sample performance of the models can be estimated.

With the combination of formal tests and forecast simulation, we should be able to ascertain the information content of stock prices for the business cycle in Nigeria. Needless to say, the ability to improve forecasts of economic activity is the *raison d'être* of a leading indicator and would indicate whether or not the stock index, in nominal or real terms, should be incorporated in a composite index of leading indicators in Nigeria.

3.2 Data

Because the All Share Index (ASI) of the Nigerian Stock Exchange (NSE) was formulated in January 1984, we use ASI data from the first quarter of 1984 to the fourth quarter of 2008, a total of 100 observations. The ASI is a market-value-weighted index representing all the stocks traded on the floor of the NSE; it is the only stock index with the coverage and vintage required to truly discern the role of the stock market as a leading indicator of economic activity in Nigeria. Nominal values of ASI are deflated with the consumer price index (CPI) to create another variable, real ASI (ASIR). CPI statistics were obtained from the National Bureau of Statistics (NBS).

Real Gross Domestic Product (GDP) and an Index of Industrial Production (IIP) are used as measures of economic activity[5] for the sample period. Both variables are produced through surveys conducted by the National Bureau of Statistics and the Central Bank of Nigeria. Economic activity in Nigeria is dominated by Agriculture, which accounted for 42.1 percent of GDP in 2008. This was followed by Industry (22 percent), Wholesale and Retail Trade (17.3 percent), Services (16.8 percent) and Building and Construction (1.8 percent). Remarkably, the share of Agriculture in Nigeria's GDP increased by 11.6 percentage points during the last twenty five years, from 30.5 percent in 1984 to 42.1 percent in 2008. During the same period, the share of industry in Nigeria's GDP declined from 42.4 percent to 22 percent, a loss of 20.4 percentage points.

4.0 Descriptive Statistics and Diagnostic Tests

4.1 Descriptive Statistics

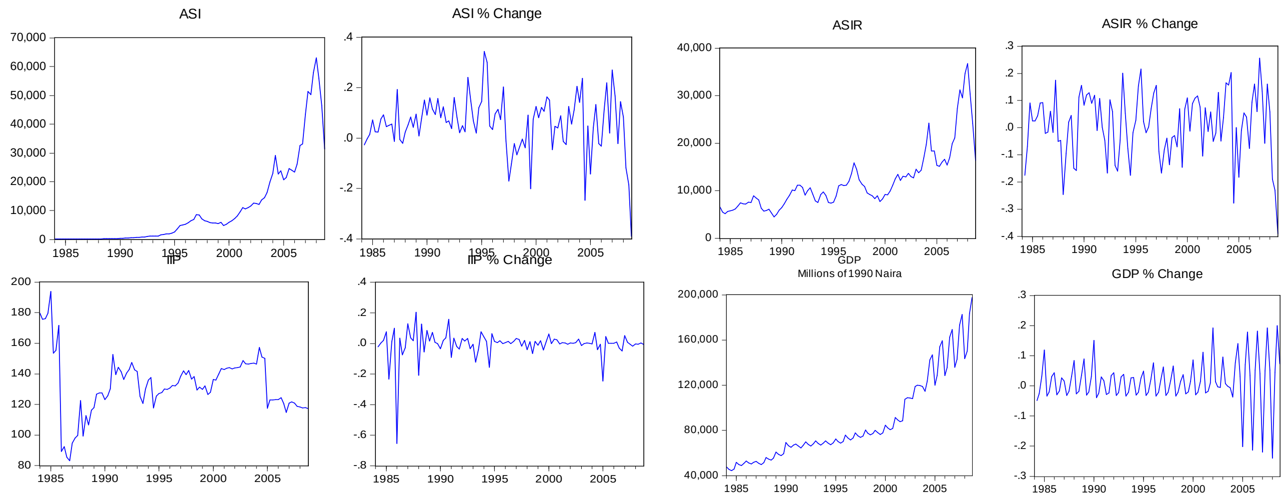
Figure 1 contains graphical representations of the variables and their quarterly growth rates. ASI, ASIR and GDP exhibit strong upward trends, while IIP seems to exhibit a substantial degree of mean-reversion. Because quarterly

[4] This model is selected on the basis of having the lowest information criteria (i.e., AIC and SIC) values.

[5] While most studies use either GDP or the Index of Industrial Production as the measure of economic activity, a number of studies, for example Fama (1981), utilize both variables. We employ both variables in this paper in the interest of completeness.

GDP surveys by the NBS commenced in 2004, annual GDP data were interpolated between 1984 and 2003 in order to derive quarterly equivalents.

Figure 1 – ASI, ASIR, GDP and IIP (1984 – 2008)



Descriptive statistics for the four time series show that are ASI and ASIR returns and GDP and IIP growth rates are negatively skewed with fat tails, judging by the kurtosis statistics. During the sample period, the ASI turned in mean quarterly returns of 5.7 percent (median of 6.7 percent); the ASIR had mean quarterly returns of 0.9 percent (median of 2.5 percent); the mean quarterly GDP growth rate was 1.43 percent (median of 1.43 percent); and the mean quarterly IIP growth rate was -0.43 percent (median of 0.35 percent). The Jarque-Bera statistics suggest that the null hypothesis of normality would be rejected for all four time series, even though the probability of 0.048 for ASIR is close to the 5 percent threshold.

4.2 Unit Root Tests

Dickey (1976) and Fuller (1976) show that the least squares estimator is biased downward in the presence of unit roots. Since the *Dickey-Fuller bias* can be expected to reduce the accuracy of forecasts, we test for the presence of this bias using the *Augmented Dickey-Fuller* (ADF) test as well as the *Phillips-Perron* (PP) test proposed by Phillips and Perron (1988).

Bierens (2003) shows that an AR (p) process as shown in equation (3):

$$y_t = \beta_0 = \sum_{j=1}^{p} \beta_j y_{t-j} + u_t, \qquad (3)$$

$$u_t \sim iidN(0, \sigma^2)$$

can be written, through recursive replacement with differenced terms, as equation (4):

$$\Delta y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j \Delta y_{t-j} + \alpha_p y_{t-p} + u_t, \qquad (4)$$

$$u_t \sim iidN(0, \sigma^2)$$

where $\alpha_0 = \beta_0$, $\alpha_j = \sum_{i=1}^{j} \beta_i - 1,\ j = 1,..., p.$

The ADF tests the null hypothesis that $\alpha_p = 0$ against the alternative hypothesis that $\alpha_p < 0$. If the AR(p) process has a unit root, then $\alpha_p = 0$. On the other hand, if the process is stationary, then $\alpha_p < 0$. In contrast to the ADF, the PP test does not require that the ARIMA process be specified and would, thus, be less subject to misspecification than the ADF test.

Table 2 – Augmented Dickey-Fuller Unit Root Tests

PP Tests - Levels

Null Hypothesis: Variable has a unit root

		Without Trend				With Trend			
Variable:		**ASI**	**ASIR**	**GDP**	**IIP**	**ASI**	**ASIR**	**GDP**	**IIP**
PP test statistics:		-0.9923	1.4539	2.1929	-3.6161	-2.3208	-0.7040	-2.9202	-3.5647
Test critical values:	1% level	-3.4977	-3.5022	-3.4977	-3.4977	-4.0534	-4.0597	-4.0534	-4.0534
	5% level	-2.8909	-2.8929	-2.8909	-2.8909	-3.4558	-3.4589	-3.4558	-3.4558
	10% level	-2.5825	-2.5836	-2.5825	-2.5825	-3.1537	-3.1555	-3.1537	-3.1537
MacKinnon prob-values:		0.7537	0.9991	0.9999	0.0071	0.4186	0.9695	0.1608	0.0382

PP Tests - First Differences

Null Hypothesis: Variable has a unit root

		Without Trend				With Trend			
Variable:		**D(ASI)**	**D(ASIR)**	**D(GDP)**	**D(IIP)**	**D(ASI)**	**D(ASIR)**	**D(GDP)**	**D(IIP)**
PP test statistics:		-4.7602	-7.2805	-10.9273	-12.7637	-4.6346	-7.4848	-13.1681	-12.7458
Test critical values:	1% level	-3.4984	-3.5030	-3.4984	-3.4984	-4.0544	-4.0609	-4.0544	-4.0544
	5% level	-2.8912	-2.8932	-2.8912	-2.8912	-3.4563	-3.4594	-3.4563	-3.4563
	10% level	-2.5827	-2.5837	-2.5827	-2.5827	-3.1540	-3.1558	-3.1540	-3.1540
MacKinnon prob-values:		0.0001	0.0000	0.0000	0.0001	0.0016	0.0000	0.0000	0.0000

Table 2 shows the results of the ADF tests on ASI, ASIR, GDP and IIP. The tests on the levels of the variables, with only a constant and no trend in the equations, show that the null hypothesis of a unit root cannot be rejected for ASI, ASIR and GDP at either the 1 percent, 5 percent or 10 percent levels; their MacKinnon (1996) one-side p-values are 0.9999, 0.9991 and 0.9998, respectively. However, with a p-value of 0.0305, the null hypothesis of a unit root can be rejected at the 5 percent level but not at the 1 percent level for IIP. ADF tests on the first differences of the variables result in a strong rejection of the null hypothesis of a unit root for ASI, ASIR and IIP. However, this is not the case for GDP, which has a p-value of 0.2381. The tests on the levels of the variables with a constant and a linear trend in the equations have similar results to those with a trend except that the p-value for IIP has increased to 0.1179. The ADF test results with first differences are not very sensitive to the addition of a linear trend to the equations, giving essentially the same results for ASI, ASIR and IIP, but with a p-value that decreases to 0.0899 for GDP.

The PP tests shown in table 3 give the same results as the ADF tests with respect to ASI and ASIR, suggesting that both times series are integrated of order one, i.e., I(*1*). With respect to GDP, the PP test is more conclusive than the ADF test, as the series becomes stationary with first differencing, suggesting an I(*1*) process. The PP test on IIP, with no trend in the equation, rejects the null hypothesis of a unit root at levels or first differencing, with p-values of 0.0071 and 0.0001, respectively. However, when a linear trend is added to the equation, the PP test on IIP cannot reject the null hypothesis of a unit root at the 1 percent level on account of the p-value of 0.0382.

Even though unit roots tests are known to have low power, one can reasonably proceed on the assumption that ASI, ASIR and GDP are I(*1*) series, while the IIP could be considered an I(*0*) series based on a 5 percent significance level. ASI, ASIR and GDP are non-stationary but could be made stationary with first differencing while IIP is stationary. Where differencing is not appropriate, ARMA terms could be used to realize white noise errors.

Table 3 – Phillips-Peron Unit Root Tests

ADF Tests - Levels

Null Hypothesis: Variable has a unit root

		Without Trend				With Trend			
Variable:		**ASI**	**ASIR**	**GDP**	**IIP**	**ASI**	**ASIR**	**GDP**	**IIP**
ADF test statistics:		2.1662	1.4539	1.9407	-3.0907	0.1366	-1.1358	-0.2014	-3.0762
Test critical values:	1% level	-3.5039	-3.5022	-3.5022	-3.4984	-4.0620	-4.0609	-4.0597	-4.0544
	5% level	-2.8936	-2.8929	-2.8929	-2.8912	-3.4600	-3.4594	-3.4589	-3.4563
	10% level	-2.5839	-2.5836	-2.5836	-2.5827	-3.1561	-3.1558	-3.1555	-3.1540
MacKinnon prob-values:		0.9999	0.9991	0.9998	0.0305	0.9972	0.9167	0.9922	0.1179

ADF Tests - First Differences

Null Hypothesis: Variable has a unit root

		Without Trend				With Trend			
Variable:		**D(ASI)**	**D(ASIR)**	**D(GDP)**	**D(IIP)**	**D(ASI)**	**D(ASIR)**	**D(GDP)**	**D(IIP)**
ADF test statistics:		-6.6700	-7.3069	-2.1182	-12.8773	-7.1430	-7.5022	-3.2047	-12.8433
Test critical values:	1% level	-3.5039	-3.5030	-3.5022	-3.4984	-4.0620	-4.0609	-4.0597	-4.0544
	5% level	-2.8936	-2.8932	-2.8929	-2.8912	-3.4600	-3.4594	-3.4589	-3.4563
	10% level	-2.5839	-2.5837	-2.5836	-2.5827	-3.1561	-3.1558	-3.1555	-3.1540
MacKinnon prob-values:		0.0000	0.0000	0.2381	0.0001	0.0000	0.0000	0.0899	0.0000

4.3 Correlation Coefficients

Correlation coefficients provide an initial look at the relationship among the variables. In order to explore the effect of the interpolation of GDP between 1984 and 2003, the coefficients were computed for three sub-samples—1984Q1-2003Q4, 2004Q1-2008Q4 and 1984Q1-2008Q4.

ASI, ASIR and GDP are highly and positively correlated. The weakest coefficient between ASI and GDP was 0.4087 in the 2004Q1 to 2008Q4 sample, and the highest was 0.9291 in the 1984Q1 to 2003Q4 sample. Likewise, the weakest coefficient between ASIR and GDP was 0.2162 in the 2004Q1 to 2008Q4 sample, and the highest was 0.8443 in the 1984Q1 to 2008Q4 sample. In contrast, ASI, ASIR and IIP had a weaker and, sometimes, negative relationship. At first glance, the negative correlation between ASI and IIP and ASIR and IIP in the 1984 to 2008 period might call into question the accuracy of IIP as a measure of economic activity. However, industrial production in Nigeria was declining during this period, while the stock market was in the midst of a boom for most of the period.

We conclude that ASI and ASIR were more highly correlated with GDP than with IIP, and that the interpolation of GDP values between 1984 and 2003 did not have an appreciable effect on the relationship among the variables.

4.4 Granger Causality Tests

Granger causality tests were conducted, using bivariate regressions as shown in equations (5) and (6), between ASI and GDP and between ASIR and GDP[6], using 1 to 10 quarterly lags, *l*.

$$\mathrm{ASI}_t = \alpha_0 + \alpha_1 \mathrm{ASI}_{t-1} + ... + \alpha_l \mathrm{ASI}_{t-l} + \beta_1 \mathrm{GDP}_{t-1} + ... + \beta_l \mathrm{GDP}_{t-l} + \varepsilon_t \tag{5}$$

$$\mathrm{GDP}_t = \alpha_0 + \alpha_1 \mathrm{GDP}_{t-1} + ... + \alpha_l \mathrm{GDP}_{t-l} + \beta_1 \mathrm{ASI}_{t-1} + ... + \beta_l \mathrm{ASI}_{t-l} + u_t \tag{6}$$

[6] Granger causality tests conducted between ASI and IPP and between ASIR and IIP showed no causality among the variables.

The null hypothesis is that GDP does not Granger-cause ASI in equation (5) and that ASI does not Granger-cause GDP in equation (6). F-tests was conducted with the joint hypothesis that β_1 through β_{10} are zero. The tests were conducted with the levels and first differences of ASI and GDP and ASIR and GDP.

The results with the variables in levels, shown in table 4, indicate that ASI causes GDP at lags 1 and 2; GDP causes ASI at lags 3 and 4, and there is bi-directional causality between ASI and GDP at lags 5 through 10. With respect to ASIR and GDP, ASI causes GDP at lag 1; GDP causes ASIR at lags 2 to 4 and 7 to 10; and there is bi-directional causality between the two variables at lags 5 and 6.

Table 4 – Granger Causality Tests – Levels

ASI vs. GDP

# of Lags	From ASI to GDP /1	From GDP to ASI /2	Test Result /3
1	0.0013	0.2429	ASI causes GDP.
2	0.0026	0.0565	ASI causes GDP.
3	0.0622	0.0016	GDP causes ASI.
4	0.1144	0.0036	GDP causes ASI.
5	0.0008	0.0058	Bi-directional causality.
6	0.0036	0.0019	Bi-directional causality.
7	0.0071	0.0005	Bi-directional causality.
8	0.0147	0.0004	Bi-directional causality.
9	0.0029	2.00E-05	Bi-directional causality.
10	0.0458	1.00E-05	Bi-directional causality.

ASIR vs. GDP

# of Lags	From ASIR to GDP /4	From GDP to ASIR /5	Test Result /3
1	0.0428	0.0599	ASIR causes GDP.
2	0.1393	0.0007	GDP causes ASIR.
3	0.1437	0.0001	GDP causes ASIR.
4	0.2157	0.0004	GDP causes ASIR.
5	0.0068	0.0009	Bi-directional causality.
6	0.0289	0.0003	Bi-directional causality.
7	0.1011	4.00E-05	GDP causes ASIR.
8	0.1830	1.00E-06	GDP causes ASIR.
9	0.4448	0.0000	GDP causes ASIR.
10	0.7402	6.00E-05	GDP causes ASIR.

1/ The numbers are p-values for the null hypothesis "ASI does not cause GDP."
2/ The numbers are p-values for the null hypothesis "GDP does not cause ASI."
3/ The test result is based on a 5 percent significance level.
4/ The numbers are p-values for the null hypothesis "ASIR does not cause GDP."
5/ The numbers are p-values for the null hypothesis "GDP does not cause ASIR."

Table 5 shows the results with first differences of the variables. GDP causes ASI at lag 1 while there is bi-directional causality between the variables at lags 2 through 10. ASIR was found to cause GDP at lags 2 and 4 to 6, while GDP was found to cause ASIR at lags 7 though 10.

These results suggest that ASI and ASI could be useful in forecasting GDP with relatively short lags. In addition, the causal relationship between ASI and GDP seems more stable than that between ASIR and GDP.

4.5 Cointegration Tests

According to Engle and Granger (1987), if two variables are both I(*1*), it is generally true that a linear combination of the variables will also be I(*1*). However, a linear combination of the variables may exist that is I(*0*). If the

Table 5 – Granger Causality Tests – First Differences

D(ASI) vs. D(GDP)

# of Lags	From D(ASI) to D(GDP) /1	From D(GDP) to D(ASI) /2	Test Result /3
1	0.1064	0.0338	D(GDP) causes D(ASI).
2	0.0108	0.0050	Bi-directional causality.
3	0.0255	0.0035	Bi-directional causality.
4	0.0002	0.0031	Bi-directional causality.
5	0.0002	0.0002	Bi-directional causality.
6	0.0025	0.0001	Bi-directional causality.
7	0.0089	4.00E-05	Bi-directional causality.
8	0.0059	2.00E-06	Bi-directional causality.
9	0.0233	2.00E-06	Bi-directional causality.
10	0.0473	4.00E-06	Bi-directional causality.

D(ASIR) vs. D(GDP)

# of Lags	From D(ASIR) to D(GDP) /4	From D(GDP) to D(ASIR) /5	Test Result /3
1	0.3186	0.1069	No causality.
2	0.0547	0.2107	D(ASIR) causes D(GDP).
3	0.1794	0.2224	No causality.
4	0.0020	0.2083	D(ASIR) causes D(GDP).
5	0.0054	0.2482	D(ASIR) causes D(GDP).
6	0.0585	0.3109	D(ASIR) causes D(GDP).
7	0.1344	0.0009	D(GDP) causes D(ASIR).
8	0.3056	0.0004	D(GDP) causes D(ASIR).
9	0.6689	0.0003	D(GDP) causes D(ASIR).
10	0.5504	0.0003	D(GDP) causes D(ASIR).

1/ The numbers are p-values for the null hypothesis "D(ASI) does not cause D(GDP)."
2/ The numbers are p-values for the null hypothesis "D(GDP) does not cause D(ASI.)"
3/ The test result is based on a 5 percent significance level.
4/ The numbers are p-values for the null hypothesis "D(ASIR) does not cause D(GDP)."
5/ The numbers are p-values for the null hypothesis "D(GDP) does not cause D(ASIR)."

variables GDP and ASI are I(1), then linear combinations of GDP and ASI will generally also be I(*1*). Nevertheless, if there is a vector such that the linear combination in equation (7)

$$z_t = \mathrm{GDP}_t - \alpha - \beta \mathrm{ASI}_t \qquad (7)$$

is I(*0*), then GDP and ASI are cointegrated of order (1,1), i.e., CI(1,1), with (1, -β) termed the cointegrating vector. Cointegration implies that there is a long-run equilibrium relationship between the two variables, and z_t is the equilibrium error.

Having established, with Granger-causality tests, that ASI and ASIR have a strong short-run relationship with GDP but that ASI and ASIR have no statistically significant relationship with IIP, we explore the long-run relationship between ASI, ASIR and GDP using three cointegration tests—the Johansen (1991, 1995) test,[7] the Engle-Granger (1987) test and the Phillips-Ouliaris (1990) test. As required by the Johansen test, ASI, ASIR and GDP are non-stationary and integrated of the same order.

Table 6 shows the results of the Johansen trace and maximum eigenvalue tests, with a linear deterministic trend[8], between nominal and real stock indices and GDP. Between ASI and GDP, both trace and maximum eigenvalue tests reject the null hypothesis of no cointegrating equation at the 1 percent and 5 percent levels, with p-values of 0.0000 for both tests. However, the null hypothesis of at most one cointegrating equation is not rejected by either test, with p-values of 0.4871 for both tests.

Between ASIR and GDP, again both trace and maximum eigenvalue tests reject the null hypothesis of no cointegrating equation at the 1 percent and 5 percent levels, with p-values of 0.0001 and .0000, respectively.

[7] Johansen and Jeselius (1990) applied this technique to money demand in Denmark and Finland.
[8] We examined the sensitivity of the Johansen tests to the trend assumptions on the cointegrating equations. The tests were not sensitive to the trend assumption, indicating the presence of one cointegrating equation in all trend specifications.

Interestingly, the null hypothesis of at most one cointegrating equation is also rejected by both the trace and maximum eigenvalue tests, with p-values of 0.0000 for both tests. The results indicate more than one cointegrating equation between ASIR and GDP.

Table 6 – Johansen Cointegration Tests

D(ASI) and D(GDP)

Trace Test

Hypothesized # of CE's	Statistic	0.05 Critical Value	Prob.**
None*	91.5042	15.4947	0.0000
At most 1	0.48303	3.84147	0.4871

Maximum Eigenvalue Test

Hypothesized # of CE's	Statistic	0.05 Critical Value	Prob.**
None*	91.0212	14.2646	0.0000
At most 1	0.48303	3.84147	0.4871

Normalized Cointegrating Coefficients
(Standard Error in Parenthesis)

D(ASI)	D(GDP)
1.0000	-1.8899
	(0.1488)

D(ASIR) and D(GDP)

Trace Test

Hypothesized # of CE's	Statistic	0.05 Critical Value	Prob.**
None*	109.7725	15.4947	0.0001
At most 1	17.98613	3.84147	0.0000

Maximum Eigenvalue Test

Hypothesized # of CE's	Statistic	0.05 Critical Value	Prob.**
None*	91.7864	14.2646	0.0000
At most 1	17.98613	3.84147	0.0000

Normalized Cointegrating Coefficients
(Standard Error in Parenthesis)

D(ASIR)	D(GDP)
1.0000	-3.4110
	(0.2766)

* Denotes rejection of the hypothesis at the 0.05 level.
**MacKinnon-Haug-Michelis (1999) p-values.

Table 7 shows the outcome of the Engle-Granger and Phillips-Ouliaris cointegration tests. With respect to the Engle-Granger test, the null hypothesis of no cointegration cannot be rejected for ASI and GDP, with p-values of 0.9658 and 0.0196, respectively. However, the null hypothesis that ASIR and GDP are not cointegrated can be rejected at the 5 percent level, with p-values of 0.0000 and 0.0189, respectively. The Phillips-Ouliaris tests strongly reject the null hypotheses of no cointegration between ASI and GDP and between ASIR and GDP, with p-values of 0.0000 throughout.

In summary, all three tests indicate that ASIR and GDP are cointegrated, while the Johansen and Phillips-Ouliaris tests indicate that ASI and GDP are cointegrated. Based on the outcome of the tests, one can conclude that there is a long run equilibrium relationship between the nominal and real stock indices and real economic activity in Nigeria.

5.0 Forecasting GDP with Stock Prices

5.1 Univariate Models

In order to ascertain the information content of stock prices for the business cycle in Nigeria, we start by estimating two univariate GDP models—an AR(1) model and an ARIMA model. All the models were estimated with data from 1984Q1-2007Q2.

Table 7 – Engle-Granger and Phillips-Ouliaris Cointegration Tests

Engle-Granger Cointegration Tests

D(ASI) and D(GDP)

Null hypothesis: Series are not cointegrated

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
D(ASI)	-1.2734	0.9573	-3.7761	0.9658
D(GDP)	-2.8677	0.3432	-29.1963	0.0196

D(ASIR) and D(GDP)

Null hypothesis: Series are not cointegrated

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
D(ASIR)	-7.5970	0.0000	-73.5708	0.0000
D(GDP)	-2.87284	0.3408	-29.35149	0.0189

*MacKinnon (1996) p-values.

Phillips-Ouliaris Cointegration Tests

D(ASI) and D(GDP)

Null hypothesis: Series are not cointegrated

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
D(ASI)	-7.6989	0.0000	-84.8174	0.0000
D(GDP)	-14.8054	0.0000	-59.5247	0.0000

D(ASIR) and D(GDP)

Null hypothesis: Series are not cointegrated

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
D(ASIR)	-7.6289	0.0000	-75.0430	0.0000
D(GDP)	-14.74491	0.0000	-59.56045	0.0000

*MacKinnon (1996) p-values.

The AR(1) model is commonly used as the benchmark for evaluating the accuracy of more sophisticated forecasting models.[9] If a GDP model with a structural variable, such as the ASI or ASIR, were to perform better than the AR(1) model in out-of-sample forecasts, then stock prices are deemed to contain information useful in predicting GDP.

Because the unit root tests conducted above suggest that GDP is I(*1*), we use the first difference of GDP with autoregressive and moving average terms, following Box and Jenkins (1976), to create the ARIMA model. Forty-eight regression models were estimated with a maximum of six AR and MA terms. The Akaike and Schwarz information criteria for the models are shown in table 8. Both model selection criteria suggest an ARIMA (6, 1, 2) model as the best of the forty-eight models estimated, with AIC and SIC statistics of 19.232 and 19.487, respectively.

Table 9 shows the coefficients and other statistics from the univariate and structural models. The highly significant coefficient of 0.9808 on the AR(1) model suggests a high degree of persistence in the GDP series, while the adjusted r-squared of 0.9314 indicates a fairly good fit, even though this may have been inflated as a result of autocorrelation. The ARIMA model shows statistically significant AR(2), AR(4), AR(6) and MA(2) terms and

[9] Stock (2003) suggests a simple rule in forecasting time series "even if your main interest is in more sophisticated models, it pays to maintain benchmark forecasts using a simple model with honest forecast standard errors evaluated using a simulated real time experiment, and to convey the forecast uncertainty to the consumer of the forecast" p. 581.

Table 8 – ARIMA Model Selection Criteria

ARMA Specification	AIC	SIC	ARMA Specification	AIC	SIC
(0, 1)	20.669	20.724	(3, 4)	19.292	19.514
(0, 2)	20.490	20.572	(3, 5)	19.289	19.539
(0, 3)	20.292	20.400	(3, 6)	19.272	19.550
(0, 4)	19.780	19.916	(4, 0)	19.405	19.545
(0, 5)	19.796	19.959	(4, 1)	19.384	19.552
(0, 6)	19.766	19.957	(4, 2)	19.340	19.536
(1, 0)	20.843	20.898	(4, 3)	19.282	19.505
(1, 1)	20.662	20.744	(4, 4)	19.260	19.512
(1, 2)	20.518	20.627	(4, 5)	19.272	19.552
(1, 3)	20.173	20.310	(4, 6)	19.271	19.579
(1, 4)	19.803	19.968	(5, 0)	19.428	19.597
(1, 5)	19.807	19.998	(5, 1)	19.419	19.616
(1, 6)	19.789	20.009	(5, 2)	19.362	19.587
(2, 0)	19.953	20.036	(5, 3)	19.238	19.492
(2, 1)	19.908	20.018	(5, 4)	19.262	19.544
(2, 2)	19.705	19.843	(5, 5)	19.284	19.594
(2, 3)	19.668	19.833	(5, 6)	19.268	19.606
(2, 4)	19.386	19.579	(6, 0)	19.293	19.491
(2, 5)	19.364	19.585	(6, 1)	19.308	19.535
(2, 6)	19.382	19.630	(6, 2)	19.232	19.487
(3, 0)	19.811	19.922	(6, 3)	19.250	19.533
(3, 1)	19.533	19.672	(6, 4)	19.266	19.577
(3, 2)	19.327	19.494	(6, 5)	19.271	19.611
(3, 3)	19.349	19.543	(6, 6)	19.286	19.654

Table 9 – Forecast Models (Excluding VECMs)

	AR(1)		**ARIMA**		**SARIMA - ASI**		**SARIMA - ASIR**	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
D(ASI(-2))	-	-	-	-	0.5120	0.0019	-	-
D(ASIR(-2))	-	-	-	-	-	-	0.5688	0.0000
AR(1)	0.9808	0.0000	-0.1319	0.3166	0.2796	0.0511	0.5647	0.0002
AR(2)	-	-	-1.1391	0.0000	-0.1003	0.4753	-0.6814	0.0000
AR(3)	-	-	-0.2895	0.0691	0.1918	0.2048	0.4270	0.0086
AR(4)	-	-	0.5685	0.0008	0.5133	0.0007	0.7174	0.0000
AR(5)	-	-	-0.1812	0.0901	-0.1258	0.4590	-0.5540	0.0008
AR(6)	-	-	0.8269	0.0000	-0.1655	0.3122	0.6380	0.0001
AR(7)	-	-	-	-	-0.0967	0.5015	-0.5294	0.0004
MA(1)	-	-	0.0915	0.5570	-0.4256	0.0000	-0.7096	0.0000
MA(2)	-	-	0.6301	0.0001	-0.3548	0.0005	0.3351	0.0759
MA(3)	-	-	-	-	-0.3361	0.0001	-0.5256	0.0022
MA(4)	-	-	-	-	0.9355	0.0000	0.3350	0.0506
MA(5)	-	-	-	-	-0.3644	0.0000	-0.2560	0.1313
MA(6)	-	-	-	-	-0.3360	0.0001	-0.1848	0.2448
MA(7)	-	-	-	-	-0.2946	0.0001	0.0691	0.6706
MA(8)	-	-	-	-	0.8771	0.0000	0.5810	0.0000
Constant	134184.5	0.1316	1222.9	0.0116	1183.083	0.0219	1342.008	0.0145
Adj. R-squared	0.9314		0.8225		0.8417		0.8594	
F-statistic	1251.02	0.0000	50.8209	0.0000	28.5766	0.0000	32.7143	0.0000
AIC	20.8312		19.2318		19.2327		19.1137	
SIC	20.8857		19.4869		19.7246		19.6057	
LM Test (NR^2)*	0.1855	0.6667	7.0809	0.2147	7.5327	0.4804	7.5962	0.4739

*Breusch-Godfrey Lagrange multiplier test of null hypothesis of no serial correlation up to highest order of ARMA process.

coefficient of determination is lower, at 0.8225, than that of the AR(1) model. However, the ARIMA model's AIC and SIC are lower than those of the AR(1) model, indicating that it is superior to the AR(1) model.

5.2 Structural ARIMA Models

We estimated two structural models by adding ASI and ASIR to the optimized ARIMA model discussed above. The structural ARIMA (SARIMA) models build on the ARIMA framework by adding more AR and MA terms, and the first differences of ASI and ASIR, lagged two periods.

Table 9 shows that, in the SARIMA models, the coefficient on D(ASI(-2)) is 0.5120 with a p-value of 0.0019, while the coefficient on D(ASIR(-2)) is 0.5688 with a p-value of 0.0000. In addition, the r-squared statistics of the SARIMA-ASI and SARIMA-ASIR models are quite similar at 0.8417 and 0.8594, respectively. While lower AIC and SIC of the SARIMA-ASIR suggest that it might be the superior model, we shall see that this is not borne out by out-of-sample forecast performance.

Table 10 – Vector Error Correction Models – ASI and GDP

Cointegrating Equations

	VECM (3)	VECM (4)
D(GDP(-1))	1.0000	1.0000
D(ASI(-1))	-2.0259 **	-1.9318 **

Error Correction Equations

	VECM (3)		VECM (4)	
	D(GDP)	D(ASI)	D(GDP)	D(ASI)
CointEq1	-0.8032 **	0.1843	-0.9781 **	0.2679 *
D(GDP(-1),2)	-0.2960 *	-0.1804 *	-0.0291	-0.3835 **
D(GDP(-2),2)	-0.6872 **	-0.1767 **	-0.4338 *	-0.3751 **
D(GDP(-3),2)	-0.8296 **	-0.0711	-0.6087 **	-0.2658 **
D(GDP(-4),2)	-	-	0.1929	-0.1881 **
D(ASI(-1),2)	-1.4237 **	-0.3757	-1.5553 **	-0.3620
D(ASI(-2),2)	0.0149	0.2380	-0.1019	0.2109
D(ASI(-3),2)	0.0915	0.2332	-0.3563	0.4918 *
D(ASI(-4),2)	-	-	-0.4308	0.2229
R-squared	0.9205	0.5069	0.9227	0.556349
Adj. R-squared	0.9147	0.4708	0.9149	0.5114
F-statistic	158.2144	14.0480	117.8737	12.3835
Akaike AIC	19.2045	17.7496	19.2301	17.7024
Schwarz SC	19.4003	17.9454	19.4835	17.9558

* Significant at the 5% level. ** Significant at the 1% level.

The correlograms and Q-statistics of the residuals of the ARIMA and SARIMA models suggested white noise error terms. The Breusch-Godfrey Lagrange multiplier tests on the residuals indicate that the null hypothesis of no serial correlation, up to the highest order of ARMA process, cannot be rejected for either the univariate or the SARIMA models.

5.3 Vector Error Correction Models

Engle and Granger (1987) show that if two variables, y_{1t} and y_{2t}, are CI(1,1), then there exists a vector error correction model (VECM) governing the behavior of the variables as shown in equations (8) and (9):

$$\Delta y_{1t} = \theta_{10} + \theta_{11} z_{t-1} + \sum_{1=1}^{p1} \theta_{12,i} \Delta y_{1,t-1} + \sum_{i=1}^{p2} \theta_{13,i} \Delta y_{2,t-1} + \varepsilon_{1t} \qquad (8)$$

$$\Delta y_{2t} = \theta_{20} + \theta_{21} z_{t-1} + \sum_{1=1}^{p3} \theta_{22,i} \Delta y_{1,t-1} + \sum_{i=1}^{p4} \theta_{23,i} \Delta y_{2,t-1} + \varepsilon_{2t} \qquad (9)$$

where Δ represents the first difference of the variables, p_i are the lag lengths, and the error terms ε_{1t} and ε_{2t} are iid $(0, \sum)$. The z_{t-1} terms represent the degree to which y_{1t} and y_{2t} deviate from their equilibrium levels in the previous period, while the θ_{11} and θ_{21} are the speed of adjustment parameters.[10] According to Engle and Granger, "for a two variable system a typical error correction model would relate the change in one variable to past equilibrium errors, as well as to past changes in both variables" (p. 254).

Table 11 – Vector Error Correction Models – ASIR and GDP

Cointegrating Equations

	VECM (3)	VECM (4)
D(GDP(-1))	1.0000	1.0000
D(ASIR(-1))	-3.7488 **	-2.9483 **

Error Correction Equations

	VECM (3)		VECM (4)	
	D(GDP)	D(ASIR)	D(GDP)	D(ASIR)
CointEq1	-0.3431 **	0.2144 **	-0.4913 **	0.2260 **
D(GDP(-1),2)	-0.6354 **	-0.1879 **	-0.4820 **	-0.2469 **
D(GDP(-2),2)	-0.9038 **	-0.1457 **	-0.7714 **	-0.2018 **
D(GDP(-3),2)	-0.9278 **	-0.0694 *	-0.8361 **	-0.1295 *
D(GDP(-4),2)	-	-	0.0611	-0.0616
D(ASIR(-1),2)	-1.1800 **	0.0389	-1.3241 **	-0.1284
D(ASIR(-2),2)	0.2555	0.2601	0.1130	0.1106
D(ASIR(-3),2)	0.2397	0.0838	0.0287	0.0073
D(ASIR(-4),2)	-	-	-0.2415	-0.0904
R-squared	0.9126	0.4263	0.9137	0.4393
Adj. R-squared	0.9062	0.3843	0.9049	0.3825
F-statistic	142.7476	10.1551	104.4866	7.7360
Akaike AIC	19.2988	17.5270	19.3408	17.5626
Schwarz SC	19.4945	17.7227	19.5942	17.8160

* Significant at the 5% level. ** Significant at the 1% level.

Given that ASI, ASIR and GDP were found to be I(1) and cointegrated, VECMs[11] were indicated. VECMs with lags lengths of 3 and 4 were estimated using both ASI and ASIR[12]. Table 10 shows the VECMs using ASI and GDP while table 10 shows the VECMs with ASIR and GDP. The coefficients in the 3-lag and 4-lag specifications are quite similar but we chose the 3-lag specifications for forecast performance testing due to their smaller information criteria statistics.

An area in which the VECMs with ASI and GDP differ significantly from those with ASIR and GDP is the estimated speed of adjustment parameters, the CointEq1 coefficients in tables 10 and 11. The coefficients are -

[10] Dolado, Gonzalo and Marmol (2003) claim that the requirement that at least one of the speed of adjustment parameters is nonzero implies "the existence of Granger causality in cointegrated systems in at least one direction" p. 638.

[11] This is a restricted version of the Vector Autogression (VAR) models described in Sims (1980) and Lutkepohl (1991), with the cointegrating equation as the restriction.

[12] The SIC suggested a lag length of four while other criteria, including the LR, FPE, AIC and HQ, suggested a lag length of twelve. In the interest of parsimony, we estimated 3-lag and 4-lag VECMs.

.8032 and -.9781 for the 3-lag and 4-lag VECMs using ASI and GDP, respectively. For the VECMs run with ASIR and GDP, the coefficients are -.3431 and -.4914 on the 3-lag and 4-lag specifications, respectively. This suggests that the speed of adjustment from deviations from long-run equilibrium in the models with nominal stock prices is approximately double that of the models with real stock prices. This property may make the models with ASI more suitable for short to medium term forecasting than the models with ASIR.

5.4 Performance of Forecast Models

We use progressively longer horizons to gauge the out-of-sample performance of the six models—AR(1), ARIMA, SARIMA-ASI, SARIMA-ASIR, 3-Lag VECM-ASI and 4-Lag VECM-ASIR. The horizons are 2007Q3-2007Q4 (two periods), 2007Q3-2008Q2 (four periods), and 2007Q3-2008Q4 (six periods). Thus, we hope to capture the short to medium term out-of-sample performance of the forecast models.

Table 12 shows the performance statistics, including the Root Mean Squared Error (RMSE), Mean Absolute Percent Error (MAPE) and Theil Inequality coefficients (TIC)[13].

Table 12 – Model Performance by Forecast Horizon

	Two Quarters (2007Q3 - 2007Q4)	**Four Quarters** (2007Q3 - 2008Q2)	**Six Quarters** (2007Q3 - 2008Q4)
AR(1) Model			
Root Mean Squared Error	35544.6	25456.2	35091.5
Mean Abs. Percent Error	19.7459	11.4095	16.0499
Theil Inequality Coefficient	0.1109	0.0833	0.1113
ARIMA Model			
Root Mean Squared Error	8785.7	6653.1	14131.6
Mean Abs. Percent Error	4.8437	3.5666	6.2209
Theil Inequality Coefficient	0.0253	0.0206	0.0421
Structural ARIMA (ASI) Model			
Root Mean Squared Error	2772.5	4107.4	6117.7
Mean Abs. Percent Error	1.5528	2.3998	2.8043
Theil Inequality Coefficient	0.0079	0.0126	0.0178
Structural ARIMA (ASIR) Model			
Root Mean Squared Error	7372.7	9376.7	14122.7
Mean Abs. Percent Error	4.1479	5.7757	7.2867
Theil Inequality Coefficient	0.0212	0.0286	0.0615
3-Lag VECM - ASI			
Root Mean Squared Error	23339.2	25633.7	37525.1
Mean Abs. Percent Error	11.4109	13.6297	16.3664
Theil Inequality Coefficient	0.0616	0.0729	0.0981
3-Lag VECM - ASIR			
Root Mean Squared Error	13549.8	12459.4	16859.9
Mean Abs. Percent Error	7.0370	6.9352	8.2257
Theil Inequality Coefficient	0.0367	0.0368	0.0465

The performance statistics are computed thus:

[13] The TIC, which lies between zero and one, is computed as the sum of the forecast error variance divided by the sum of a naïve forecast variance, where the naïve forecast is the previous period's value of the forecast object (this could be a random walk model). A value of zero indicates a perfect fit for the forecast model while a value of one indicates that the model is not better than the naïve forecast. The bias, variance and covariance proportions decompose the forecast error into the distance between the mean of the forecast compared the mean of the forecast object, the distance between the variation of the forecast compared to that of the forecast object, and the remaining unsystematic error, respectively. A "good" forecast would have a low TIC and a higher covariance proportion than bias or variance proportions. See Theil (1966), Armstrong and Fildes (1995) and Diebold (2007).

$$\text{RMSE} = \sqrt{\sum_{t=T+1}^{T+h} (\hat{\text{GDP}}_t - \text{GDP}_t)^2 \, / \, h} \tag{10}$$

$$\text{MAPE} = 100 \sum_{t=T+1}^{T+h} \left| \frac{\hat{\text{GDP}}_t - \text{GDP}_t}{\text{GDP}_t} \right| / \, h \tag{11}$$

$$\text{TIC} = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{\text{GDP}}_t - \text{GDP}_t)^2 \, / \, h}}{\sqrt{\sum_{t=T+1}^{T+h} \hat{\text{GDP}}_t^{\,2}} + \sqrt{\sum_{t=T+1}^{T+h} \text{GDP}_t^{\,2}}} \tag{12}$$

where the forecast sample is T + h, with h (the forecast horizon) taking the values of 2, 4 and 6, and the forecast and actual values in period t are GDP_t – hat and GDP_t, respectively.

Table 13 – Ranking of Forecast Models*

Rank	**Two Quarters** (2007Q3 - 2007Q4)	**Four Quarters** (2007Q3 - 2008Q2)	**Six Quarters** (2007Q3 - 2008Q4)
1	SARIMA-ASI	SARIMA-ASI	SARIMA-ASI
2	SARIMA-ASIR	ARIMA	ARIMA
3	ARIMA	SARIMA-ASIR	SARIMA-ASIR
4	VECM-ASIR	VECM-ASIR	VECM-ASIR
5	VECM-ASI	AR(1)	AR(1)
6	AR(1)	VECM-ASI	VECM-ASI

* Using mean absolute percentage error (MAPE) as the ranking criterion.

Within the two-quarter horizon, the ARIMA model improves substantially on the performance of the AR(1) model, with a MAPE of 4.84 percent versus 19.75 percent for the AR(1) model. The SARIMA-ASI model performs better than either the AR(1) or the ARIMA model, with a MAPE of 1.55 percent; this is 92.14 percent and 67.94 percent lower than the MAPEs of the AR(1) and ARIMA models, respectively. The SARIMA-ASIR model, with a MAPE of 4.14 percent, performed better than the AR(1) and ARIMA models, but not as well as the SARIMA-ASI model. The VECMs outperformed the AR(1) model but had higher error rates than the ARIMA and SARIMA models.

Over four quarters, the ARIMA model, with a MAPE of 3.57 percent outperforms the AR(1) model which has a MAPE of 11.41 percent. However, the SARIMA-ASI model outperforms both models with a MAPE of 2.39 percent. The SARIMA-ASIR model outperformed the AR(1) model and VECMs but had higher error rates than the ARIMA and SARIMA-ASI models. With a MAPE of 6.94 percent, the VECM-ASIR outperformed the AR(1) model but the VECM-ASI had a MAPE of 13.63 percent versus the AR(1) model’s 11.41 percent.

The results over a six-quarter horizon mirror those for the four-quarter; the SARIMA-ASI model has the lowest MAPE of 2.80 percent, followed by the ARIMA model’s 6.22 percent, the SARIMA-ASIR model’s 7.29 percent, the VECM-ASIR’s 8.23 percent, the AR(1) model’s 16.05 percent and the VECM-ASI’s 16.37 percent.

We summarize the out-of-sample forecast performance of the models in table 13, which ranks the models by MAPE. Regardless of the forecast horizon, the SARIMA-ASI model consistently outperforms the other five models. In addition, the VECM-ASI model ranked fifth over the two-quarter horizon and last over the other horizons. The results suggest that stock market prices contain information that could be used to improve GDP forecasts in the short- to medium term and that a structural ARIMA specification with the nominal stock index is likely to perform better than an ARIMA specification with a deflated stock index or a VECM with either the nominal or real stock index.

6.0 Conclusions and Policy Implications

The goal of this paper was to determine whether or not stock prices contained information which could be used to improve predictions of economic activity in Nigeria. Granger causality tests indicated that the All Share Index is a leading indicator of real GDP but had no relationship with the Index of Industrial Production. In addition, no causality was found between GDP and IIP. Johansen cointegration tests also suggested a long-run equilibrium relationship between nominal and real stock prices and real GDP in Nigeria.

The finding of bi-directional causality between stock prices and GDP is not surprising in light of the fact that, while stock prices reflect the expectations of investors, they ultimately must also reflect economic fundamentals. A high rate of economic growth will lead to an increase in firms' earnings and higher earnings will buoy stock prices. Thus, there is evidence that the stock market in Nigeria is not only a leading indicator of the real economy but that Nigerian stock prices are, at least partly, based on economic fundamentals. Other studies, including Pilinkus (2009), have found bi-directional causality between stock prices and economic activity.

Figure 2 shows average price-earnings (PE) ratios of Nigerian stocks between January 2001 and December 2009. Nigerian stocks seemed to have become decoupled from fundamentals during the boom that began around January 2007; the average PE ratio reached an all-time high of 48.9 in February 2008 before the ensuing crash. However, by December 2009, the average PE ratio had fallen to 19.30, which is quite close to the nine-year average of 18.37. As such, the evidence suggests that, while the Nigerian stock market is not immune to bubbles, it is, to a large extent and in the long run, governed by economic fundamentals.

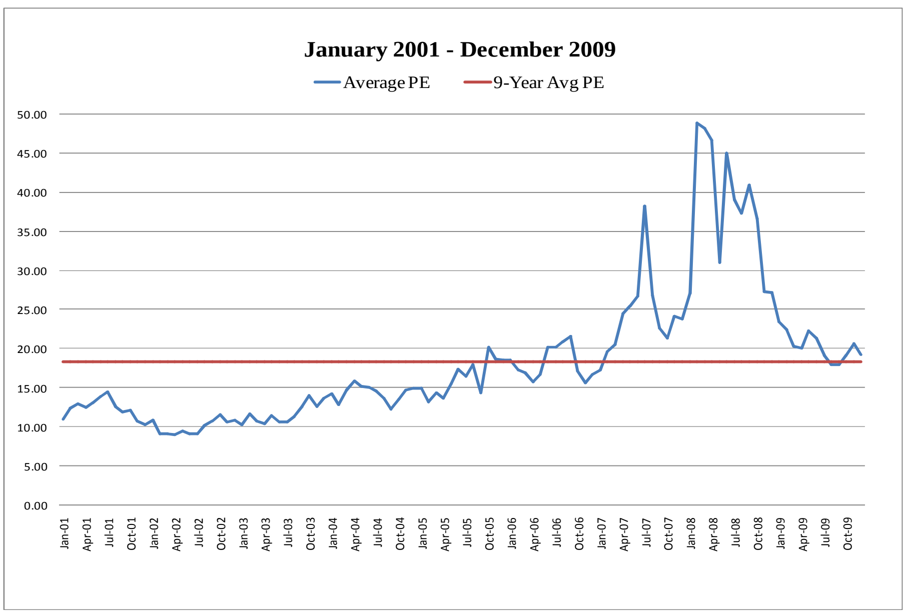


Figure 2 – Nigerian Stock Exchange Price-Earnings Ratios

The "acid test" of a leading indicator is its ability to improve the performance of forecasts of GDP or other macroeconomic variables of interest. Tests conducted with short to medium term forecast horizons show that the information in stock prices can reduce forecast errors by up to 92 percent compared to an AR(1) model and up to 68 percent compared to an ARIMA model. Deflating the All-Share Index using the CPI did not improve the performance of the models. Also, VECMs performed poorly in comparison to models based on an ARIMA framework.

The evidence presented in this paper suggests that the All Share Index should be added, in nominal form, to a composite index of leading economic indicators (CILEI) for Nigeria, with a two-quarter lag[14]. This is likely to improve the accuracy of the composite index of leading economic indicators. Other financial variables should also be evaluated for inclusion in the CILEI since they embody expectations of economic agents in the same manner that the ASI does. A leading candidate among financial variables is the Treasury bond yield curve, as operationalized by the spread between a benchmark long maturity bond (e.g., the 10-year federal government bond) and a short maturity security (e.g., the three month government bill). Estrella and Mishkin (1996) show that

[14] Most financial variables in composite indices of leading indicators are incorporated in nominal form.

the yield spread outperforms most other macroeconomic variables in predicting U.S. recessions two to six quarters ahead. The Federal Reserve Bank of New York has documented the reliability of the slope of the yield curve as a leading indicator of economic activity in the U.S.[15]

The methodology utilized in this paper could be replicated in order to investigate the information content of the yield curve in Nigeria. The addition of the stock index and yield curve to the CILEI is in keeping with international best practice, as several nations, including the U.S., UK, Japan and South Africa have both financial variables in their composite indices of leading economic indicators.

In addition to leading indices, other approaches could be explored in order to improve GDP forecasts. Further research could investigate the efficacy of using monetary aggregates, credit to the private sector, oil revenues, rainfall statistics and surveys of economists to improve predictions of the future path of economic activity. More accurate forecasts of economic activity will enhance our ability to manage the economy via monetary and fiscal policies.

References

Adebiyi, M. A., (2005), "Capital Market Performance and Nigerian Economic Growth" in *Issues in Money, Finance and Economic Management in Nigeria*, Essays in Honor of Professor Obasanmi Olakanpo, Edited by Fakiyesi, O. and O. Akano. Unilag Press Section 1, Chapter 5, pp. 146-176.

Akinlo, A.E., & D. O. Akinlo (2009), "Stock Market Development and Economic Growth: Evidence from Seven Sub-Saharan African Countries," *Journal of Economics and Business*, Vol. 61, No. 2, pp. 162-171.

Armstrong, J.S. and R. Fildes (1995), "On the Selection of Error Measures for Comparisons among Forecasting Methods," *Journal of Forecasting*, Vol. 14, pp. 67-71.

Atmadja, A.S. (2005), "Granger Causality Tests for Five ASEAN Countries' Stock Markets and Macroeconomic Variables During and Post the Asian Financial Crisis," *Jurnal Manajemen & Kuwirausahaan*, Vol. 7, No. 1, pp. 1 -21.

Bahadur, S. and S. Neupane (2006), "Stock Market and Economic Development: A Causality Test," The Journal of Nepalese Business Studies, Vol. 3, No. 1, pp. 36-44.

Baker, Malcolm, and Jeffrey Wurgler (2000), "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance*, Vol. 55, pp. 2219-2257.

Bernanke, B., M. Gertler and S. Gilchrist (1996), "The Financial Accelerator and the Flight to Quality," *The Review of Economics and Statistics*, Vol. 78, No. 1, pp. 1-15.

Bierens, H. (2003), "Unit Roots," in Baltagi, B., (editor), *A Companion to Theoretical Econometrics*, Blackwell Publishing, 2003.

Bostic, R., S. Gabriel and G. Painter (2009), "Housing Wealth, Financial Wealth, and Consumption: New Evidence from Micro Data," *Regional Science and Urban Economics*, Vol. 39, pp. 79-89.

Box, George E.P. and Gwilyn M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, Revised Edition, Oakland, CA: Holden-Day.

Brigham, E. F. and J.F. Houston, *Fundamentals of Financial Management*, Thomson-Southwestern, Eleventh Edition, 2007.

Burgstaller, J. (2002), "Are Stock Returns a Leading Indicator for Real Macroeconomic Developments?" Jonaness Kepler University of Linz, working paper no. 0207.

Burns, A. F. and W. C. Mitchell (1946), *Measuring Business Cycles*. New York: National Bureau of Economic Research.

[15] See http://www.newyorkfed.org/research/capital_markets/ycfaq.html

Campbell, H.R. (1989), "Forecasts of Economic Growth from the Bond and Stock Markets." *Financial Analysts Journal*, September/October, pp. 38-45.

Case, K. E., J. M. Quigley, and R. J. Shiller (2005), "Comparing Wealth Effects: The Stock Market versus the Housing Market," *Advances in Macroeconomics*: Vol. 5 Iss. 1, Article 1.

Central Bank of Nigeria, *Statistical Bulletins*, 2007 - 2008.

Choi, J., S. Hauser and K. Kopecky (1999), "Does the Stock Market Predict Real Activity? Times Series Evidence from the G-7 Countries," *Journal of Banking & Finance*, Vol. 23, pp. 1771-1792.

Christoffersen, P.F. and T. M. Sløk (2000), "Do Asset Prices in Transition Countries Contain Information about Future Economic Activity*?" IMF Working Paper No. 00/103*.

Comincioli, B. (1996), "The stock market as a leading indicator: An application of Granger causality," *The University Avenue Undergraduate Journal of Economics*, Sample Issue.

Conference Board (2009), "The Conference Board Leading Economic Index for the United States," http://www.conference-board.org/pdf_free/economics/bci/stillsnow.pdf

Dickey, D.A. (1976), *Estimation and Hypothesis Testing in Nonstationary Time Series*, Ph.D. dissertation, Iowa State University.

Dickey, D.A. and W.A. Fuller (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, Vol. 49, pp. 1057-1072.

Diebold, F.X. (2007), *Elements of Forecasting*, fourth edition, Thomson Southwestern.

Dolado, J., J. Gonzalo, and F. Marmol (2003), "Cointegration," in Baltagi, B., (editor), *A Companion to Theoretical Econometrics*, Blackwell Publishing, 2003.

Engle, R. F. and C.W. Granger (1987), "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, Vol. 55, pp. 251-276.

Estrella, A. and F. S. Mishkin (1996), "The Yield Curve as a Predictor of U.S. Recessions," *Current Issues in Economics and Finance*, Federal Reserve Bank of New York, Vol. 2, pp. 1-6.

Fama, E. F. (1981), "Stock Returns, Real Activity, Inflation, and Money," *The American Economic Review*, Vol. 71, No. 4, pp. 545-565.

Fama, E. F. (1990), "Stock Returns, Expected Returns, and Real Activity," Journal of Finance, Vol. 45, pp. 1089-1108.

Fazzari S., R. Hubbard and B. Peterson (1988), "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, pp. 141-195.

Federal Reserve Bank of New York (2010) *The Yield Curve as a Leading Indicator*, http://www.newyorkfed.org/research/capital_markets/ycfaq.html

Foresti, P. (2007), "Testing for Granger Causality Between Stock Prices and Economic Growth," MPRA Paper No. 2962.

Fuller, W.A. (1976), *Introduction to Statistical Time Series*. New York: Wiley.

Gan, C., M. Lee, H. Yong and J. Zang (2006), "Macroeconomic Variables and Stock Market Interactions: New Zealand Evidence," Investment Management and Financial Innovations, Vol. 3, pp. 89-101.

Gordon, M. J. (1959), "Dividends, Earnings and Stock Prices," *Review of Economics and Statistics*, Vol.41, pp. 99–105.

Granger, C. J. (1969), "Investigating Causal Relationships by Econometrics Models and Cross Spectral Methods," *Econometrica*, Vol. 37, pp. 425-435.

Hirshleifer, D. (2001), “Investor Psychology and Asset Pricing,” *Journal of Finance* Vol. 56, pp.1533- 1597.

Huang, R.D. and W. A. Kracaw (1984), “Stock Market Returns and Real Activity: A Note,” *The Journal of Finance*, Vol. 39, No. 1, pp. 267-273.

Husain, F. and T. Mahmood (2001), “The Stock Market and the Economy in Pakistan,” *The Pakistan Development Review*, Vol. 40, pp. 107–114.

Ikoku, A. E. and C. T. Okany (2010), “Can Price-Earnings Ratios Predict Stock Prices? Evidence from the Nigerian Equity Market,” Monetary Policy Department Working Paper, Central Bank of Nigeria, July 2010.

Jefferis, K., C. Okeahalam and T. Matome (2001), “*International Stock Market Linkages in Southern Africa*,” AERC Research Paper 105.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, Vol. 59, pp. 1551-1580.

Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.

Johansen, S. and K. Jeselius (1990), “Maximum Likelihood Estimation and Inference on Cointegration-with Applications to the Demand for Money,” *Oxford Bulletin of Economics and Statistics*, Vol. 52, pp. 169-210.

Lee, Bong-Soo (1992), “Causal Relations Among Stock Returns, Interest Rates, Real Activity and Inflation,” *The Journal of Finance*, Vol. 47, No. 4, pp.1591-1603.

Leigh, L. (1997), “Stock market equilibrium and macroeconomic fundamentals,” IMF Working Paper WP/97/15.

Lutkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag.

MacKinnon, J. G. (1996), "Numerical Distribution Functions for Unit root and Cointegration tests,” *Journal of Applied Econometrics*, Vol. 11, pp. 601–618.

MacKinnon, James G., Alfred A. Haug, and Leo Michelis (1999), "Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration," *Journal of Applied Econometrics*, 14, 563-577.

Mauro, P. (2003), “Stock Returns and Output Growth in Emerging and Advanced Economies,” *Journal of Development Economics*, Vol.71, No.1, pp. 129-153.

Mitchell, W.C. and A. F. Burns (1938), *Statistical Indicators of Cyclical Revivals*, NBER Bulletin 69, NY. Reprinted in Business Cycle Indicators. G.H. Moore, ed. 1961. Princeton: Princeton U. Press.

Moolman, E. and J. Jordaan (2005), “Can Leading Business Cycle Indicators Predict the Direction of the South African Commercial Share Price index,” *South African Journal of Economics*, Vol. 73, No. 1, pp. 68-78.

Mun, H.W., E. C. Siong and T.C. Thing (2008), “Stock Market and Economic Growth in Malaysia: Causality Test,” *Asian Social Science*, Vol. 4, No. 4, pp.86-92.

Narudeen, A. (2009), “Does Stock Market Development Raise Economic Growth? Evidence from Nigeria,” *The Review of Finance and Banking*, Vol. 1, pp. 15-26.

Nishat, M. and R. Shaheen (2004), “Macroeconomic Factors and the Pakistani Equity Market,” *Pakistan Development Review*.

Osinubi, T. S. (2004), “Does Stock Market Promote Economic Growth in Nigeria?” *The ICFAI Journal of Applied Finance*, IJAF Volume 10, No. 3, pp. 17-35.

Otoo, M.W. (1999), “Consumer Sentiment and the Stock Market,” Board of Governors of the Federal Reserve System paper.

Paiella, M. (2007), “Does Wealth Affect Consumption? Evidence for Italy,” *Journal of Macroeconomics*, 29, pp. 189-205.

Pearce, Douglas K. (1983), "Stock Prices and the Economy." *Federal Reserve Bank of Kansas City Economic Review*, November, pp. 7-22.

Phillips, P.C. and P. Perron (1988), “Testing for a Unit Root in Times Series Regression,” *Biometrika*, Vol. 75, pp. 335-346.

Phillips, P.C. and S. Ouliaris (1990), “Asymptotic Properties of Residual Based Tests for Cointegration,” *Econometrica*, Vol. 58, No. 1, pp.165–193.

Pilinkus, D. (2009), “Stock Market and Macroeconomic Variables: Evidence from Lithuania,” *Economics and Management*, Vol. 14, pp. 884-891.

Padhan, P. C. (2007), “The nexus between stock market and economic activity: an empirical analysis for India,” *International Journal of Social Economics*, Vol. 34, pp. 741-753.

Ritter, J. R. (1991), “The Long-run Performance of Initial Public Offerings,” *Journal of Finance*, Vol. 46, pp. 3-27.

Schwert, G. W. (1990), "Stock Returns and Real Activity: A Century of Evidence," *The Journal of Finance*, Vol. 45, No. 4, pp. 1237-1257.

Sims, C. A. (1980), “Macroeconomics and Reality,” *Econometrica*, Vol. 48, pp. 1-46.

Simone, S. (2009), “Wealth Effect in the US: Evidence from Brand New Micro-Data,” *Working Paper, Department of Economics*, University of Siena.

Stock, J.H. (2003), “Forecasting Economic Time Series,” in Baltagi, B., (editor), *A Companion to Theoretical Econometrics*, Blackwell Publishing, 2003.

Stock, J.H. and M.W. Watson (2003a), “Forecasting Output and Inflation: The Role of Asset Prices,” *Journal of Economic Literature*, Vol. 41, No. 3, pp. 788-829 Stock, J.H. and M.W. Watson (2003b), “How Did Leading Indicator Forecasts Do During the 2001 Recession?,” National Bureau of Economic Research Working Paper.

Theil, H. (1966), *Applied Economic Forecasting*, Amsterdam: North-Holland.

Monetary and Fiscal Policy Interactions in Nigeria: An Application of a State-Space Model with Markov-Switching

Chuku A. Chuku[1]

This paper uses quarterly data to explore the monetary and fiscal policy interactions in Nigeria between 1970 and 2008. As a preliminary exercise, the paper examines the nature of fiscal policies in Nigeria using a vector autoregression (VAR) model. The simulated generalized impulse response graphs generated from the VAR estimation provides evidence of a non-Ricardian fiscal policy in Nigeria. Further, the paper analyzes the interactions between monetary and fiscal policies by applying a State-space model with Markov-switching to estimate the time-varying parameters of the relationship. The evidence indicates that monetary and fiscal policies in Nigeria have interacted in a counteractive manner for most of the sample period (1980-1994). At other periods, we do not observe any systematic pattern of interaction between the two policy variables, although, between 1998 and 2008, some form of accommodativeness can be inferred. Overall, the results suggest that the two policy regimes (counteractive and accommodative) have been weak strategic substitutes during the post 1970 (Civil War) period. For the policy maker, our results imply the existence of fiscal dominance in the interactions between monetary and fiscal policies in Nigeria, implying that inflation, predominantly results from fiscal problems, and not from lack of monetary control.

Keywords: Monetary-fiscal policy interaction; State-space models; Markov-switching, Fiscal Theory of the Price Level (FTPL).

JEL Classification: E31, E63, H5

1. Introduction

Monetary and fiscal policies are the two most important tools for managing the macroeconomy in other to achieve high employment rates, price stability and overall economic growth. An important issue that has exercised the minds of macroeconomist is the understanding of how the dependence, independence and interdependencies between monetary and fiscal policies could lead the economy closer or further away from set goals and targets.

In a poorly co-coordinated macroeconomic environment, fiscal policies might affect the chances of success of monetary policies in various ways, such as: its eroding impact on the general confidence and efficacy of monetary policy, through its short-run effects on aggregate demand, and by modifying the long-term conditions for economic growth and low inflation. On the other hand, monetary policies may be accommodative or counteractive to fiscal policies, depending on the prevailing political and economic paradigms.

After the prosecution of the Nigerian Civil war in 1970, diverse monetary and fiscal policies measures were employed to reconstruct the economy and to put it on a sustainable growth trajectory. These efforts may have been bolstered or undermined by the nature of the interactions between monetary and fiscal policies in Nigeria. This paper hypothesizes that the interactions between monetary and fiscal policies in Nigeria, have been characterized by regime-shifts, which can be demarcated into two phases of accommodative and counteractive policies.

The objective of the paper is therefore, to examine the hypothesis of regime-shifts in the interactions between monetary and fiscal policies in Nigeria during the Post Civil War era (1970-2008). To that end, we employ a state-space (Ss) model with Markov-switching (Ms) properties to examine this behaviour. This exercise is justified because to the best of my knowledge, it does not only pioneer the application of the Ss-Ms model for the analysis of policy interactions in Nigeria, it inherently provides insights about the validity or otherwise of the fiscal theory of the price level in Nigeria.

[1] Department of Economics, University of Uyo, P.M.B. 1017, Uyo, Nigeria. chukuachuku@gmail.com; +2348067247177

The rest of the paper is organized as follows. In Section 2, the paper discusses the issues in the literature and theory of monetary-fiscal policy interactions. Section 3 examines the preliminary evidence on the fiscal theory of price level determination in Nigeria. Section 4 specifies the State-space model with Markov-switching properties. The Kalman algorithm for the one-step ahead forecast is also described. Section 5 presents the results and the synthesis from the results, while Section 6 contains the conclusion.

2. Issues on Monetary-Fiscal Policy Interactions

Numerous studies have examined the interactions between monetary and fiscal policies (see for example Semmler and Zhang, 2003; Fialho and Portugal, 2009; Sargent, 1999 and Leith and Wren-Lewis, 2000). Most of these studies have focused on three basic issues (theoretical and empirical) on the interactions between monetary and fiscal policies. These issues include: the fiscal theory of price level determination, strategic interaction, and time-varying regime changes in policy interactions. The major issue that has been prominent in most of these studies is the issue of the "fiscal theory of price level determination" (FTPL). The FTPL has been studied by Leeper (1991), Sims (1994, 1997 and 2001), Woodford (1994, 1995 and 2000), Semmler and Zhang (2003), among others. These studies seek to analyze the "non-Ricardian" fiscal policy, which specifies the time paths of government's debt, expenditure and taxes, without considering the government's intertemporal solvency, such that, in equilibrium, the price level has to adjust to ensure solvency (Semmler and Zhang, 2003).

The introduction of the non-Ricardian fiscal policy into a standard New-Keynesian monetary sticky price model alters the stability conditions associated with the central bank's interest rate policy[2]. The process through which this occurs is simple. First, fiscal policies affect the equilibrium price-level. An increase in the price level reduces the real value of the net assets of the private sector or, equivalently, the net government liability. The reduction of private sector wealth reduces private-sector demand for goods and services through direct wealth effect. As a result, there will be only one price level that results in aggregate demand that equals aggregate supply. Changes in expectations regarding future government budget also have similar wealth effects that require an off-setting change in the price level in order to maintain equilibrium.

Under this non-Ricardian fiscal policy, one thus arrives at a theory of price-level determination in which fiscal policy plays the crucial role, because the effects of price-level changes on aggregate demand depends on the size of the government budget and also due to the off-setting wealth effects of expected future government debt (Semmler and Zhang, 2003).

Following Woodford (1995), the fiscal theory of price level determination can be presented thus: Let P_t denote the price level at time t, W_t the nominal value of beginning-of-period wealth, g_t government expenditure in period *t*, T_t the nominal value of net taxes paid in period *t*, R_t^b the gross nominal return on bonds held from period *t*, to $t+1$ and R_t^m the gross nominal return on the monetary base. Other variables are defined thus:

$$\tau_t = T_t/P_t \ (real\ tax)$$
$$\Delta_t = (R_t^b - R_t^m)/R_t^b \ ('price'\ of\ holding\ money)$$
$$r_t = \left(\frac{P_t}{P_{t+1}}\right) - 1 \ (real\ rate\ of\ return\ on\ bonds)$$
$$m_t = M_t/P_t \ (real\ balances)$$

Under this circumstances the equilibrium condition that determines the price level P_t at time *t*, given the predetermined nominal value of net government liabilities W_t, and the expectations at date *t*, regarding the current and future values of real quantities and relative prices can be expressed as:

[2] Benhabib et al. (2001) demonstrate the conditions under which interest rate feedback rules that are used to set the nominal interest rate as an increasing function of the inflation rate induces aggregate instability. They find that these conditions are partly affected by the monetary-fiscal policy regime emphasized in the fiscal theory of the price level.

$$\frac{W_t}{P_t} = \sum_{s=t}^{\infty} \frac{(\tau_s - g_s) + \Delta_s \mathrm{m}_s}{\prod_{j=t}^{s-1}(1 + r_j^b)} \quad (1)$$

Woodford (1995) explains the mechanisms by which the price level adjusts to satisfy the equilibrium condition in Equation (1) under assumptions of long-run price flexibility[3]. The mechanism is such that an increase in the nominal value of outstanding government liabilities or size of real government budget deficit expected at some future date is inconsistent with equilibrium at the existing price level. Either change causes households to believe that their budget set has expanded, and so, they demand additional consumption immediately. The consequence will be an excess demand for goods, and price level will therefore increase, to the extent that the capital loss to the value of private-sector assets restores household's estimates of their wealth to ones that just allow them to purchase the quantity of goods that the economy can supply. Woodford (1995) emphasized that in the special case of the "Ricardian" policy regime, the fiscal mechanism described above, fails to play any role in the price level determination.

An examination of the monetary-fiscal policy interactions within the FTPL framework is essential for a country like Nigeria, where government's fiscal deficits as a ratio of GDP have largely been significant, averaging around -3.89 and government's debt as a proportion of GDP has fluctuated between 9 and 41% from 1970 to 2008[4]. These significant ratios, intuitively suggests that fiscal policies may have a significant influence on the price level in Nigeria. Overall, the principle of the fiscal theory of the price level (FTPL) implies that unless specific measures are taken to implement a coordinated fiscal policy, the objective of price stability may not be achieved even with a committed, independent and "non-discretionary" monetary policy regime.

Despite its popularity and general acceptability, the FTPL has come under intense criticisms on the theoretical and empirical formulations. Buiter (2001), Semmler and Zhang (2003), and Canzoneri et al. (2000) provide some detailed criticism on the FTPL. Another prominent issue in the literature on the monetary-fiscal policy relations is the analysis of the "strategic-interactions" between monetary and fiscal policies. Some examples of studies that explore the strategic interactions between monetary and fiscal policies include Cantenaro (2000), van Anarle et al. (2002) and Wyplosz (1999). The work by van Anarle et al. (2002) was particularly interesting because they considered the interactions between monetary and fiscal authorities in a differential game framework. They derived explicit solutions for the dynamics of fiscal deficit, inflation and government debt in a cooperative and Nash open-loop equilibrium framework. From their results, they identified three alternative policy interactions: (1) non-cooperative monetary and fiscal policies, (2) partial-cooperation and (3) full-cooperation; both in the symmetric and asymmetric settings.

Although the work by van Anarle et al. (2002) and most other works on monetary –fiscal policy interactions are theoretical, recent studies on this relationship have been empirical. For example, Fialho and Portugal (2009) studies the interactions between monetary and fiscal policies in Brazil using a Markov-switching vector autoregression model. By applying the fiscal theory of the price level, they propose that there is a relationship between public debts (a measure for fiscal policy) and Selic (their measure for monetary policy). They also assume the existence of two regimes and possibility for switching between the two. From their results, they conclude that the nature of macroeconomic coordination between monetary and fiscal policies in Brazil follows a "substitution- approach", throughout the period of the study, with a dominant monetary regime, in opposition to the non-Ricardian policies of the fiscal theory of the price level.

Another fascinating empirical study is the one by Canzoneri et al. (2000) who studies the fiscal regime of the U.S with VAR models, arguing that Ricardian regimes are as empirically plausible as non-Ricardian regimes, and provide interpretations of certain aspects of monetary and fiscal policy interactions. Melitz (1997) uses pooled data for 15 member states of the European Union (EU) to undertake some estimation, and find that coordinated macroeconomic policies are in practice in the region. Specifically, they conclude that "easy-fiscal" policy leads to "tight-monetary" policy and "easy-monetary" policy, to "tight-fiscal" policy.

[3] This assumption may not always be plausible because within the Keynesian framework, prices may be sticky in the short-run.
[4] Computed by author with data sourced from CBN Statistical Bulletin, Anniversary Edition.

In a very influential paper, Muscatelli et al. (2002) estimated VAR models with both constant and time varying parameters for G7 countries and found that monetary and fiscal policies where used as strategic complements, and that the strategic interdependence between monetary and fiscal policies can be captured using Bayesian VAR models. The finding and recommendation by Muscatelli et al. (2002) influence the study by Semmler and Zhang (2003) that use both a VAR and a State-space model with Markov-switching to analyze the interactions between monetary and fiscal policies in the Euro-Area. Their results reveal that there exist some regime changes in the monetary and fiscal policy interactions in France and Germany.

The approach that is adopted in this work is substantially influenced by the recommendations that emerged from the findings of Muscatelli et al. (2002). That is, the interdependence between monetary and fiscal policies can be adequately captured in a Bayesian VAR model with Markov-switching characteristics. The approach we adopt also draws from the "State-space" refinement introduced by Semmler and Zhang (2003). Thus, this paper analyzes the monetary-fiscal policy interactions in Nigeria, using a State-space Markov-switching VAR model.

3. Preliminary Evidence from Nigeria

Before analyzing the hypothesized regime switching nature of the interactions between monetary and fiscal policies in Nigeria, we first undertake some preliminary empirical research on the nature of fiscal policies in Nigeria, using a simple VAR framework. The rationale behind our preliminary investigation is to test whether the fiscal regime in Nigeria has followed the "Ricardian" or "non-Ricardian" approach, to enable us ascertain whether the assumptions for the fiscal theory of price level determination are valid or invalid for Nigeria. The approach we adopt is in the spirit of Canzoneri et al. (2000) and Semmler and Zhang (2003). Thus, we exaime the interaction between two fiscal variables: fiscal balance and government liabilities. Government liabilities are measured by the Federal Government's domestic debt outstanding, and the fiscal balance is the overall surplus or deficit of government finances. We scale the two variables by dividing with nominal GDP. All the data sets are compiled from the CBN Statistical Bulletin, Special Anniversary Edition and are converted to quarterly frequencies by means of the cubic spline technique (see Lisman and Sandee, 1964 and Denton, 1971 for a description of this frequency conversion technique). Figure 1 plots the scatter between fiscal balance to GDP and government liability to GDP in Nigeria.

The Figure indicates that there exist a negative correlation between fiscal balances and government liability in Nigeria, with the correlation coefficient being -0.672. This suggests that net borrowing does not decrease when the fiscal balance decreases. Rather, it increases when the fiscal balance decreases. This observed relationship suggests the existence of "non-Ricardian" fiscal policy in Nigeria.

Further, we undertake VAR estimation for the two variables. The VAR model with order (*k*) is presented thus:

$$\boldsymbol{Y}_t = \boldsymbol{C_0} + \sum_{\boldsymbol{i=1}}^{k} \Phi_{\boldsymbol{i}} \boldsymbol{Y}_{t-i} + \boldsymbol{\epsilon}_t \qquad (3.1).$$

Where $\boldsymbol{Y}_t = (Y_{1t}, Y_{2t,}\ )'$ is a *2*1* vector of endogenous variables, i.e., fiscal balance to GDP ratio (*FSB*), and government liability to GDP ratio (*GL*), while $\boldsymbol{Y}_{t-i}$ is the corresponding lag term for order *i*. $\Phi_{\boldsymbol{i}}$, is an *n* n* matrix of autoregressive coefficients, for *i = 1, 2, ...k*. $\boldsymbol{C_0} = (C_1, C_2\,)'$ is the *C* intercept vector of the VAR model. $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \epsilon_{2t,})'$ is an *n*1* vector of white noise processes. *K* is the number of lagged terms. VAR estimations are very sensitive to lag structure of variables. Using a sufficient lag length may help to reflect the long-term impact of variables on others. However, including longer lag lengths will lead to multicollinerarity problems and will increase the degrees of freedom (DOF). Empirical simulations show that for any $K \geq 11$, the model will become divergent with at least one autoregressive root that is greater than one. According to sequential modified Likelihood Ratio test statistic (LR), lag orders between 1 and 3 are recommended for models of this nature (Wooldridge, 2006). Here, we use lag order 2, determined by the Hannan-Quinn information criterion.

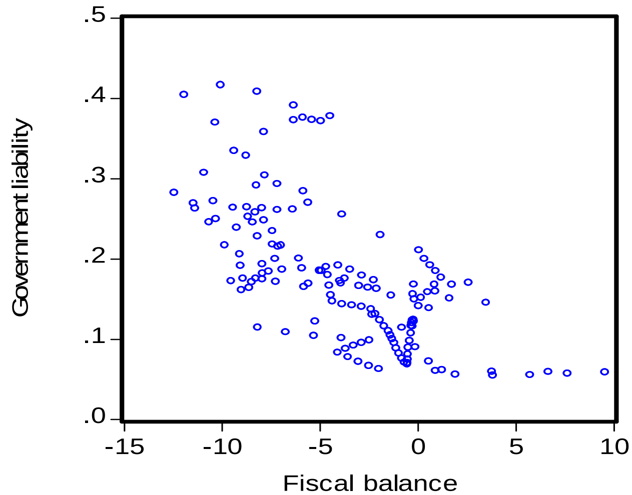


Figure 1 Scatter of government liability/GDP and fiscal balance/GDP.

Before undertaking the VAR estimation, we test for stationarity of the variables, using the ADF unit root test[5]. The results indicate that the variables are stationary at their first-differences. Hence, we use the first differences of the fiscal balance and government liability series in the VAR estimation. With two lags of the variables (determined by the Hannan-Quinn information criterion), the results obtained from the estimation are thus:

$$\underset{(0.599)}{\Delta FSB = 0.098} + \underset{(27.89)}{1.601\Delta FSB_{t-1}} - \underset{(-12.46)}{0.717\Delta FSB_{t-2}} - \underset{(-1.58)}{8.495\Delta GL_{t-1}} + \underset{(1.027)}{5.389\Delta GL_{t-2}}$$

$$\underset{(3.660)}{\Delta GL = 0.004} - \underset{(-0.470)}{2.09 * 10^{-4}\Delta FSB_{t-1}} - \underset{(-0.385)}{1.72 * 10^{-4}\Delta FSB_{t-2}} + \underset{(44.19)}{1.831\Delta GL_{t-1}} - \underset{(-21.37)}{0.866\Delta GL_{t-2}}$$

$$R^2 = 0.96 \qquad Log\ Likelihood = -186.79$$

Where ΔFSB and ΔGL denotes the first difference of fiscal balance/GDP and government liability/GDP respectively, and the values in parenthesis are the *t*-values. The results from the VAR estimation lend credence to the negative relationship observed in the scatter diagram plotted in Figure 1. Following this estimation, we simulate the impulse responses for the two variables, and present then in Figure 2. The impulse response graphs indicate that one-standard deviation innovation in ΔFSB causes a negative response in ΔGL (see Figure 2, Panel C), and similarly, one S.D innovation in ΔGL also induces negative some kind of negative response in ΔFSB (see Panel B). This relationship provides preliminary evidence of the existence of the non-Ricardian fiscal regime in Nigeria.

4. Model Specification

We draw from Muscatelli et al. (2002) and Semmler and Zhang (2003) by specifying a State-Space (SS) model with Markov-Switching (MS) characteristics. The reason for applying this model is to enable us test the hypothesis of regime changes (accommodative and counteractive) and the nature of the interactions (i.e., substitutes or complements) between monetary and fiscal policies in Nigeria, and if yes, to find out how they may have interacted, i.e., as substitutes or complements. The peculiar advantage of the SS-MS model is in the fact that it allows us to take into account multiple structural breaks in a given time series, and to explain non-linearities in the data. Though powerful, the SS-MS model is restrictive, because it only permits the existence of two time-regimes (Maddala and Kim, 1998). This limitation does not undermine the objective of our work, since we hypothesize that monetary-fiscal policies in Nigeria can be categorized into accommodative or counteractive regimes.

---

[5] See Appendix for the results

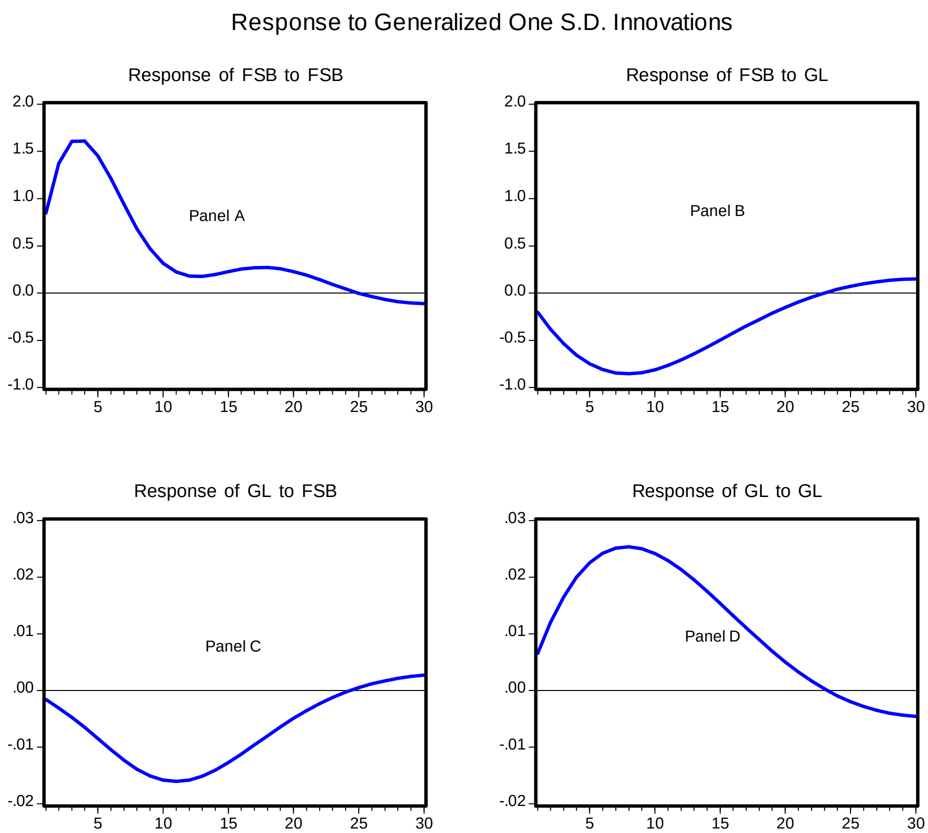


Figure 2 Generalized impulse responses of ΔFSB to ΔGL and vice versa

The procedure we follow is to set up a VAR model with the fiscal and monetary variables as endogenous variables, and then estimate the time-varying parameters in a State-Space model with Markov-Switching. We use the minimum rediscount rate (denoted by ***MRR***) as our measure of the central bank's monetary policy, and the budget balance to GDP ratio (denoted by ***FSB***) as our measure for fiscal policy. Thus, we estimate the following simple equation:

$$FSB_t = a_{1t} + \alpha_{2t} FSB_{t-1} + \alpha_{3t} MRR_{t-1} + \varepsilon_t \tag{2}$$

Where ε_t is a shock with normal distribution and zero mean. We assume that the coefficients a_i are time-varying, and the variance of the shock ε_t is not constant, but rather, has Markov-Switching properties. Hence we define X_t and ϕ_t as

$$X_t = (1\ FSB_{t-1}\ MRR_{t-1}) \tag{3}$$

$$\phi_t = (a_{1t}\ a_{2t}\ a_{3t})' \tag{4}$$

Equation (2) can be rearranged as

$$FSB_t = X_t \phi_t + \varepsilon_t \tag{5}$$

Following Kim (1993), Kim and Nelson (1999) and Maddala and Kim (1998), we assume that ε_t has two states of variance with Markov-switching properties, hence:

$$\varepsilon_t \backsim N\left(0\ \sigma^2_{\varepsilon,SS_t}\right) \tag{6}$$

with

$$\sigma^2_{\varepsilon,SS_t} = \sigma^2_{\varepsilon,0} + \left(\sigma^2_{\varepsilon,1} - \sigma^2_{\varepsilon,0}\right) SS_{t,}\ \sigma^2_{\varepsilon,1} > \sigma^2_{\varepsilon,0} \tag{7}$$

and

$$Pr[SS_t = 1|SS_{t-1} = 1] = p$$
$$Pr[SS_t = 0|SS_{t-1} = 0] = q$$

Where SS_t = 0 or 1, indicates the state of the variance of ε_t and Pr stands for probability. The time-varying vector ϕ_t is assumed to have the following path.

$$\phi_t = \bar{\Phi}_{SS_t} + F\phi_{t-1} + \eta_t, \ \ \eta_t \backsim N\left(0\ \sigma^2_{\eta,SS_t}\right) \qquad (8)$$

Where $\bar{\Phi}_{SS_t}$ denotes the drift of ϕ_t under different states. F is a diagonal matrix with constant elements. η_t is a vector of shocks of normal distribution with zero mean and Markov-switching variance. σ^2_{η,SS_t} is assumed to be a diagonal matrix. If we assume that $E(\varepsilon_t, \eta_t) = 0$, then the State-space model with Markov-switching transition probabilities can be expressed thus:

$$FSB_t = X_t\phi_t + \varepsilon_t, \ \ \varepsilon_t \backsim N\left(0\ \sigma^2_{\varepsilon,SS_t}\right) \qquad (9)$$
$$\phi_t = \bar{\Phi}_{SS_t} + F\phi_{t-1} + \eta_t, \ \ \eta_t \backsim N\left(0\ \sigma^2_{\eta,SS_t}\right) \qquad (10)$$

Equation (9) is called the "Signal" or "Observation" equation, while Equation (10) is referred to as the "State" or "Transition" equation[6].

The estimation of the SS-MS VAR model is done by the maximum likelihood ratio method. The maximization of the likelihood of an MsVAR model results in an iterative process to obtain estimates of autoregressive parameters and of the transition probabilities controlled by the unobserved states of a Markov Chain.

The Ss-MsVAR model is estimated using the Kalman filter, which is a recursive algorithm for sequentially updating the one-step ahead estimates of the state mean and variances, given new information (see Harvey, 1989 and Hamilton and Susmel, 1994 for more details). In a State-space model with Markov-switching, the goal is to form a forecast of ϕ_t based not only on Y_{t-1} (where Y_{t-1} denotes the vector of observations available as at time *t-1*), but also conditional on the random variable SS_t, taking on the value *j* and on SS_{t-1}, taking on the value *i*. Where *i* and *j* equal 0 or 1 respectively. Hence,

$$\phi^{(i,j)}_{t|t-1} = E[\phi_t|Y_{t-1}, SS_t = j, SS_{t-1} = i] \qquad (11)$$

While the corresponding mean square error of the forecast is

$$P^{(i,j)}_{t|t-1} = E\left[\left(\phi_t - \phi_{t|t-1}\right)\left(\phi_t - \phi_{t|t-1}\right)'\middle|Y_{t-1}, SS_t = j, SS_{t-1} = i\right] \qquad (12)$$

Based on the conditions that $SS_{t-1} = i$ and $SS_t = j$ (*i, j = 0, 1*), the Kalman filter algorithm for our model is as follows:

$$\phi^{(i,j)}_{t|t-1} = \bar{\Phi}_j + F\phi^{i}_{t-1|t-1,} \qquad (13)$$
$$P^{(i,j)}_{t|t-1} = FP^{i}_{t-1|t-1}F' + \sigma^2_{\eta,j} \qquad (14)$$
$$\xi^{(i,j)}_{t|t-1} = FSB_t - X_t\phi^{(i,j)}_{t|t-1} \qquad (15)$$
$$\upsilon^{(i,j)}_{t|t-1} = X_tP^{(i,j)}_{t|t-1}X_t{}' + \sigma^2_{\varepsilon,j} \qquad (16)$$
$$\phi^{(i,j)}_{t|t} = \phi^{(i,j)}_{t|t-1} + P^{(i,j)}_{t|t-1}X_t{}'[\upsilon^{(i,j)}_{t|t-1}]^{-1}\xi^{(i,j)}_{t|t-1} \qquad (17)$$
$$P^{(i,j)}_{t|t} = \left(I - P^{(i,j)}_{t|t-1}X_t{}'\left[\upsilon^{(i,j)}_{t|t-1}\right]^{-1}X_t\right)P^{(i,j)}_{t|t-1} \qquad (18)$$

If we observe the sequence of data up to point *T*, then, the process of using this information to form expectations for any time period up to time *T* is known as "fixed-interval smoothing". Additional details on the smoothing procedure can be found in Maddala and Kim (1998) and Eviews 5.1 User's guide.

[6] The complete set of specifications for the Signal and State equations as implemented in Eviews is presented in the Appendix.

5. Empirical Results and Synthesis

The results from the State-space model with Markov-switching are presented in Tables 1, 2, 3 and Figure III. Table 1 presents qualified[7] evidence, which suggests that two different distinct regimes have characterized the interactions between monetary and fiscal policies in Nigeria. The point estimates of the regime dependent means, μ_1 for regime 1 and μ_2 for regime two are statistically different. The estimated mean in regime 1 is negative at -0.1286 and for regime 2, it is positive at 0.5591. These signs validate our hypothesis that within the sample period, the variables dichotomises into phases that exhibit declining and growing interactions. We label the growing phase as the period of accommodative monetary-fiscal policies (i.e. regime 2), and the declining phase as the period of counteractive monetary-fiscal policies (i.e. regime 1). Since the signs assumed by regime 1 and regime 2 are opposing (i.e. negative and positive), it implies that during the early stages of our sample period, both policies where counteractive and that latter on, they were accommodative. Muscatelli et al. (2002) refer to this kind of behaviour of monetary and fiscal policy as being strategic substitutes and complements, respectively.

The fact that we obtained a lower mean for regime (2), indicates that regime (1) (counteractive monetary-fiscal policy) has been the predominant phase during the sample period under review. Whereas, regime (2) can be interpreted as an adjustment strategy, originating from macroeconomic disturbances in the economy. This relationship is clearly depicted in the one-step ahead smoothed estimates of the signal series shown in Figure 3. From the figure, we observe that between 1998 and 2004, the smoothed estimates of fiscal policy were largely expansionary, with increasing government borrowings and liabilities. Whereas, during the same period, the smoothed estimates of monetary policy was contractionary. This kind of policy interaction may be unique to Nigeria's history, as the converse of this relationship is found by Fialho and Portugal (2009) for Brazil between 1995:6 and 1999:12.

Table 1 Parameter Estimates of the SS-MS Model

Parameters	Coefficients	z-statistics
μ_1	-0.1286	-0.2872
μ_2	0.5591	0.3134
ϕ_1	0.4666	0.0006
ϕ_2	-0.0404	-5.72E-05
ϕ_3	0.5452	$7.71 * 10^{-4}$
ϕ_4	0.0512	7.24E-05
ϕ_5	0.10957	$1.55 * 10^{-4}$
ϕ_6	-0.3801	$-5.37 * 10^{-4}$
ϕ_7	-0.9999	$-1.41 * 10^{-3}$
ϕ_8	0.6047	$8.55 * 10^{-4}$
Log likelihood -491.74		

By analyzing regime (1) more closely, we observe that this regime is feasible in more turbulent moments in the history of the Nigerian economy. The period between 1980 and 1994, which was predominantly counteractive, coincides with the oil price crunch of the 1980's, and the period when Nigeria implemented the structural adjustment programme.

[7] The evidence is qualified because it does not provide a clear-cut demarcation

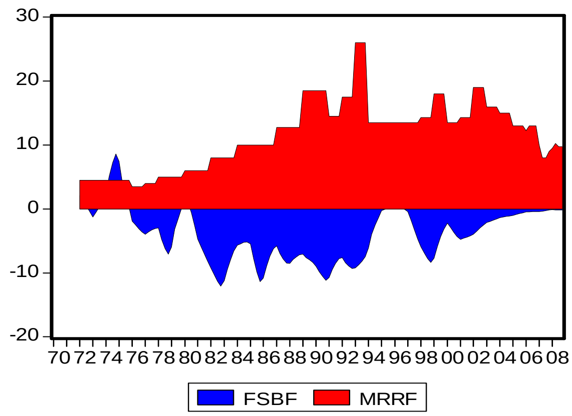


Figure 3 One-step ahead smoothed estimates of the signal series

Where FSBF and MRRF are the one-step ahead forecasts of fiscal balance and the minimum rediscount rate respectively. According to the time-varying transition probability coefficients presented in Table 2, not all the estimated coefficients in the data generating process (DGP) of the transition probabilities are significant.

Table 2 Estimates of Time-Varying Transition Probability Parameters

Parameters	Coefficients	RMSE[8]
α_1	-0.3608	0.3056
β_1	0.3504	0.6641
α_2	1.95505	0.4155
β_2	0.41488	0.2718

The parameters which govern the time-variation of the transition probabilities, namely, α_1 and α_2, have opposite signs. This is consistent with the intuition that an increase in the monetary measure (MRR) decreases the probability of remaining in a counteractive regime and increases the probability of switching regime. The parameters β_1 and β_2 determine the unconditional mean duration of staying in the accommodative or counteractive regimes of monetary and fiscal policy.

Table 3 Transition Probability Matrix

	Regime 1	Regime 2
Regime 1 (Accommodative)	0.8632	0.1368
Regime 2 (Counteractive)	0.2215	0.7785

The probability of transition from regime (1) to (2) and vice versa are displayed in Table 3. The Table shows that if the policy paradigm is in regime (1), at time *t*, then the probability that regime (1) will be maintained at time (*t+1*) is 0.8632, and the probability that the policy regime will shift from (1) to (2) at time (*t+1*) is 0.1368. For an initial state regime of (2), the probability of maintaining regime (2) in the next time period is 0.7785, and that of transiting to regime (1) is 0.2215. These probability values reinforce the results that we obtained from the time-varying coefficients displayed in Table 2.

[8] Where RMSE is the root mean square error.

Overall, we summit that the empirical evidence obtained here are qualified and should be interpreted with caution. This is because the point estimates of the regime dependent means μ_1 and μ_2 both have z-statistics that are not significant.

6. Conclusion

This paper uses quarterly data to explore the monetary and fiscal policy interactions in Nigeria between 1970 and 2008. The paper first examined the salient issues in the theory and literature of the interactions between monetary and fiscal policies. As a preliminary exercise, the paper examined the nature of fiscal policies in Nigeria using a VAR model. The simulated generalized impulse response graphs generated from the VAR estimation provides evidence of a non-Ricardian fiscal policy in Nigeria. These results suggest the validity of the fiscal theory of the price level determination, which postulates that changes in prices are driven by fiscal policies, and that the price level has to adjust to ensure equilibrium in private sector wealth, and government solvency (Woodford, 1995).

Further, the paper analyzes the interactions between monetary and fiscal policies by applying a State-space model with Markov-switching to estimate the time-varying parameters of the relationships. The evidence indicates that monetary and fiscal policies in Nigeria have interacted in a counteractive manner for most of the sample period (1980-1994). At other periods we do not observe any systematic pattern of interaction between the two policy variables, although between 1998 and 2008, some form of accommodativeness can be inferred (see Figure 3). Overall, the results suggest that the two policy regimes- counteractive and accommodative- were weak strategic substitutes during the post 1970 (Civil War) period. This is because the z-statistics of the coefficients of the regime means were not significant. With this kind of result, we identify a game were the fiscal authorities play first, while the monetary authorities are reactive, managing the monetary instrument based on fiscal activities.

For the policy maker, our results imply the existence of fiscal dominance in the interactions between monetary and fiscal policies in Nigeria. The evidence on the implementation of the non-Ricardian fiscal policy and the fiscal theory of the price level, implies that inflation, predominantly results from fiscal problems, and not from lack of monetary control. Based on the results obtained, government should pay attention to monetary activities before embarking on fiscal policies, especially with respect to government liabilities.

We submit that the empirical observations regarding the conclusions presented here are subject to criticisms, especially because of the insignificant z-statistics. However, it will be difficult to criticize the paper now, because to the best of my knowledge, the empirical literature on the interactions between monetary and fiscal policies in Nigeria, with regime switching factored in, is non-existent. This has made it difficult to compare results and conclusions.

The methodology used here can be improved by applying a special kind of Markov-Switching regression model with more than two regimes (see Maddala and Kim, 1998), and introducing another leg to the equation, which will analyze the sensitivity of fiscal policies to the exchange rate dynamics. Another suggestion is to apply a gradual switching State-Space model for two countries (Nigeria and a major trading partner, say China).

ACKNOWLEGDEMENT

I acknowledge the support received from my colleagues in the Department of Economics, University of Uyo, Uyo, Nigeria, and the editorial assistance provided by Ekpeno Effiong. I am also grateful to the two anonymous reviewers of this paper for their useful comments. The responsibilities for any other errors are mine.

REFERENCES

Benhabib, J., S. Schmitt-Grohe and M. Uribe (2001). "Monetary-fiscal policyand multiple equillibria." The American Economic Review, 91 (1) 166-186.

Buiter, W. H. (2001) "The fiscal theory of the price level: A critique" Working Paper, European Bank for Reconstruction and Development.

Cantenaro, M. (2000). "Macroeconomic policy interactions in the EMU: A case for fiscal policy coordination" Mimeograph, Department of Economics Discussion Papers, University of Surry.

Canzoneri, M. B., R. E. Cumby and B. T. Diba (2000). "Is the price level determinacy by the needs of fiscal solvency?" The American Economic Review, 91(5) 1221-1238.

Denton, F.T. (1971). Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. Journal of American Statistical Association, 66, (333) 99-102.

Fialho, M., and M. S. Portugal (2009) "Monetary and fiscal policy interactions in Brazil: An application of the fiscal theory of the price level." Mimeo, Universidade Federal do Rio Grande do Sul- UFRGS.

Hamilton, J.D., and R. Susmel (1994). "Autoregressive conditional heteroskedasticity and changes in regimes" Journal of Econometrics, 64, 307-333.

Harvey, A.C. (1989) "Forecasting Structural Time Series Models and the Kalman Filter," Cambridge University Press, Cambridge.

Kim, I.M. (1993). A dynamic programming approach to the estimation of Markov-switching regression models" Journal of Statistical Computation and Simulations, 45, 61-76.

Kim, I.M., and C. R. Nelson (1999) "State-Space Models with Regime Switching," MIT Press.

Leeper, E. M. (1991) "Equillibrium under 'active' and 'passive' monetary and fiscal policies" Journal of Monetary Economics 27, 129-147.

Leith, C., and S. Wren-Lewis (2000). "Interactions between monetary and fiscal policies", The Economic Journal, 110, March, C93-C108.

Lisman, J.H., and Sandee, J. (1964). Derivation of quarterly figures from annual data. Journal of the Royal Statistical Society, 13(2), 87-90.

Maddala, G. S., and I. Kim (1998). "Unit Roots, Cointegration, and Structural Change" Cambridge University Press, Cambridge.

Melitz, Jacques (1997) "Some cross-country evidence about debt, deficits and the behaviour of monetary and fiscal authority," CEPR Discussion Paper No. 1653.

Muscatelli, V. A., and P. T. Trecroci (2000) "Monetary and fiscal policy interactions over the cycle: some empirical evidence" Discussion Paper, Department of Economics, University of Glasglow.

Sargent, T. (1999). "A premier on monetary and fiscal policy", Journal of Banking and Finance, 23; 1463-1482.

Semmler, W., and W. Zhang (2003) "Monetary and fiscal policy interactions: Some empirical evidence from the Euro-Area". Mimeo, Bielefeld University Working Paper No. 48, March, 2003.

Sims, C. A. (1994). "A simple model for study of determinacy of price level and the interactions of monetary and fiscal policy" Economic Theory 4; 381-399.

Sims, C. A. (1997). "Fiscal foundations of price stability in open economics" Paper presented at the Hong-Kong meeting of the regional Econometric Society on July 24, 1997.

Sims, C. A. (2001). "Fiscal aspects of central bank independence" Mimeograph, Princeton University.

van Aarle, B., J. Engwerda and J. Plasmans (2002). "Monetary and fiscal policy interactions in the EMU: A dynamic game approach" Anals of Operations Research 109; 229-264.

Woodford, M. (1994) "Monetary policy and price level determinancy in a cash in advance economy" Economic Theory 4; 345-380.

Woodford, M. (1995) "Public debt and the price level". Mimeograph, Princeton University.

Woodford, M. (2000). "Fiscal requirements for price stability." NBER Working Paper No. W8072.

Wyplosz, C. (1999), "Economic policy coordination in EMU: strategies and intuitions" CEPR.

APPENDIX

Detailed Specification of the State-Space Markov-switching model

@signal fsb = c(1)*c(1) + sv1*fsb(-1) + sv2*mrr(-1) + sv3*fsb(-1) + sv4*mrr(-1) + [var = exp(c(2))]
@signal mrr = c(3)*c(1) + sv5*fsb(-1) + sv6*mrr(-1) + sv7*fsb(-1) + sv8*mrr(-1) + [var = exp(c(4))]
@state sv1 = sv1(-1)
@state sv2 = sv2(-1)
@state sv3 = c(6) + sv3(-1) + [var = exp(c(5))]
@state sv4 = c(8) + sv4(-1) + [var = exp(c(7))]
@state sv5 = sv5(-1)
@state sv6 = sv6(-1)
@state sv7 = c(10) + sv7(-1) + [var = exp(c(9))]
@state sv8 = c(12) + sv8(-1) + [var = exp(c(11))]

Unit root test results

Variable	**ADF**		**KPSS**		**Conclusion**
	Level	1st Difference	Level	1st Difference	
FSB	-2.66[9]	-3.66[8]***	0.27[9] ***	0.06[1]	I(1)
GL	-2.15[12]	-3.26[11]***	0.26[10] ***	0.10[5]	I(1)
MRR	-2.05[0]	-12.14[0]***	1.10[10] ***	0.19[20]	I(1)

Notes: ***,** and * indicates significance at the 1%, 5%, 10% levels respectively. The values in bracket for the ADF test indicates the optimal lag length selected by the SIC within a maximum lag of 13. The values in bracket for the KPSS test indicate the bandwidth selection, using the Newey-West's Bartlett Kernel.

Optimal Designs Approach to Portfolio Selection

I.A. Etukudo[1]

In order to obtain the best tradeoff between risk and return, optimization algorithms are particularly useful in asset allocation in a portfolio mix. Such algorithms and proper solution techniques are very essential to investors in order to circumvent distress in business outfits. In this paper, we show that by minimizing the total variance of the portfolio involving stocks in two Nigerian banks which is a measure of risk, optimal allocation of investible funds to the portfolio mix is obtained. A completely new solution technique – modified super convergent line series algorithm which makes use of the principles of optimal designs of experiment is used to obtain the desired optimizer.

Keywords: Portfolio selection, minimum variance, optimal designs, optimal allocation.

1. Introduction

In every investment, there is a tradeoff between risk and returns on such investment. An investor therefore must be willing to take on extra risk if he intends to obtain additional expected returns. However, there must be a balance between risk and returns that suits individual investors, Neveu (1985).

Great care must be taken by any investor in the allocation of his investible funds to a list of investments open to him in order to minimize the total risk involved. A mathematical model to suit a problem of this nature and in particular, a quadratic programming model for portfolio selection was developed by Markowitz (1952, 1959).

A portfolio mix is a set of investments that an investor can invest in while a portfolio risk refers to the risk common to all securities in the portfolio mix and this is equated with the standard deviation of returns, Ebrahim (2008).

The purpose of the investment of cash in portfolios of securities is to provide a better return than would be earned if the money were retained as cash or as a bank deposit. The return may come in the form of a regular income by way of dividends or interest or by way of growth in capital value or by a combination of both regular income and growth in capital value, Cohen and Zinbarg (1967). Thus, the real objective of portfolio construction becomes that of achieving the maximum return with minimum risk, Weaver (1983).

Grubel (1968) showed that higher returns and lower risks than the usual are obtained from international diversification. Arnott and Copeland (1985) have also shown that the business cycle has a significant effect on security returns. On their part, Chen, Roll and Ross (1986) determined that certain macroeconomic variables are significant indicators of changes in stock returns. Contributing further, Bauman and Miller (1995) showed that the evaluation of portfolio performance should take place through a complete stock market cycle because of differences in performance during the market cycle. Macedo (1995) demonstrates that switching between relative strength and relative value strategies can increase returns in an international portfolio.

Since portfolio selection problem is a quadratic programming problem which involves a minimization of risk associated with such investment by minimizing the total variance which is a measure of the risk involved, Francis (1980), suitable solution technique should be adopted to obtain optimal solution. Etukudo and Umoren (2009) have shown that it is easier and in fact better to use modified super convergent line series algorithm ($MSCLS_Q$) which uses the principles of optimal designs of experiment in solving quadratic programming problems rather than using the traditional solution technique of modified simplex method. This paper therefore focuses on optimal designs approach to optimal allocation of investible funds in a portfolio mix.

[1] Department of Mathematics/Statistics & Computer Science, University of Calabar, Calabar, Nigeria, nseidorenyin@gmail.com

2. A quadratic programming model for portfolio selection

For a quadratic programming model for portfolio selection, let

n = number of stocks to be included in the portfolio
x_j = number of shares to be purchased in stocks j, j = 1, 2, …, n
Y_j = returns per unit of money invested in stocks j at maturity

Assuming the values of Y_j are random variables, then

$$E(Y_j) = \overline{Y}_j; \quad j = 1,2,\ldots,n \tag{1}$$
$$V = \sigma_{ij} = E\left[\left(Y_i - \overline{Y}_i\right)\left(Y_j - \overline{Y}_j\right)\right] \tag{2}$$

where $E(Y_j)$ is the mathematical expectation of Y_j and V is the variance – covariance matrix of the returns. See Gruyter (1987), Parsons (1977) and Etukudo et al (2009). Hence, the variance of the total returns or the portfolio variance is given by

$$f(\mathbf{x}) = \mathbf{X'VX} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j \tag{3}$$

which measures the risk of the portfolio selected. The non-negativity constraints are

$$x_j \geq 0,\ j = 1, 2, \ldots, n \tag{4}$$

Assuming the minimum expected returns per unit of money invested in the portfolio is B, then

$$\sum_{j=1}^{n}\overline{Y}_j x_j \geq B \tag{5}$$

2.1 Minimization of the total risk involved in the portfolio

By minimizing the total variance, f(**x**) of the portfolio, the total risk involved in the portfolio is minimized. In order to obtain a minimum point of equation 3, f(**x**) must be a convex function, Hillier and Lieberman (2006). That is,

$$\frac{\partial^2 f(x_j)}{\partial x_1^2}\frac{\partial^2 f(x_j)}{\partial x_2^2}\cdots\frac{\partial^2 f(x_j)}{\partial x_j^2} - \left[\frac{\partial^2 f(x_j)}{\partial x_i \partial x_j}\right]^2 - \ldots - \left[\frac{\partial^2 f(x_j)}{\partial x_i \partial x_j}\right]^2 \geq 0 \tag{6}$$

where

$$\begin{gathered}\frac{\partial^2 f(x_j)}{\partial x_1^2} \geq 0\\ \vdots \\ \frac{\partial^2 f(x_j)}{\partial x_j^2} \geq 0\end{gathered} \tag{7}$$

where $i \neq j = 1, 2, \ldots, n$. Strict inequalities of 6 and 7 imply that f(**x**) is strictly convex and hence, has a global minimum at **x***. From equation 3 and inequalities 4 and 5, the portfolio selection model is given by;

$$\text{Min } f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j$$

subject to:

$$\sum_{j=1}^{n}\overline{Y}_j x_j \geq B \qquad ; \quad x_j \geq 0,\ j = 1, 2, \ldots, n$$

Remark

The expected values, $\overline{Y}_j$ and the variance – covariance matrix, σ_{ij} are based on data from historical records.

3. Modified super convergent line series algorithm ($MSCLS_Q$), Umoren and Etukudo (2009)

The sequential steps involved in $MSCLS_Q$ are given as follows:

Step 1: Let the response surface be

$$y = c_0 + c_1x_1 + c_2x_2 + q_1x_1^2 + q_2x_1x_2 + q_3x_2^2 \qquad x_1, x_2 \in G_i, \; i = 1, 2, \ldots, k^*$$

Select N support points such that $3k^* \le N \le 4k^*$ where $2 \le k^* \le 3$ is the number of partitioned groups desired. By arbitrarily choosing the support points as long as they do not violate any of the constraints, make up the initial design matrix

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1N} & x_{2N} \end{bmatrix}$$

Step 2: Partition X into k* groups with equal number of support points and obtain the design matrix, X_i, i = 1, 2, …, k* for each group. Obtain the information matrices $M_i = X_i'X_i$, i = 1, 2, …, k* and their inverses M_i^{-1}, i = 1, 2, …, k* such that

$$M_i^{-1} = \begin{bmatrix} v_{i11} & v_{i21} & v_{1i1} \\ v_{i12} & v_{i22} & v_{i32} \\ v_{i13} & v_{i23} & v_{i33} \end{bmatrix}$$

Step 3: Compute the matrices of the interaction effect of the variables for the groups. These are

$$X_{iI} = \begin{bmatrix} x_{i11}^2 & x_{i11}x_{i21} & x_{i21}^2 \\ x_{i12}^2 & x_{i12}x_{i22} & x_{i22}^2 \\ \vdots & \vdots & \vdots \\ x_{i1N}^2 & x_{i1N}x_{i2N} & x_{i2N}^2 \end{bmatrix}$$

where i = 1, 2, …, k* and the vector of the interaction parameters obtained from f(**x**) is given by

$$\mathbf{g} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

The interaction vectors for the groups are given by $\mathbf{I}_i = M_i^{-1}X_i'X_{iI}\mathbf{g}$ and the matrices of mean square error for the groups are

$$\overline{M}_i = M_i^{-1} + \mathbf{I}_i\mathbf{I}_i' = \begin{bmatrix} \overline{v}_{i11} & \overline{v}_{i21} & \overline{v}_{i31} \\ \overline{v}_{i12} & \overline{v}_{i22} & \overline{v}_{i32} \\ \overline{v}_{i13} & \overline{v}_{i23} & \overline{v}_{i33} \end{bmatrix}$$

Step 4: Compute the optimal starting point, $\overline{\mathbf{x}}_1^*$ from

$$\overline{\mathbf{x}}_1^* = \sum_{m=1}^{N} u_m^* \mathbf{x}_m; \quad u_m^* > 0; \quad \sum_{m=1}^{N} u_m^* = 1, \quad u_m^* = \frac{a_m^{-1}}{\sum_{m=1}^{N} a_m^{-1}}, \quad a_m = \mathbf{x}_m'\mathbf{x}_m, \quad m = 1, 2, \ldots, N$$

Step 5: The matrices of coefficient of convex combinations of the matrices of mean square error are

$$H_i = \mathrm{diag}\left\{\frac{\bar{v}_{i11}}{\sum \bar{v}_{i11}}, \frac{\bar{v}_{i22}}{\sum \bar{v}_{i22}}, \frac{\bar{v}_{i33}}{\sum \bar{v}_{i33}}\right\} = \mathrm{diag}\{h_{i1}, h_{i2}, h_{i3}\}\ ,\ i = 1, 2, \ldots, k^*$$

By normalizing H_i such that $\sum H_i^* H_i^{*'} = I$, we have

$$H_i^* = \mathrm{diag}\left\{\frac{h_{i1}}{\sqrt{\sum h_{i1}^2}}, \frac{h_{i2}}{\sqrt{\sum h_{i2}^2}}, \frac{h_{i3}}{\sqrt{\sum h_{i3}^2}}\right\}$$

The average information matrix is given by

$$M(\xi_N) = \sum_{i=1}^{k} H_i^* M_i H_i^{*'} = \begin{bmatrix} \bar{m}_{11} & \bar{m}_{21} & \bar{m}_{31} \\ \bar{m}_{12} & \bar{m}_{22} & \bar{m}_{32} \\ \bar{m}_{13} & \bar{m}_{23} & \bar{m}_{33} \end{bmatrix}$$

Step 6: From f(**x**), obtain the response vector

$$\mathbf{z} = \begin{bmatrix} z_0 \\ z_1 \\ z_2 \end{bmatrix} \text{ where } z_0 = f(\bar{m}_{12} - \bar{m}_{13});\ z_1 = f(\bar{m}_{22} - \bar{m}_{23});\ z_2 = f(\bar{m}_{32} - \bar{m}_{33})$$

Hence, we define the direction vector

$$\mathbf{d} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \end{bmatrix} = M^{-1}(\xi_N)\mathbf{z}$$

and by normalizing **d** such that $\mathbf{d}^{*'}\mathbf{d}^{*} = 1$, we have

$$\mathbf{d}^* = \begin{bmatrix} d_1^* \\ d_2^* \end{bmatrix} = \begin{bmatrix} \dfrac{d_1}{\sqrt{d_1^2 + d_2^2}} \\ \dfrac{d_2}{\sqrt{d_1^2 + d_2^2}} \end{bmatrix}$$

Step 7: Obtain the step length, ρ_1^* from $\rho_1^* = \min_i \left\{\dfrac{\mathbf{c}_i' \bar{\mathbf{x}}_1^* - b_i}{\mathbf{c}_i' \mathbf{d}^*}\right\}$ where $\mathbf{c}_i'\mathbf{x} = b_i$, i = 1, 2, … , m is the i^{th} constraint of the quadratic programming problem.

Step 8: Make a move to the point $\mathbf{x}_2^* = \bar{\mathbf{x}}_1^* - \rho_1^* \mathbf{d}^*$

Step 9: Compute f($\mathbf{x}_2^*$) and f($\mathbf{x}_1^*$). Is $|f(\mathbf{x}_2^*) - f(\mathbf{x}_1^*)| \le \varepsilon$ where $\varepsilon = 0.0001$, then stop for the current solution is optimal, otherwise, replace $\bar{\mathbf{x}}_1^*$ by $\mathbf{x}_2^*$ and return to step 7. If the new step length, ρ_2^* is negligibly small, then an optimizer had been located at the first move.

4. A Numerical Example

An investor has a maximum of N10, 000.00 to invest by purchasing shares in Oceanic Bank and First City Monument Bank. Below is the historical data of prices per share in the banks for 25 days.
We are required to obtain optimal allocation of the investible funds for purchase of shares in the portfolio in order to minimize the total risk in the portfolio mix. From the data on table 3.1, the mean prices per share for First City Monument Bank and Oceanic Bank are 18.13 and 28.19 respectively.

Table 3.1: Price per share

Day	FCMB (Y_1)	Oceanic Bank (Y_2)	Day	FCMB (Y_1)	Oceanic Bank (Y_2)
1	17.6	29.89	14	18.5	26.95
2	17	29.61	15	18.78	26
3	17.55	28.95	16	18.4	27.3
4	17.9	27.95	17	18.74	28.86
5	17.5	28	18	18.74	30.09
6	17.7	28.61	19	19.1	30.6
7	17.74	28.6	20	18.71	29.6
8	17	26.49	21	18.9	28.6
9	16.8	25.99	22	18.75	29.01
10	16.99	25.95	23	18.53	28.98
11	18.5	27.3	24	18.5	28.55
12	18.8	27.17	25	18.49	28.75
13	18.1	26.99			
			$\sum Y_i$	453.32	704.79
			$\overline{Y}$	18.13	28.19

Source: The Nigerian Stock Exchange

Table 3.2: Mean deviation

Day	$(Y_1 - \overline{Y}_1)$	$(Y_2 - \overline{Y}_2)$	Day	$(Y_1 - \overline{Y}_1)$	$(Y_2 - \overline{Y}_2)$
1	-0.5328	1.6984	14	0.3672	-1.2416
2	-1.1328	1.4184	15	0.6472	-2.1916
3	-0.5828	0.7584	16	0.2672	-0.8916
4	-0.2328	-0.2416	17	0.6072	0.6684
5	-0.6328	-0.1916	18	0.6072	1.8984
6	-0.4328	0.4184	19	0.9672	2.4084
7	-0.3928	0.4084	20	0.5772	1.4084
8	-1.1328	-1.7016	21	0.7672	0.4084
9	-1.3328	-2.2016	22	0.6172	0.8184
10	-1.1428	-2.2416	23	0.3972	0.7884
11	0.3672	-0.8916	24	0.3672	0.3584
12	10.6672	-1.0216	25	0.3572	0.5584
13	-0.0328	-1.2016			

The expected return per share is the difference between the mean price of that share and its price on the 25th day. The investor assumes that his expected returns would be at least N100.00. Since his objective is to minimize his total risk, the problem involves obtaining optimal portfolio mix where the investment is done at the 25th day prices.

The share price deviations are obtained from table 3.1 as shown in table 3.2 while the variance – covariance matrix table for the share price are obtained from table 3.1 as shown in table 3.3.

From the table 3.3, the variance- covariance matrix is given by

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 0.4863 & 0.3078 \\ 0.3078 & 1.7846 \end{pmatrix}$$

Hence, the model for minimizing the total risk of the portfolio is

$$\text{Min } f(x) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0.4863 & 0.3078 \\ 0.3078 & 1.7846 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.4863x_1^2 + 0.6156x_1x_2 + 1.7846x_2^2$$

Subject to:

$$\begin{aligned} 18.49x_1 + 28.75x_2 &\le 10{,}000 \\ 0.3572x_1 + 0.5584x_2 &\ge 100 \\ x_1, x_2 &\ge 0 \end{aligned}$$

where x_1 and x_2 are respectively the number of shares purchased from First City Monument Bank and Oceanic Bank in the Portfolio.

Table 3.3: Variance- covariance matrix value

Day	$(Y_1-\bar{Y}_1)^2$	$(Y_1-\bar{Y}_1)(Y_2-\bar{Y}_2)$	$(Y_2-\bar{Y}_2)^2$	**Day**	$(Y_1-\bar{Y}_1)^2$	$(Y_1-\bar{Y}_1)(Y_2-\bar{Y}_2)$	$(Y_2-\bar{Y}_2)^2$
1	0.28387584	-0.90490752	2.88456256	14	0.13483584	-0.45591552	1.54157056
2	1.28323584	-1.60676352	2.01185856	15	0.41886784	-1.41840352	4.80311056
3	0.33965584	-0.44199552	0.57517056	16	0.07139584	-0.23823552	0.79495056
4	0.05419584	0.05624448	0.05837056	17	0.36869184	0.40585248	0.44675856
5	0.40043584	0.12124448	0.03671056	18	0.36869184	1.15270848	3.60392256
6	0.18731584	-0.18108352	0.17505856	19	0.93547584	2.32940448	5.80039056
7	0.15429184	-0.16041952	0.16679056	20	0.33315984	0.81292848	1.98359056
8	1.28323584	1.92757248	2.89544256	21	0.58859584	0.31332448	0.16679056
9	1.77635584	2.93429248	4.84704256	22	0.38093584	0.50511648	0.66977856
10	1.30599184	2.56170048	5.02477056	23	0.15776784	0.31315248	0.62157456
11	0.13483584	-0.32739552	0.79495056	24	0.13483584	0.13160448	0.12845056
12	0.44515584	-0.68161152	1.04366656	25	0.12759184	0.19946048	0.31181056
13	0.00107584	0.03941248	1.44384256				
				$\sum_{i=1}^{2}\sum_{j=1}^{2}(Y_i-\bar{Y}_i)(Y_j-\bar{Y}_j)$	11.670504	7.387288	42.830936
				σ_{ij}	0.486271	0.307803667	1.78462233

5. Test for Convexity

Since
$$\frac{\partial^2 f(x_j)}{\partial x_1^2}\frac{\partial^2 f(x_j)}{\partial x_2^2} - \left[\frac{\partial^2 f(x_j)}{\partial x_1 \partial x_2}\right]^2 = 3.0924 > 0$$

$$\frac{\partial^2 f(x_j)}{\partial x_1^2} = 0.9726 > 0 \text{ and } \frac{\partial^2 f(x_j)}{\partial x_2^2} = 3.5692 > 0$$

f(**x**) is strictly a convex function and its global minimum point, **x*** is obtained by solving the above portfolio selection problem.

6. Solution to the portfolio selection problem by optimal designs approach

$$\text{Minimize } f(x) = 0.4863x_1^2 + 0.6156x_1x_2 + 1.7846x_2^2$$

Subject to:

$$18.49x_1 + 28.75x_2 \quad \leq 10{,}000$$
$$0.3572x_1 + 0.5584x_2 \quad \geq 100$$
$$x_1, x_2 \geq 0$$

Let $\tilde{X}$ be the area defined by the constraint. Hence

$$\tilde{X} = \{x_1, x_2\}$$

Step 1: Select N support points such that $3k^* \leq N \leq 4k^*$ where $2 \leq k^* \leq 3$ is the number of partitioned groups desired. By choosing $k^* = 2$, we have $6 \leq N \leq 8$

Hence, by arbitrarily choosing 8 support points as long as they do not violate the constraints (within the feasible region), the initial design matrix is

$$x = \begin{bmatrix} 1 & 125 & 110 \\ 1 & 100 & 130 \\ 1 & 100 & 120 \\ 1 & 110 & 110 \\ 1 & 130 & 100 \\ 1 & 110 & 120 \\ 1 & 125 & 100 \\ 1 & 95 & 120 \end{bmatrix}$$

Step 2: Partition X into 2 groups such that

$$G_1 = \{x_1, x_2 ; 100 \leq x_1 \leq 125, 110 \leq x_2 \leq 130\}$$
$$G_2 = \{x_1, x_2 ; 95 \leq x_1 \leq 130, 100 \leq x_2 \leq 120\}$$

and the design matrices for the two groups are

$$X_1 = \begin{bmatrix} 1 & 125 & 110 \\ 1 & 100 & 130 \\ 1 & 100 & 120 \\ 1 & 110 & 110 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 & 130 & 100 \\ 1 & 110 & 120 \\ 1 & 125 & 100 \\ 1 & 95 & 120 \end{bmatrix}$$

The respective information matrices are

$$M_1 = X_1^{'}X_1 = \begin{bmatrix} 4 & 435 & 470 \\ 435 & 47725 & 50850 \\ 470 & 50850 & 55500 \end{bmatrix} \text{ and } M_2 = X_2^{'}X_2 = \begin{bmatrix} 4 & 460 & 440 \\ 460 & 53650 & 50100 \\ 440 & 50100 & 48800 \end{bmatrix}$$

Step 3: The matrices of the interaction effect of the variables are

$$X_{1I} = \begin{bmatrix} 15625 & 13750 & 12100 \\ 10000 & 13000 & 16900 \\ 10000 & 12000 & 14400 \\ 12100 & 12100 & 12100 \end{bmatrix} \text{ and } X_{2I} = \begin{bmatrix} 16900 & 13000 & 10000 \\ 12100 & 13200 & 14400 \\ 15625 & 125000 & 10000 \\ 9025 & 14400 & 14400 \end{bmatrix}$$

and the vector of the interaction parameters obtained from f(x) is given by

$$\mathbf{g} = \begin{bmatrix} 0.4863 \\ 0.6156 \\ 1.7846 \end{bmatrix}$$

The interaction vectors for the groups are

$$\mathbf{I_1} = M_1^{-1}X_1^{'}X_{1I}\mathbf{g} = \begin{bmatrix} -40534 \\ 186 \\ 499 \end{bmatrix}$$

and

$$\mathbf{I_2} = M_2^{-1}X_2^{'}X_{2I}\mathbf{g} = \begin{bmatrix} -34541 \\ 175 \\ 459 \end{bmatrix}$$

The matrices of mean square error for the groups are respectively

$$\overline{M}_1 = M_1^{-1} + I_1I_1^{'} = \begin{bmatrix} 1643005496.62 & -7539324.31 & 20226464.31 \\ -7539324.31 & 34596.00 & 94311.00 \\ 20226464.31 & 94311.00 & 249001.00 \end{bmatrix}$$

$$\overline{M}_2 = M_2^{-1} + I_2I_2^{'} = \begin{bmatrix} 1193081221.55 & 6044672.98 & 15854316.20 \\ 6044672.98 & 30625.01 & 80325.01 \\ 15854316.20 & 80325.01 & 210681.02 \end{bmatrix}$$

Step 4: Obtain the optimal starting point

$$\overline{x}_1^* = \sum_{m=1}^{N} u_m^* x_m; \; u_m^* > 0; \; \sum_{m=1}^{N} u_m^* = 1, \; u_1^* = \frac{a_m^{-1}}{\sum_{m=1}^{N} a_m^{-1}}, \; a_m = x_m^{'} x_m, \; m = 1, 2, ..., N$$

Now,

$$a_1 = \mathbf{x_1^{'}x_1} = \begin{bmatrix} 1 & 125 & 110 \end{bmatrix} \begin{bmatrix} 1 \\ 125 \\ 110 \end{bmatrix} = 27726, \; a_1^{-1} = 0.00003607$$

$$a_2 = \mathbf{x_2^{'}x_2} = \begin{bmatrix} 1 & 100 & 130 \end{bmatrix} \begin{bmatrix} 1 \\ 100 \\ 130 \end{bmatrix} = 26901, \; a_2^{-1} = 0.00003717$$

$$a_3 = \mathbf{x_3^{'}x_3} = \begin{bmatrix} 1 & 100 & 120 \end{bmatrix} \begin{bmatrix} 1 \\ 100 \\ 120 \end{bmatrix} = 24401, \; a_3^{-1} = 0.00004098$$

$$a_4 = \mathbf{x_4^{'}x_4} = \begin{bmatrix} 1 & 110 & 110 \end{bmatrix} \begin{bmatrix} 1 \\ 110 \\ 110 \end{bmatrix} = 24201, \; a_4^{-1} = 0.00004132$$

$$a_5 = \mathbf{x_5^{'}x_5} = \begin{bmatrix} 1 & 130 & 100 \end{bmatrix} \begin{bmatrix} 1 \\ 130 \\ 100 \end{bmatrix} = 26901, \; a_5^{-1} = 0.00003717$$

$$a_6 = \mathbf{x_6^{'}x_6} = \begin{bmatrix} 1 & 110 & 120 \end{bmatrix} \begin{bmatrix} 1 \\ 110 \\ 120 \end{bmatrix} = 26501, \; a_6^{-1} = 0.00003773$$

$$a_7 = \mathbf{x_7^{'}x_7} = \begin{bmatrix} 1 & 125 & 100 \end{bmatrix} \begin{bmatrix} 1 \\ 125 \\ 100 \end{bmatrix} = 25626, \; a_7^{-1} = 0.00003902$$

$$a_8 = \mathbf{x_8^{'}x_8} = \begin{bmatrix} 1 & 95 & 125 \end{bmatrix} \begin{bmatrix} 1 \\ 95 \\ 125 \end{bmatrix} = 23426, \; a_8^{-1} = 0.00004269$$

$$\sum_{m=1}^{8} a_m^{-1} = 0.00003607 + 0.00003717 + 0.00004098 + 0.00004132 + 0.00003717 + 0.00003773$$
$$+ 0.00003902 + 0.00004269 = 0.0003122$$

Since

$$u_1^* = \frac{a_m^{-1}}{\sum_{m=1}^{N} a_m^{-1}}, \; m = 1, 2, ..., N$$

$$u_1^* = \frac{0.00003607}{0.0003122} = 0.1155\,,\quad u_2^* = \frac{0.00003717}{0.0003122} = 0.0119$$
$$u_3^* = \frac{0.00004098}{0.0003122} = 0.1313 \quad u_4^* = \frac{0.00004132}{0.0003122} = 0.1324$$
$$u_5^* = \frac{0.00003717}{0.0003122} = 0.1191\,,\quad u_6^* = \frac{0.00003773}{0.0003122} = 0.1209$$

$$u_7^* = \frac{0.00003902}{0.0003122} = 0.1250\,,\quad u_8^* = \frac{0.00004269}{0.0003122} = 0.1367$$

Hence, the optimal starting point is $\bar{\mathbf{x}}_1^* = \sum_{m=1}^{8} u_m^* x_m$

$$\bar{x}_1^* = 0.1155\begin{bmatrix}1\\125\\110\end{bmatrix} + 0.0119\begin{bmatrix}1\\100\\130\end{bmatrix} + 0.1313\begin{bmatrix}1\\100\\120\end{bmatrix} + 0.1324\begin{bmatrix}1\\110\\110\end{bmatrix}$$
$$+\ 0.1191\begin{bmatrix}1\\130\\100\end{bmatrix} + 0.1209\begin{bmatrix}1\\110\\120\end{bmatrix} + 0.1250\begin{bmatrix}1\\125\\100\end{bmatrix}$$
$$+\ 0.1367\begin{bmatrix}1\\95\\120\end{bmatrix} = \begin{bmatrix}1.0000\\111.4350\\113.8299\end{bmatrix}$$

Step 5: Obtain the matrices of coefficients of convex combinations from $\overline{M}_1$ and $\overline{M}_2$ as follows:

$$H_1 = \mathbf{diag}\left\{\frac{1193081221.55}{1193081221.55 + 1643005496.62},\ \frac{30625.01}{30625.01 + 34596.01},\ \frac{210681.02}{210681.02 + 249001.02}\right\} = \mathrm{diag}\{0.4207,\ 0.4696,\ 0.4583\}$$

$$H_2 = I - H_1 = \mathbf{diag}\{0.5793,\ 0.5304, 0.5417\}$$

and by normalizing H_1 and H_2 such that $H_1^* H_1^{*'} + H_2^* H_2^* = 1$, we have

$$H_1^* = \mathbf{diag}\left\{\frac{0.4207}{\sqrt{0.4207^2 + 0.5793^2}},\frac{0.4696}{\sqrt{0.4696^2 + 0.5304^2}},\frac{0.4583}{\sqrt{0.4583^2 + 0.5417^2}}\right\}$$

$$= \mathbf{diag}\{0.5876, 0.6629, 0.6459\}$$

$$H_2^* = \mathbf{diag}\left\{\frac{0.5793}{\sqrt{0.4207^2 + 0.5793^2}},\frac{0.5304}{\sqrt{0.4696^2 + 0.5304^2}},\frac{0.5417}{\sqrt{0.4583^2 + 0.5417^2}}\right\}$$

$$= \mathbf{diag}\{0.8091, 0.7487, 0.7634\}$$

The average information matrix is given by

$$M(\xi_N) = H_1^* X_1^{'} X_1 H_1^{*'} + H_2^* X_2^{'} X_2 H_2^{*'}$$

$$= \begin{bmatrix} 4 & 448 & 450 \\ 448 & 51046 & 50408 \\ 450 & 50408 & 51595 \end{bmatrix}$$

Step 6: From $f(x_1, x_2)$, obtain the response vector $\mathbf{z} = \begin{bmatrix} z_0 \\ z_1 \\ z_2 \end{bmatrix}$

$$z_0 = f(448, 450) = 0.4863(448)^2 + 0.615644(448)(450)$$
$$+1.7846(450)^2 = 583090$$
$$z_1 = f(51046,\ 50408) = 0.4863(510\,46)^2 + 0.615644(5\,1046)(5040\,8)$$
$$+1.7846(504\,08)^2 = 7385800000$$
$$z_2 = f(50408, 51595) = 0.4863(50408)^2 + 0.615644(50408)(51595)$$
$$+1.7846(51595)^2 = 7587400000$$

Therefore,

$$\mathbf{z} = \begin{bmatrix} 583090 \\ 7385800000 \\ 7587400000 \end{bmatrix}$$

Here, we define the direction vector

$$\mathbf{d} = \begin{bmatrix} \underline{d_0} \\ d_1 \\ d_2 \end{bmatrix} = M^{-1}(\xi_N)\mathbf{z} = \begin{bmatrix} \underline{-1936900000} \\ 8900000 \\ 8300000 \end{bmatrix}$$

and by normalizing **d** such that $\mathbf{d}^{*'}\mathbf{d}^* = \mathbf{1}$, we have

$$\mathbf{d}^* = \begin{bmatrix} d_1^* \\ d_2^* \end{bmatrix} = \begin{bmatrix} \dfrac{8900000}{\sqrt{8900000^2 + 8300000^2}} \\[2ex] \dfrac{8300000}{\sqrt{8900000^2 + 8300000^2}} \end{bmatrix} = \begin{bmatrix} 0.7313 \\ 0.6820 \end{bmatrix}$$

Step 7: Obtain the step length, ρ_i^* from

$$\rho_i^* = \min_i \left\{ \frac{c_i^{'}\bar{\mathbf{x}}_1^* - b_i}{c_i^{'} d^*} \right\}$$

where $c_i^{'}x = b_i$, $i = 1, 2, ..., m$ is the ith constraint of the portfolio selection problem.

For $c_1 = \begin{bmatrix} 18.49 \\ 28.75 \end{bmatrix}$ and $b_1 = 10000$, we have $\rho_1^* = \left\{ \dfrac{[18.49 \quad 28.75]\begin{bmatrix} 111.4350 \\ 113.8299 \end{bmatrix} - 10000}{[18.49 \quad 28.75]\begin{bmatrix} 0.7313 \\ 0.6820 \end{bmatrix}} \right\} = \text{-}140.8659$

For $c_2 = \begin{bmatrix} 0.3572 \\ 0.5584 \end{bmatrix}$ and $b_2 = 100$, we have $\rho_2^* = \left\{ \dfrac{\begin{bmatrix} 0.3572 & 0.5584 \end{bmatrix}\begin{bmatrix} 111.4350 \\ 113.8299 \end{bmatrix} - 100}{\begin{bmatrix} 0.3572 & 0.5584 \end{bmatrix}\begin{bmatrix} 0.7313 \\ 0.6820 \end{bmatrix}} \right\} = 5.2443$

Step 8: Make a move to the point

$$\mathbf{x}_2^* = \bar{\mathbf{x}}_1^* - \rho_1^* d^* = \begin{bmatrix} 111.4350 \\ 113.8299 \end{bmatrix} - [-140.8659]\begin{bmatrix} 0.7313 \\ 0.6820 \end{bmatrix} = \begin{bmatrix} 214.4242 \\ 209.9040 \end{bmatrix},$$

since $\rho_1^* = -140.8659$ is the minimum step length.

Step 9

$$\begin{aligned} f(\mathbf{x}_2^*) &= 0.4863(214.4242)^2 + 0.6156(214.4242)(209.9040) \\ &\quad + 1.7846(209.9040)^2 = 128700 \\ f(\bar{\mathbf{x}}_1^*) &= 0.4863(111.4350)^2 + 0.6156(111.4350)(113.8299) \\ &\quad + 1.7846(113.8299)^2 = 36971 \end{aligned}$$

$$\text{since } \left| f(\mathbf{x}_2^*) - f(\bar{\mathbf{x}}_1^*) \right| = |128700 - 36971| = 91729$$

Make a second move by replacing $\bar{\mathbf{x}}_1^* = \begin{bmatrix} 111.4350 \\ 113.8299 \end{bmatrix}$ by $\mathbf{x}_2^* = \begin{bmatrix} 214.4242 \\ 209.9040 \end{bmatrix}$

The new step length is obtained as follows: $\rho_3^* = \left\{ \dfrac{\begin{bmatrix} 18.49 & 28.75 \end{bmatrix}\begin{bmatrix} 214.4242 \\ 209.9040 \end{bmatrix} - 100}{\begin{bmatrix} 18.49 & 28.75 \end{bmatrix}\begin{bmatrix} 0.7313 \\ 0.6820 \end{bmatrix}} \right\} = -0.000125$

Since the new step length is negligible, the optimal solution was obtained at the first move and hence,

$$\mathbf{x}_2^* = \begin{bmatrix} 214.4242 \\ 209.9040 \end{bmatrix} \text{ and } f(\mathbf{x}_2^*) = 128700$$

The portfolio selection problem which is a minimization of portfolio variance was solved using modified super convergent line series algorithm which gave

$$x_1 = 214$$
$$x_2 = 210$$

as the number of shares to be purchased from Oceanic Bank and First City Monument Bank respectively in order to obtain a minimum risk or minimum variance.

7. Summary and Conclusion

In this paper, we assumed that the portfolio has already been selected by the investor from a list of available investments. Using historical data prices (25 days) of stocks from First City Monument Bank and Oceanic Bank, we showed how optimal allocations of investible funds could be made to each Bank's stocks by minimizing the portfolio variance thereby minimizing the total risk using optimal designs approach.

The approach adopted in obtaining optimal solution is recommended for use by potential investors as a way out of business collapse.

References

Arnott, R. D. and Copeland, W. A. (1985): The Business Cycle and Securities Selection. *Financial Analysts Journal*, V41(2), 26 – 32.

Bauman, W. S. and Miller, R. E. (1995): Portfolio Performance Rankings in Stock Market Cycles. *Financial Analysts Journal*, V51(2), 79 - 87.

Chen, Nai – Fu, Roll, R. and Ross, S. (1986): Economic Forces and the Stock Market. *Journal of Business*, V59 (3), 83 – 404

Cohen, J. B. and Zinbarg E. D. (1967) Investment Analysis and Portfolio Management , Irwin. 111.

Ebrahim, Zain (2008): Modern Portfolio Theory, 2008 (http://en.wikipedia.org.)

Etukudo, I. A, Effanga, E. O., Onwukwe, C. E. and Umoren, M. U. (2009): Application of Portfolio Selection Model for Optimal Allocation of Investible Funds in a Portfolio Mix. *Scientia Africana*, Faculty of Science, University of Port Harcourt, Vol. 8, No. 1, 96-102.

Etukudo, I. A. and Umoren, M. U. (2009): A Comparison of a Modified Super Convergent Line Series Algorithm and Modified Simplex Method for Solving Quadratic Programming Problems, *ICASTOR Journal of Mathematical Sciences*, Kolkata, India, Vol. 3, No. 1, 41-61.

Francis, J. C. (1980): Investment Analysis and Management (Third Edition). McGraw – Hill, Inc., U.S.A.

Grubel, H. G. (1968): Internationally Diversified Portfolios – Welfare Gains and Capital Flows. *American Economic Review*, V58 (5), 1299 – 1314.

Gruyter, W. D. (1987): Operations Research. Theory, Techniques and Applications. Berlin, New York.

Hillier, F. S. and Lieberman, G. J. (2006): Introduction to Operations Research. Eighth Edition, Tata McGraw-Hill, New Delhi.

Macedo, R. (1995): Value, Relative Strength and Volatility in Global Equity Country Selection. Financial Analysts Journal, V51(2), 70 – 78.

Markowitz, H. M. (1952): Portfolio Selection. *Journal of Finance* 7, 77 – 91.

Markowitz, H. M. (1959): Portfolio Selection. Efficient Diversification of Investments. Cowles Foundation Monographs. Yale University Press, New Haven.

Neveu, R. P. (1985): Fundamentals of Managerial Finance. South – Western Publishing Co., Cincinative, Ohio

Parsons, J. A. (1977) The Use of Mathematics in business. Alexander Hamilton Institute Inc., U. S. A

.

Umoren, M. U. and Etukudo, I. A. (2009): A Modified Super Convergent Line Series Algorithm for Solving Quadratic Programming Problems, *Journal of Mathematical Sciences*, Kolkata, India, Vol. 20, No. 1, 55-66.

Weaver, D. (1983): Investment Analysis, Longman In Association With The Society of Investment Analysis. Great Britain.

A Distributional Analysis of out-of-pocket Healthcare Financing in Nigeria Using a New Decomposable Gini Index

[*]H. Eme Ichoku[1], William M. Fonta[1] and Abdelkrim Araar[2]

This study applies a new method of decomposing total redistributive effect of taxation proposed by Duclos et al. (2003) to assess the redistributive effects of direct healthcare financing in Nigeria. This new framework makes it possible not only to introduce into the conventional Gini Index estimation framework a flexible ethical measure of aversion to inequality but also a novel concepts of horizontal inequity and re-ranking. The empirical results indicate that when the decision to utilize healthcare is always linked to the decision to pay for healthcare, as is the case in Nigeria, out-of-pocket payment, contrary to existing literature, may indeed be progressive with high levels of horizontal inequity and re-ranking effect. But the progressivity may underlie the lack of ability to pay by poorer households. All the components of the redistributive effect are also likely to vary with the level of the social aversion to inequity

Keywords: DJA decomposition, Gini index, horizontal inequity and re-ranking.

JEL Classification: B41, C52, C81, D63, I11.

1. Introduction

It is well established in the literature that social inequalities lead to unequal health outcomes. However, it is also possible that health institutions, and in particular; the method of financing health services, could feed inequality back into the social space and exacerbate the existing inequalities and impede the social capacity to care. Methods of financing health services could, for example; give rise to unequal claims and different experiences of using the health system[3]. Health system can feed into and reinforce existing social and health inequalities (Mackintosh, 2001 and 2006). In particular health care financing system has potential to deepen and widen the existing levels of social inequalities (McIntyre et al., 2006 and Sauerborn et al., 1996). There is therefore the need for rigorous analysis of inequalities embedded in healthcare financing system in order to make explicit their implied redistributive consequences, the extent health institutions aggravate social inequalities and whether or not such redistributive consequences are justifiable on the basis of equity (Acocela, 1998). While equity in health financing has been the subject of several studies, such analyses have usually applied descriptive analysis in the analysis of equity and therefore often fail to isolate the components of inequities where they exist.

This study has two objectives. Firstly, it is an empirical study aimed at estimating the total redistributive effect of the prevailing health financing mechanism in Nigeria, namely; out-of-pocket payment (oop), and the components of this redistributive effect[4]. Secondly, it aims to test the performance of a new decompositional framework developed by Duclos, Jalbert and Araar -DJA (Duclos et al., 2003) in contrast to the prevailing Aronson, Johnson, and Lambert - AJL (Aronson et al., 1994) decomposition framework which has largely dominated the literature.

[*] This work was carried out with funding from the Poverty and Economic Policy (PEP) Research Network, which is financed by the International Development Research Centre (IDRC), the Canadian International Development Agency and the Australian Agency for International Development (AusAID). We are sincerely grateful to Jean-Yves Duclos and three anonymous reviewers whose comments and suggestions have helped greatly to improve the quality of our paper. However, all errors and omissions are the sole responsibility of the authors

[1] Centre for Demographic and Allied Research (CDAR), Department of Economics, University of Nigeria, Nsukka, Enugu State, Nigeria, (e-mails: fontawilliam@yahoo.com, hichoku@yahoo.com), +2348035408395

[2] 2190 Pavillon JA Desève, Université Laval, Ste-Foy, Québec, Canada, G1K 7P4, Phone: 418-656-7507, (e-mail :aabd@ecn.ulaval.ca)

[3] Tibandebage and Mackintosh (2001) described such an experience in Tanzania

[4] Equity in health financing is only one aspect of overall equity goal that a health system may strive to achieve (See, for e.g., WHO 2000). Other forms of equity goal include equity in access and utilization, financial risk protection etc.

2. Redistribution as a Central Policy Issue in Nigeria

Unequal access to economic and social resources is central to understanding the political economy of Nigeria. This is not only because of the increasing scope and severity of poverty in the country estimated at about 54%, but also because it is central to political instability, claims of marginalization and resource control agitations amidst, frequent conflicts in many parts of the country (Egwu, 1998). Similarly, there is also large disparity in income as recorded by the Gini coefficient estimates of about 50.6 (Adams 2004, Canagarajan et al. 1997 and Okojie et al. 2000). This has led to what is now termed politics of prebendalism[5] and state capture by entrenched interests. This system of power relation determines access to social services including health services which are generally characterized by privilege and patronage (Alubo, 2001 and 1987).

Although health expenditure in Nigeria is made at the three tiers of governments (i.e., federal, sate and local government levels), nevertheless, available statistics still show that households largely bear the responsibility of financing their health needs through direct payments (USAID-FMOH, 2009). Over 68% of total health expenditure comes from direct out-of-pocket payments (FMoH 2003 and WHO 2009,). This proportion is very high even in the context of other poorer African countries[6]. Despite increased funding to the health sector averaging over 6% of the total budget since 2003; health expenditure still lags behind the 15% of commitment to the Abuja Declaration 2000 and Gaborone Declaration 2005. In the absence of mechanisms for risk-sharing and pooling of resources for the majority of the population, households have to pay for nearly every healthcare cost directly on a ‘cash and carry’ basis. The dominance of the Nigerian health system by for-profit providers could interact with poor public financing and out-of-pocket to escalate the potential disequalizing and impoverishing effects of health care.

The key question being addressed is why should the post payment income distribution be of policy concern? The concern arises in the first place, from the fact that healthcare payments may eat so deep into the pocket of a household that it has little or nothing left to provide other basic life necessities such as food, shelter, and the education of the children (McIntyre et al. 2006, Wagstaff and van Doorslaer, 2001). Secondly, since the experience of ill-health is random among households, the cost implications of treatment would also differ in the absence of a mechanism for pooling of resources and risk-sharing. Such would be the case, for instance, where every household is expected to pay for its healthcare needs directly from out-of-pocket. In that case, the prepayment income distribution could differ significantly from the post-payment distribution. The concern here, therefore, is how far the healthcare financing mechanism worsens or improves the prepayment inequality in income distribution at the post-payment period.

3. Decomposition Method

3.1 The DJA Decomposition Framework

Achieving the objectives of this study requires the decomposition of the total redistributive effect of healthcare financing into its various components. Although decomposition methods have been developed in the context of analysis of equity effects of fiscal system, we apply the methods in the context of equity effects of healthcare financing. Following Arrow (1963) we reason that out-of-pocket healthcare payments, like the tax system, represent further deductions from household income, albeit, more idiosyncratically. Indeed, its effects on household are less systematic than the tax system since ill-health and the cost of treatment are stochastic distributions.

[5] Politics of prebendalism is defined as the intense struggle by interest groups to capture state resources for group welfare only (Ayogu, 1999). The term ‘Prebendalism’ was originally coined by Joseph (1987), with particular reference to the Nigerian political economy (see http://en.wikipedia.org/wiki/Prebendalism)

[6] In most OECD countries, the fraction of direct healthcare payment to total payment is generally below 25%. US and Ireland with 22% and 23% respectively, have some of the highest proportions, while Germany has 9.3%, UK has 17% and Denmark has 14%.

The AJL approach to decomposition of total redistributive effects of taxation and transfer programs into vertical, horizontal and re-ranking effects has dominated the literature since it was first developed. However, the AJL approach has certain weaknesses that make the consideration of alternative decomposition frameworks necessary. Foremost among the weaknesses is the implicit assumption of a rigid ethical social welfare function that is insensitive to the policy-maker`s concern or level of social aversion to inequality[7].

To overcome the problem of ethical rigidity associated with the AJL approach as well as other shortcomings, DJA propose the use of an inequality index with two flexible ethical parameters of aversion to inequality (Duclos and Araar, 2006). One can recall here that the flexible forms of the generalised Gini index was already discussed by Yitzhaki (1983). The DJA single parameter, ethically sensitive weighting scheme uses differences in income rank of individuals. A brief summary of the framework is presented here in other to highlight the conceptual assumptions behind the estimated parameters in the empirical model.

For $p \in (0,1)$ denoting the percentile or fractional rank of an individual in the income distribution, a non-negative weight v may be defined such that:

$$w(p,v) = v(1-p)^{(v-1)}, \quad v \geq 1 \tag{1}$$

Where v is a parameter reflecting social preferences about aversion to inequality using differences in the ranks of individuals in the income parade. Equation (1) has the following restrictions:

(i) $$\int_0^1 w(p)dp = 1 \tag{2}$$

That is, the weights are normalized to 1. Furthermore,

(ii) $$w(p_i) \leq w(p_j) \qquad for \quad p_i \geq p_j \tag{3}$$

In other words, the weights are sensitive to the individuals' ranks in the income distribution such that if a small transfer is made from a richer to a poorer person, inequality is perceived to be reduced (the Pigou-Dalton transfer principle). For $v = 2$, equation (1) reduces to the ethical weights of the standard Gini index which measures income disparity within a region or group of people. For values of v less than 2, the policy maker is assumed to prefer inequality favouring the rich while for values of v greater than 2, the policy maker is averse to inequality that disadvantages the poor.

DJA model specifies a social welfare function that is concave, additive, and linear in levels of income and that can generate relative inequality indices as:

$$W_X(\varepsilon, v) = \int_0^1 U_\varepsilon(X(p))w(p,v)dp \tag{4}$$

Equation (4) is the Social Welfare Function (SWF) of the gross income (X) where U_ε is the Atkinson (1970) concave utility function. The parameter ε is an index of aversion to uncertainty in post-payment income of those within any given income level y. In other words, ε may be interpreted as a measure of aversion to horizontal inequity. One can recall here that, with the fiscal system framework, the concept of horizontal inequity refers to the unequal treatment of equals. DJA uses the concept of Equally Distributed Equivalent Income (EDE) as in Atkinson (1970) in order to analyze the cost of inequality to the society.

[7] See for example, Wagstaff (2005), Gerdtham and Sundberg (1996) , Wagstaff et al. (1999) van Doorslaer et al. (1999) , Wagstaff and van Doorslaer (2001) , Wagstaff (2001), for the applications of this framework and Lambert and Ramos (1997a; 1997b) for a modified version.

Since EDE is determined by the level of ε, the inequality index can be expressed in the form (Atkinson 1970, Blakorby and Donalson, 1978):

$$I_X = 1 - \frac{\xi_X}{\mu_X} \tag{5}$$

Equation (5) expresses the cost of inequality in terms of proportion of total income. It is the fraction of total income that could be used to restore equality or remove existing inequality without loss in social welfare. By implication, if risk or uncertainty in the postpayment income increases (reduces), the ratio $\frac{\xi x}{\mu x}$ falls (rises) and thus I_x (inequality) rises (falls) requiring more (less) level of equally distributed income (relative to the mean of the actual distribution) to achieve the same level of social welfare as before. This concept of EDE is central to the DJA model estimation of horizontal inequity and is sensitive to the parameter ε.

Inequality can similarly be defined for post-payment income to obtain (I_N), for expected post-payment income (I_N^E), and for expected net income utility (I_N^P). These relationships provide the key to understanding the DJA decomposition framework:

$$\Delta I = \underbrace{I_X - I_N}_{RE} = \underbrace{I_X - I_N^E}_{V} - \underbrace{(I_N^P - I_N^E)}_{H \geq 0} - \underbrace{(I_N - I_N^P)}_{R \geq 0} \tag{6}$$

Like in the case of the AJL framework, the total redistributive effect ΔI is composed of three main parts: the vertical equity (V), horizontal equity (H) and reranking effect (R). V is the conventional index of progressivity of payments. The index of horizontal inequity (H) aggregates the over all increase in inequality that arises due to unequal payment by those at the same level of prepayment income. It is assumed that at any fixed point in the distribution of prepayment income, say point x, there is a group of individuals having exactly income x at prepayment distribution who can be denoted as W_x This is the group of prepayment equals at point x. This is different from what is done in the AJL framework where 'income equals' could more appropriately be regarded as 'near equals'. Horizontal inequity at point x is the measured level of inequality induced among the group W_x located at point x by the payment system. An individual who belonged to this class of prepayment income may now find her income at point x_1 or point x_2 due to the fact that the payment system has treated equals unequally. This is a source of uncertainty and risk factor. The average post-payment income among the set of prepayment equals with x is μ_x^a. However, a risk-averse individual may prefer to have ξ_x^a with certainty rather than take a gamble that may give her x_1 at the worst and x_2 at the best with respective probabilities. Thus, the H index measures the risk premium:

$$H_x = \mu_x^a - \xi_x^a \tag{7}$$

For the policy-maker, H_x represents the amount that has to be given up among those with post-payment income W_x so as to remove the uncertainties in the post-payment income of this group without loss of welfare.

To obtain the global measure of H involves aggregation of the form: $\sum H_x \phi_x, \quad \forall x \in R$ and where ϕ_x denotes the population share with income x. In the DJA framework, like in Lambert and Ramos (1997b) but unlike the AJL framework, the weight chosen is pure weight: the proportion of the population at the given x. In other words, the weighting factor is independent of the income at point x. This is important because it ensures that horizontal inequity at point x is not contaminated by vertical considerations as is the case with the AJL framework where the weighting factor is the product of the population and income shares at point x. R index, measures the reranking effect. R arises because individuals may move out of their prepayment income class to other classes due to effects of payment. It arises as a result of any changes in rank induced by the healthcare payment system (or fiscal policy).

The DJA approach estimates *H* as loss in social welfare arising from the fact that the payment system generates uncertainty and is a source of dis-utility. Inequality itself is a loss of utility to the social welfare because it generates resentment and a feeling of deprivation among peers (Runciman, 1966). It is the cost of these resentments that is captured by *H*.

The empirical estimation of the DJA framework is based on the Gaussian kernel function which is nonparametric because of the well known properties of the Gaussian distribution. For example it does not need apriori assumptions about the distribution of income of the sample population. However, it seems that, in general, the choice of the kernel function is not as important as the choice of the window width which determines the smoothness of the distribution (Yatchew, 1998, Silverman 1986, *Stata Reference Manual*, 2002). By adopting this statistical approach, therefore, the DJA framework transfers the normative decision of determining income equals from the decision maker to statistical exercise: the choice of the window width is now determined by the optimal trade-off between bias and minimization of the squared mean error. The only assumptions required are statistical assumptions such as the smoothness and continuity of the joint distribution of gross and net incomes.

3.2 Social Determinants of Values of ε and v

For direct health payments, the values of ε and v may be determined by, among other things, the health system bureaucracy. The cost of removing such inequalities are suggested to lie between 0.25 and 1.0 of the total amount, implying that ε may lie between this range (i.e., $0.25 \le \varepsilon < 1$). More efficient health systems would have the value of ε near to the minimum while inefficient system would have values of ε close to 1. On the same basis, the social value of *v* is suggested to range between 1 and 4. In the estimation that follows we follow the Duclos et al. (2003), example where they used the value of $\varepsilon = 0.4$. But we also estimate for $\varepsilon = 0$, which amounts to the assumption of the null hypothesis: $H_{DJA} = 0$; where H_{DJA} refers to the value of horizontal inequity obtained under the DJA framework. As was noted earlier on, the DJA approach generalizes the value of *v* implied by the standard Gini index which is $v = 2$. For the following estimations we use a broad range of values of *v* from 1.5 to 5.0.

4. The Data

The study used a cross-sectional survey data generated between April and August 2004. The data set was part of a large set of data aimed at generating information on a wide range of social welfare issues including household health-seeking behavior, general household welfare and access to social services, health financing, self-assessed health, among others. The absence of such vital health statistics necessitated a field survey in order to generate a fresh set of data that could be used to inform health policy in the state. However, instead of a national survey, the survey was limited to one of the 36 states in Nigeria, Enugu state with a population of about 3 million in 2004. The selection of the state was among other things, based on cost considerations.

Also, familiarity of the researchers with the terrain of the state was an added impetus as this knowledge assisted in no small measure in the survey design and its actual execution. Furthermore, the state mirrored in many ways the health problems of most states in Nigeria: heavy disease burden, heavy out-of-pocket financing, dominance of the small-size private for-profit health facilities, and general lack of purchasing power among the population, among others. The state tier of governance also plays a core role in health policy decisions as it has responsibility for primary and secondary care in Nigeria.

The actual field survey was conducted after series of preparation that included design of questionnaire, ethical approval from the Enugu State Ministry of Health, training of fieldworkers, and pre-testing of the survey instrument. The standard multi-stage sampling design was adopted. The entire state was stratified along urban-rural divide. Pre-existing clustering arrangements used by the Federal Office of Statistics (FOS)[8] and the National Population Commission (NPC) were adopted and this provided the frames not only for the clusters but also for the

[8] The Federal Office Statistics has been reorganized and renamed Nigerian Bureau of Statistics (NBS)

households. The urban and rural clusters served also as the enumeration areas (EAs) or primary sampling unit (PSUs). One hundred EAs were selected at random and sampled intensively.

The decision on the sample size of 1500 households for this study was largely guided by Yamene (1967) in which the sample size depends both on the size of the population and the researcher's acceptable margin of error. For this survey, given that the average household size in the state was about 4 (Nigeria Demographic and Health Survey - NDHS, 2003), the margin of error was fixed for 0.01%. In other words, our confidence interval was more than 99%. Each EA was made up of approximately 20 households. Fifteen households were selected from each cluster making a total of 1500 households with about 5814 individuals.

The heads of household or, in their absence, their spouses were the main respondents on household level questions however; adult household members were also required to respond to individual level questions. The main variables used for the estimation include household gross expenditure defined as the total expenditure of the household inclusive of healthcare expenditure in the four weeks preceding the interview. Current literature suggests that expenditure is a better reflection of a household's permanent income (Bollen et al. 2001; Deaton 1997; Hentschel and Lanjouw, 1996 and Stevenson et al., 1988). The healthcare expenditure variable was defined as the cost of healthcare to a household in the four weeks preceding the interview. This includes cost of treatment, cost of consultation card, transportation to and for the treatment facility, and any other incidental expenses connected with ill-health in the household. The net households expenditure is the gross expenditure defined above net of healthcare costs. While there are arguments for considering household scale economies in studies of this nature, there are equally strong arguments against (Deaton, 1997). This study uses households' per capita expenditure for the estimations and the results that are presented in subsequent section.

In consideration of the multi-stage sampling design and the potential negative effect of improper weighting on point estimates, and the effects of stratification and clustering on estimated standard errors, the data analysis adopted the bootstrap method in Distributive Analysis software (DAD). The results, therefore, provide information also on bootstrapped asymptotic standard errors.

4.1 Empirical Results

4.1.1 Descriptive Statistics

The total number of households interviewed was 1500; three households were dropped because of inconsistent and incomplete information. Thus, the analysis is based on 1497 representing almost 100% of the total. About 65% of the respondents in respect of household questions was male and about 20% was urban households. In analyzing the effect of household health care expenditure we used the representative individual in each household. 522 representing 35% of the total sample reported financing healthcare during the reference period. The headcount poverty index on prepayment and based on updated poverty line in Aigbokan (2000) was 57.3% and the postpayment headcount poverty was about 61.4%, implying that out-of-pocket increased poverty by about 7%. The sample shows that 38% of poor households incurred out-of-pocket while 30% of the non-poor reported out-of-pocket during the period. Mean out-of-pocket for different quintiles for the sick-only and for all households are shown in Figure 1 while Figure 2 reports the percentage of total household expenditure spent as out-of-pocket by different income quintiles of households. However, as would be expected, the mean amount spent on healthcare financing by the poor and non poor differed widely. The poorest 75% of the households spent on average N342.4 while the top richest 5% of households spent on average N2389.40 or 700% of the amount by poorest 75% of households.

Figure 1 indicates that among households that reported incurring out-of-pocket the poorest quintiles spent less than $1.00 on health care while the richest quintile spent nearly $40 per capita on health in the preceding 4 weeks of the interview. The graph is flatter when the populations that reported no out-of-pocket were included. Figure 2 similarly shows that relative burden of out-of-pocket measured by the proportion of total household income spent by each quintile as out-of-pocket and indicates that the rich spend the larger proportion of their total expenditure as out-of-pocket.

This seems to indicate that out-of-pocket could indeed be progressive; however, when the health needs of the various quintiles are taken into account, it would seem to show that the poorer households were grossly under spending on health, probably because they could not afford the cost of care.

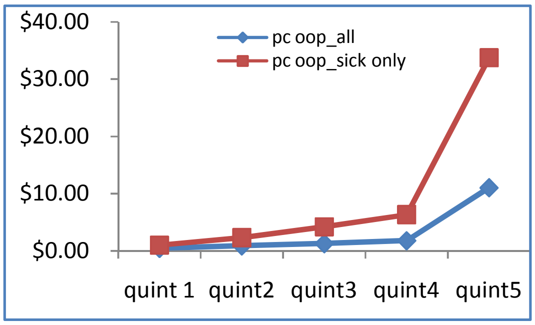


Figure 1: Per capita out-of-pocket (all) and per capita out-of-pocket (sick only)

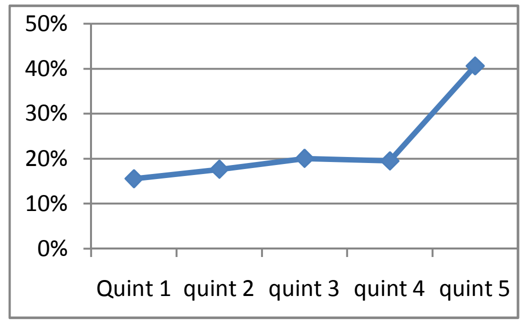


Figure 2: Distribution of out-of-pocket as % of per capita expenditure (sick only)

4.1.2 DJA Decomposable Results

Results of the DJA Model decomposition based on an assumed value of aversion to uncertainty in net payment income $\varepsilon = 0.4$ and aversion to inequality $v = 1.5, 2, 3, 5$ are presented in Table 1.

Table 1: Decomposition Results Based on DJA Framework ($\varepsilon = 0.4$)

Indices	$\varepsilon = 0.4,\ v = 1.5$	$\varepsilon = 0.4,\ v = 2$	$\varepsilon = 0.4,\ v = 3$	$\varepsilon = 0.4,\ v = 5$
G_X	**0.3807****	**0.4966****	**0.6145****	**0.7150****
	0.1173	0.1498	0.1849	0.2169
G_N	**0.3829****	**0.4999****	**0.6201****	**0.7224****
	0.1182	0.1503	0.1871	0.2199
RE	**-0.0033**	**-0.0033**	**-0.0056**	**-0.0075**
	0.0055	0.0040	0.0054	0.0071
V	**0.0284****	**0.0334****	**0.0398****	**0.0509****
	0.0102	0.0120	0.0133	0.0158
H	**0.0202****	**0.0253****	**0.0342****	**0.0473****
	0.0059	0.0076	0.0104	0.0144
R	**0.0103****	**0.0114****	**0.0111****	**0.0111****
	0.0031	0.0036	0.0047	0.0046

NB. Bootstrapped standard errors at the bottom of estimated values, ** Statistically significant at 5% level

Column 1 of Table 1 presents the parameters. Columns 2 through 5 report the estimated parameters under the different simulations around the values of ε and v and their respective bootstrapped standard errors. Since the definition of income band is statistically determined in the DJA framework, it obviates the need for the arbitrary definition of income bands which is one of the major concerns in the AJL framework. The estimated prepayment Gini Index for $v = 2.0$ represents the implicit value assumption in the conventional Gini Index. Under this assumption the policy maker is inequality-neutral. It does not matter whether inequality favours the rich or the poor. The estimated inequality is close to the estimated Gini Index of 51 obtained by Okojie et al. (2000) and 50.6 by Adams (2004). When $v = 1.5$, the policy maker in fact favours inequality and so perceived less inequality as reflected in the Gini Index value of 0.38. On the other hand, when the policy maker is averse to inequalities disadvantaging the poor as would be the case when $v \geq 0$, then she/he perceives more inequality for every given value of ε. This is reflected in the higher estimated prepayment Gini Index (G_X) as the value of v increases beyond 2. The same result holds for the estimated postpayment Gini Index ($\boldsymbol{G_N}$). For all the values of v (given ε), it is observed that despite their negative magnitudes, the overall redistributive effects are actually not different from zero because, as noted earlier in the descriptive statistics, richer households actually seem to devote relatively higher proportion of their total expenditure on out-of-pocket. This also accounts for the statistically insignificant values of vertical inequality particularly at low values of v. On the other hand the estimated values of horizontal inequality and reranking effects are statistically significant for all the values of v.

Tables 1 and 2 also show that under a constant value of ε, the estimated values of G_X, G_{X-T} and the absolute values of RE increase consistently as the value of v increases from 1.50 to 5.0. In none of the values of v within the tabulated range is the redistributive effect significant. However, the increasing negative value of RE shows that as pro-poor aversion to inequality increases the redistributive effect becomes more pro-rich. In other words, the more sensitive or 'equality-minded' the policy-maker is the more pro-rich the redistributive effect appears. Similarly vertical and horizontal inequities each increase with increases in the value of v keeping the value of ε constant. In all cases, vertical inequity is higher than horizontal inequity and reranking effect. This indicates that vertical inequity is a more serious problem in the population. But the rate of increase in the ratio V/H decreases with increase in the value of v (keeping $\varepsilon = 0.4$); though this decrease is not very steep. This is shown in Fig 3. Surprisingly, and contrary to other studies (see, for example; van Doorslaer et al. 1999, Wagstaff and van Doorslaer 2001), it is found that vertical equity is positive and significant, implying that out-of-pocket can at times and in different contexts be progressive and redistributive in its effect but this is at a cost as shall be discussed later. This result was already anticipated in the descriptive statistics, as illustrated by Figures 1 and 2 where it was shown that the rich spend higher proportion of their total expenditure on health than the poor.

Table 2: Decomposition Results Based on DJA Framework $(\varepsilon = 0)$

Indices	$\varepsilon = 0,\ v = 1.5$	$\varepsilon = 0,\ v = 2.0$	$\varepsilon = 0,\ v = 3.0$	$\varepsilon = 0,\ v = 5.0$
$\boldsymbol{G_X}$	**0.3104****	**0.4474****	**0.5830****	**0.6970****
	0.0909	0.1349	0.1750	0.2088
$\boldsymbol{G_N}$	**0.3105****	**0.4482****	**0.5863****	**0.7025****
	-0.0915	-0.1349	-0.1764	-0.2108
RE	**-0.0001**	**-0.0008**	**-0.0033**	**-0.0056**
	0.0054	0.0054	0.0061	0.0056
V	**0.0186****	**0.0250****	**0.0314****	**0.0430****
	0.0092	0.0111	0.0119	0.0150
H	**0.0056****	**0.0119****	**0.0221****	**0.0372****
	-0.0019	-0.0036	-0.0068	-0.0112
R	**0.0131****	**0.0139****	**0.0126****	**0.0113****
	-0.0045	-0.0054	-0.0044	-0.0044

NB. Bootstrapped standard errors at the bottom of estimated values, ** statistically significant at 5% level

Tables 1 and 2 also show that under a constant value of ε, the estimated values of G_X, G_{X-T} and the absolute values of RE increase consistently as the value of v increases from 1.50 to 5.0. In none of the values of v within the tabulated range is the redistributive effect significant. However, the increasing negative value of RE shows that as pro-poor aversion to inequality increases the redistributive effect becomes more pro-rich. In other words,

the more sensitive or 'equality-minded' the policy-maker is the more pro-rich the redistributive effect appears. Similarly vertical and horizontal inequities each increase with increases in the value of v keeping the value of ε constant. In all cases, vertical inequity is higher than horizontal inequity and reranking effect. This indicates that vertical inequity is a more serious problem in the population. But the rate of increase in the ratio V/H decreases with increase in the value of v (keeping $\varepsilon = 0.4$); though this decrease is not very steep. This is shown in Fig 3. Surprisingly, and contrary to other studies (see, for example; van Doorslaer et al. 1999, Wagstaff and van Doorslaer 2001), it is found that vertical equity is positive and significant, implying that out-of-pocket can at times and in different contexts be progressive and redistributive in its effect but this is at a cost as shall be discussed later. This result was already anticipated in the descriptive statistics, as illustrated by Figures 1 and 2 where it was shown that the rich spend higher proportion of their total expenditure on health than the poor.

It is to be noted that contrary to what the theory suggests under the DJA assumptions, namely that when $\varepsilon = 0,\ H \to 0,$ we still find that the estimated values of H_{DJA} under the different values of v are still statistically different from zero. Thus, it seems that there is a difference between what is predicted by theory and what is obtained in actual implementation of the model[9].

Comparing Tables 1 and 2, we find that for all values of v, the value of H is greater in Table 1 with $\varepsilon = 0.4$ than in Table 2, with $\varepsilon = 0$. This is obviously so because the cost of uncertainty in post-payment income distribution given prepayment rank increases with increases in the value of aversion to uncertainty in the post-payment income measured by ε. This result can easily be seen from equation (5) and by noting the inverse relationship between ε and x. That is, the higher the level of uncertainty in net income the less the EDE. Higher values of ε are associated with higher values of H.

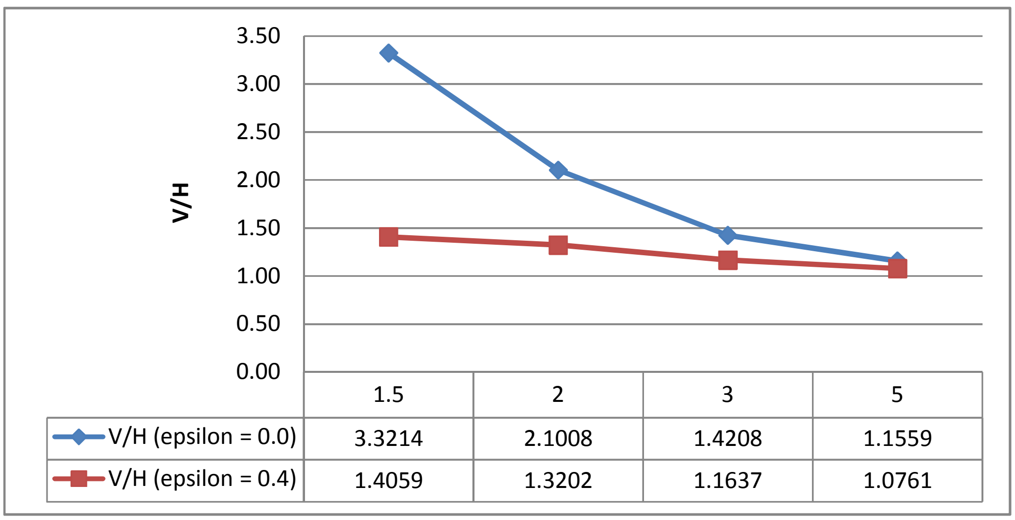


Figure 3: The Ratio between vertical and Horizontal Inequities (V/H) ($\varepsilon = 0$ and $\varepsilon = 0.4$)

Furthermore, setting $\varepsilon = 0$, is a test of the null hypothesis that $H = 0$. This constraint is rejected as can be seen from Table 2. It is to be noted that the constraint becomes increasingly non-binding as the value of v increases. This perhaps explains why the value of the ratio V/H in Table 2 declines more dramatically than in Table 1 (see Fig. 3). The relations between V and H is exhibited when the value of ε is constrained to zero as in Table 2. The divergence between the values of V and H is even more significant at lower values of v for $\varepsilon = 0$. However, it seems likely that increases in the value of ε must be accompanied by not only increases in the value of horizontal inequity but also increases in the value of V if marginal increases in the payment rate are also accompanied by higher levels of deadweight loss or costs of inefficiency associated with ε.

[9] The attention of authors has been drawn to this apparent divergence between what is predicted by the theoretical model and what is obtained in actual implementation

What can be the policy implications of significance of the two components H and V? In clear way, the higher level of horizontal inequity confirms the idiosyncratic nature of health shocks for households that have practically the same level of prepayment wellbeing. It follows that the promotion of mechanism that can help to share the risk of health shocks according to household capabilities, must contribute significantly in reducing inequality coming from this stochastic distribution of health shocks. The higher level of vertical inequity confirms the need to empower poor households to expend relatively more on health-producing goods. To treat this form of inequality, the design of policies targeting the poor in the use of health services such as exemptions and deferrals should be appropriate. For instance, among the policies that can deal with these two inequity components at the same are those designed to promote or to finance partially the heath risk sharing institution such community and social health insurance in poor countries or regions.

The value of R appears rather constant with increases in the value of v. This is consistent with the DJA theoretical framework: the larger the value of v the more weight that is given to "the reranking resentment of the poorest" (Duclos et al., 2003). But giving more weight to the resentment of the poorest need not necessarily increase the reranking. Therefore, in spite of the increase in weight given to the reranking resentment of the poorest, the reranking index R changes but very little. Irrespective of this sluggish movement in R in response to increases in v, it is clear from the results in both Tables 1 and 2 that reranking is a major effect that follows direct healthcare payment. In fact at the lower values of v in Table 2, R has higher values than H. The question of which of these is a more serious problem to a society may depend on social values. It is clear, nevertheless, that reranking is a major problem in healthcare financing through direct payment in Nigeria.

5. Conclusion

The main conclusions from this study could be summarized as follows: It is possible to extend the decomposition frameworks developed in the context of fiscal studies to the analysis of equity effects of healthcare financing. The DJA decomposition framework has been successfully applied in this study to decompose the equity effects of healthcare financing in a developing country. Unlike the AJL model, this framework overcomes the problem of arbitrariness in defining income equals. It also offers the possibility that the healthcare financing policy maker can build into policy considerations the level of social aversion to inequality.

A healthcare financing system that is dominated by direct out-of-pocket payment has strong equity implications. It may be progressive or proportional but such progressivity or proportionality underlies the fact that some of the poor may avoid the use of health services because they cannot pay the cost. They therefore have unmet health need which may not be captured by the distributional analysis since such analysis focuses on effects of payment on prepayment distribution. Some others may be using health services at the cost of great displacement effects for other critical household needs. Thus, progressivity of out-of-pocket may be bought at a high opportunity cost. In an environment where the majority of the people are poor, the system of direct healthcare financing may lead to the exclusion of the majority of the people from the use of healthcare facilities, given that in such a system a household's decision to use healthcare service implies the decision to finance healthcare. Therefore, it is possible that the Nigerian healthcare system is excluding proportions of the population from the use of healthcare on the grounds that they cannot afford the cost of treatment.

The results further show very mild vertical inequity under the v = 2, which is implicit in the conventional Gini Index. The perceived level of inequity becomes significant as the policy maker's sensitivity to inequality increase with increasing value of v. However, the levels of inequity vary with the level of the social aversion to inequity as well as the level of aversion to uncertainty in the post-payment period. The higher the level of pro-poor aversion to inequities, the more visible the vertical and horizontal inequities appear. While the elimination of all these forms of inequities and reranking is important, it is likely that policy instrument that addresses one may exacerbate the other. For example, a policy aimed at eliminating vertical equity may worsen reranking or horizontal equity. Thus, it seems that in the final analysis, there is need to prioritize among these conflicting objectives depending on the elimination of which of these would contribute more to the improvement of social welfare. This is a normative question. It is possible for example, that one society might consider reranking a more serious form of inequity because it is not based on productive criterion while another society may consider vertical equity the more serious because it worsens the ante level of inequity based on the random misfortune of falling ill.

The DJA framework provides an alternative interpretation of horizontal inequity as different from what prevails in literature. H_{DJA} is an index of the risk premium an economic agent in a given prepayment income class is willing to pay to prevent the uncertainty that accompanies the transition from pre- to post-payment income distribution that arises from the intervention of the fiscal system. Thus, H_{DJA} seems to bring into sharper focus, the random effects of healthcare financing in a system dominated by direct healthcare payments. The higher the level of random effects in healthcare financing, the higher the risk premium a risk adverse policy maker is willing to pay to obtain the certainty equivalent in the post-payment income and, therefore, the more costly to the social fabric of such arbitrariness. This clearly suggests that prepayment schemes, and health insurance schemes that lessen the uncertainties associated with healthcare financing system are more likely to be preferred by healthcare financiers to the prevailing system of health financing through out-of-pocket. The cost to the social fabric might be higher if individuals compare themselves with those with similar productive characteristics but who have been better treated by the payment system. Thus, apart from introducing a flexible ethical parameter into the decomposition of the Gina index, the DJL framework also provides a novel interpretation of the concept of horizontal inequity.

Note finally that the proposed theoretical implementations of the DJA model in the context of financing healthcare services are made to contribute to enrich the empirical studies tools and to derive interesting results. This may assist policymakers in shaping policies designed to fight simultaneously the different negative aspects of distribution.

References

Adams, R.H. (2004), ‘Economic Growth, Inequality and Poverty: Estimating the Growth Elasticity of Poverty’, World Development 32(12):1989–2014

Acocellla, N. (1999), *The Foundation of Economic Policy: Values and Techniques*, Cambridge University Press, UK.

Aigbokhan B.E. (2000), *Poverty, Growth and Inequality in Nigeria: a Case Study*, African Economic Consortium, Nairobi.

Alubo, S.O. (1987), 'Power and influence in medical care: an analysis of medical practice in post colonial Nigeria' *Social Science & Medicine*, **24,** (5), 453-462.

Alubo, S. O. (2001), 'The promise and limits of private medicine: health policy dilemmas in Nigeria' *Health Policy and Planning*, **16** (3), 313-321.

Aronson, J. R., Johnson, P. and Lambert, J. (1994), 'Redistributive effect and unequal income tax treatment' *The Economic Journal,* **104** (423), 262-270

Arrow, K. J. (1963), 'Uncertainty and the welfare economics of medical care' *Social Science & Medicine,* **44** (6), 773-787

Atkinson, A. B. (1970), 'On the measurement of inequality' *Journal of Economic Theory,* **2**,244-263.

Ayogu, M. D. (1999), 'Before prebendalism: a positive analysis of core infrastructure investment in a developing fiscal federalism' *African Development Review,* **11** (2), 169-198

Bollen K., Glanville J., Stecklov G. (2001) ‘Socioeconomic status and class in studies of fertility and health in developing countries’, *Annual Review of Sociology* 27: 153-185

Canagarajan, S., Nwafor, J. and Thomas, S. (1997), ‘The evolution of poverty and welfare in Nigeria 1985-1992’, Washington DC, World Bank Mimeo.

Deaton, A. (1997), *Analysis of Household Surveys: A Microeconomic Approach to Development Policy*, The Johns Hopkins University Press, Baltimore.

Duclos, J. Y. and A. Araar (2006): Poverty and Equity Measurement, Policy, and Estimation with DAD, Ottawa Cairo Dakar Montevideo Nairobi New Delhi Singapore: Springer.

Duclos, J. Y., Jalbert, V. and Araar, A. (2003), 'Classical horizontal inequity and reranking: an integrating approach' *CIRPEE Working paper 03-06*, University of Laval, Canada

Egwu, S. (1998), *Structural Adjustment, Agrarian Change and Rural Ethnicity in Nigeria*, Nordiska Afrikanstitutet, Uppsala

Federal Ministry of Health (2003) Priorities for Action within the Context of Health Sector Reform Agenda, Federal Ministry of Health, Abuja

Gerdtham U-G and Sundberg, G. (1996), 'Redistributive effects of the Swedish health care financing system' Stockholm School of Economics Working Paper No. 113

Hentschel, J. and Lanjouw, P. (1996), 'Constructing an indicator of consumption for the analysis of poverty' *LSMS Mimeo*,.The World Bank., Washington D.C

Lambert, P. J. and Aronson, J. (1993), 'Inequality decomposition analysis and the Gini coefficient revisited' *The Economic Journal*, **103** (Sept), 1221-1227

Lambert, P. J. and Ramos, X. (1997a), 'Vertical redistribution and horizontal inequity' *International tax and Public Finance*, **4**,25-37.

Lambert, P. J. and Ramos, X. (1997b), 'Horizontal inequity and reranking' In D. Slottje and Zandvakili (eds), *Inequality and Taxation*, JAI Press, Greenwich.

Mackintosh, M. (2001), Do health care systems contribute to inequalities? In Leon D. and Walt G (eds.,) Poverty Inequality and Health: An International Perspective, Oxford: Oxford University Press

Mackintosh, M. (2006), Commercialization, inequality and the limits to transition in health care: a Polanyian framework for policy analysis, *Journal of International Development*, 18:393-406

McIntyre, D., Thiede, M, Dahlggren G, and Whitehead, M. (2006), 'What are the economic consequences for households of illness and of paying for health care in low- and middle-income country contexts? *Social Science & Medicine*, 62 (2006) 858–865

Narayan, D., Patel, R., Schafft, K., Rademacher, A. and Koch-Schulte, S. (2000), *Voices of the Poor: Can Anyone Hear Us?*, Oxford University Press, New York.

National Population Commission of Nigeria, (2003), *Nigeria Demographic and Health Survey 2003*, National Population Commission and ORC/Macro, Calverton Maryland

Nwosu, E. J. (2000), *The Challenge of Poverty in Africa*, Skillmark Media Ltd., Owerri.

Okojie, C. E., Ogwumike, F. O., Anyawu, J. C. and Alayande, B. A. (2000), 'Poverty in Nigeria - an analysis of gender issues, access to social services and the labour market' A paper presented at AERC conference in Nairobi - Kenya.

Runciman, W.G. (1966), *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*, University of California Press, Berkeley and Los Angeles.

Sauerborn, R., Nougtara, A., Hien, M., & Diesfeld, H. J. (1996), Seasonal variations of household costs of illness in Burkina Faso. Social Science and Medicine, 43(3), 281–290.

Silverman, R.W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

STATA Corporation, (2002), *Reference Manual*, STATA Corporation, Texas.

Stevenson, A., Muscatelli, V. and Gregory, M. (1988), *Macroeconomic Theory and Stabilization Policy*, Philip Allan, New York

Tibandebage, P. and Mackintosh, M. (2001), The market shaping of capabilities in low income contexts and some implications for social policy: liberalized health care in Tanzania, Paper presented at the conference on Justice and Poverty: Examining Sen's Capability Approach, Von Hugel Institute, Cambridge, 5-7 June 2001

USAID-FMOH (2009) Nigeria Health System Assessment 2008, Abuja

van Doorslaer, E., Wagstaff, A., vanderBurg, et al., (1999), 'The redistributive effect of health care finance in twelve OECD countries' *Journal of Health Economics*, **18**,291-313.

Wagstaff, A. (2001), 'Measuring equity in health care financing' World Bank. Mimeo, The World Bank, Washington D.C,

Wagstaff, A., and van Doorslaer, E. (2001), 'Paying for health care: quantifying fairness, catastrophe and impoverishment with applications to Vietnam 1993–98', World Bank Policy Research Working Paper No. 2715 Washington, DC: World Bank.

Wagstaff, A., van Doorslaer, E., vanderBurg et al., (1999), 'Redistributive effect, progressivity and differential tax treatment: personal income taxes in twelve OECD countries' *Journal of Health Economics*, **17**, 73-98.

Wagstaff A. (2005) 'Decomposing changes in income inequality into vertical and horizontal redistribution and reranking, with applications to China and Vietnam' World Bank Mimeo, The World Bank, Washington D.C.

WHO (2009) World Health Statistics, www.who.int/whosis/whostat/EN_WHSO (Table7.pdf, Accessed Nov. 11, 2009

WHO (2005), The *World Health Report 2005: Make Every Mother and Child Count*, World Health Organization, Geneva.

WHO (2000), The World Health Report 2000: Health Systems: Improving Performance, World Health Organization, Geneva.

Yamane, T. (1967) *Statistics, An Introductory Analysis*, 2nd Ed., New York: Harper and Row.

Yatchew, A. (1998), 'Nonparametric regression techniques in economics' *Journal of Econometric Literature*, **XXXVI** (June 1998), 669-721

Yitzhaki, S. (1983), 'On an extension of the Gini inequality index' *International Economic Review*, **24**,617-628

Factors, Preventions and Correction Methods for Non-Response in Sample Surveys

GODWIN NWANZU, AMAHIA[1]

Missing survey data occur because of unit and item non-response. This is practically independent of the method of data collection. As a result of the bias that non-response sometimes introduces in survey estimates, identifying factors that promote it, and taking measures of prevention and correction methods are clearly necessary. The standard method to compensate for unit non-response is by weighting adjustment, while item non-responses are handled by some form of imputation. This paper reviews factors that give rise to non-response and the corresponding methods used for its prevention and control. It also discusses their properties.

Keyword: Non-response; Unit non-response; Weighting adjustment; Imputation

1.0 Introduction

Surveys usually collect responses to a large number of items for each sampled unit. One of the most obvious problems in surveys is the inability to collect responses on some or all of the items for a sampled unit or when some responses are deleted because they fail to satisfy edit constraints. This is called the problem of non-response. It indicates a clearly visible "flaw" in the survey operation and has important implications during design and analysis. This is because the sample respondents alone do not validly depict the population investigated and analysis based on respondents may result in misleading inference. It is common practice to distinguish between unit non-response when none of the survey responses are collected for a sampled unit, and item non-response when some but not all of the responses are available. Unit non-response arises because of refusals, inability to participate, not-at-homes, units closed, away on vacation, unit vacant or demolished, and untraced units. Item non-response arises because of item refusals, "don't knows", omissions and answers deleted in editing.

This paper identifies factors that promote survey non-response and reviews the methods available for handling it. The distinction between unit and item non-response is useful in this paper since different adjustment methods are used for these two cases. Generally, the only information available about unit non-respondents is that on the sampling frame from which the sample was drawn. For example, in a two-staged stratified sampling scheme, the primary sampling units, secondary sampling units and the strata in which the non-respondents are located are important. The importance of this information is usually incorporated into weighting adjustments that attempt to compensate for the missing data. As a rule, weighting adjustments are used for unit non-response. In the case of item non-response, a great deal of additional information is available for the element involved. Responses for other survey items are available, in addition to information from the sampling frame. In order to retain all survey responses for elements with some item non-responses, the usual adjustment procedure produces analysis records that incorporate the actual responses to items for which the answer were acceptable and inputed responses for other items.

1.2 Reasons for Non-Response

Reasons explaining why units fail to respond in a survey are often reported, although the words used to describe them may vary. Terminology here seems to depict the type of units being studied and the mode of data collection used in the survey. Durbin (1954) and Kish (1965) discuss some of the general reasons for non-response in household surveys. Research has found that three types of unit non-response have distinctive causes and, for many surveys, distinctive effects on the quality of survey statistics. These are failure to deliver the survey request, refusal to participate in the survey, and inability to participate in the survey.

1.3 Non-Response due to Failure to Deliver the Survey

Non-response due to non-contact or failure to deliver the survey request misses the sample persons whose activities make them unavailable in the specific mode of data collection. The key concept here is the "contactability" of sample units. That is, whether the sample unit is accessible to the survey researcher. In figure 1 below, we present a basic diagram of the influences acting on the contactability of sample units in a survey. In household surveys for

[1] Department of Statistics, University of Ibadan, Nigeria

example, if the researcher knows when people are at home and accessible, successful contact would be made in the first attempt. However, the accessible times of units are generally unknown; hence, interviewers are asked to make multiple calls (a maximum of five) on a sample unit. Some sampled units have “access impediments” that prevent interviewers from contacting them (e.g., locked apartment buildings). People who are rarely at home often remain uncontacted even after repeated call attempts by intervierwers. Similarly, people who have call blocking services on their telephone often are not aware of the attempts of telephone interviewers to reach them.

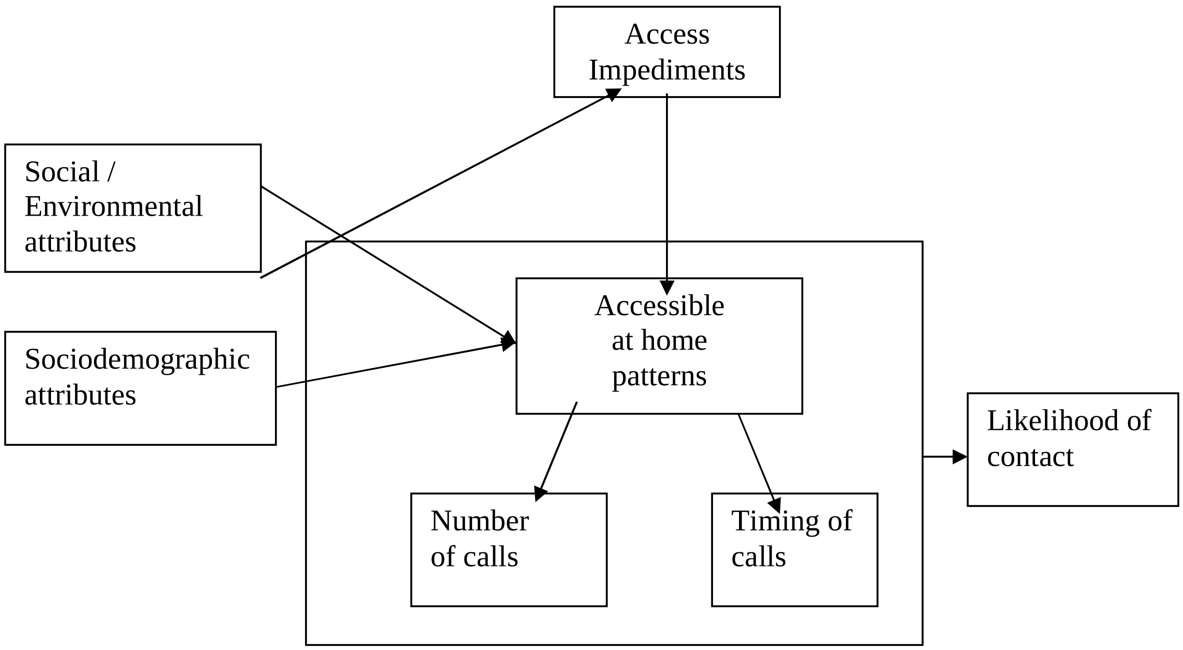


Figure 1: Causal influences on contact with sample household

In practice, the percentage of successful calls declines with each successful call. For example, figure 2 below presents the percentage of sample Agbowo community households contacted by call number among those yet never contacted in a demographic household survey conducted by the author in 2008. About 58% of the contacted households were reached in the first call. With each succeeding call, smaller and smaller percentages were reached.

It was observed that two principal factors predict the number of calls required to gain first contact in household surveys: calls in the evenings and on weekends were found to be more productive than calls at other times; different populations were found to have different accessibility likelihoods.

Generally, sample persons tend to be more accessible to interviewers when they are at home. The problem is to predict when sample persons would be at home. For those who are employed out of the home, most are away from home at set times, often the same periods each week. Most employed persons in Nigeria are away from home from 7.00 a. m. to 6.00 p.m, Mondays through Fridays. However, exceptions may be found in Lagos and Abuja as a result of poor traffic situations. If interviewers call at those times, proportionally fewer persons would be reached. The best times to meet people at home are Saturdays and Sundays and in the evenings from 6.00 p.m. to 9.00 p.m. local time. The easiest households to contact tend to be those in which someone is almost always at home. These include households with persons who are not employed outside the house, either because they care for young children not yet in school, or because they are too old to work. On the other hand, persons in households with access impediments are the most difficult to reach. These include persons in apartment buildings with locked central entrances (e.g. old and new Bodija in the city of Ibadan), and gated residences.

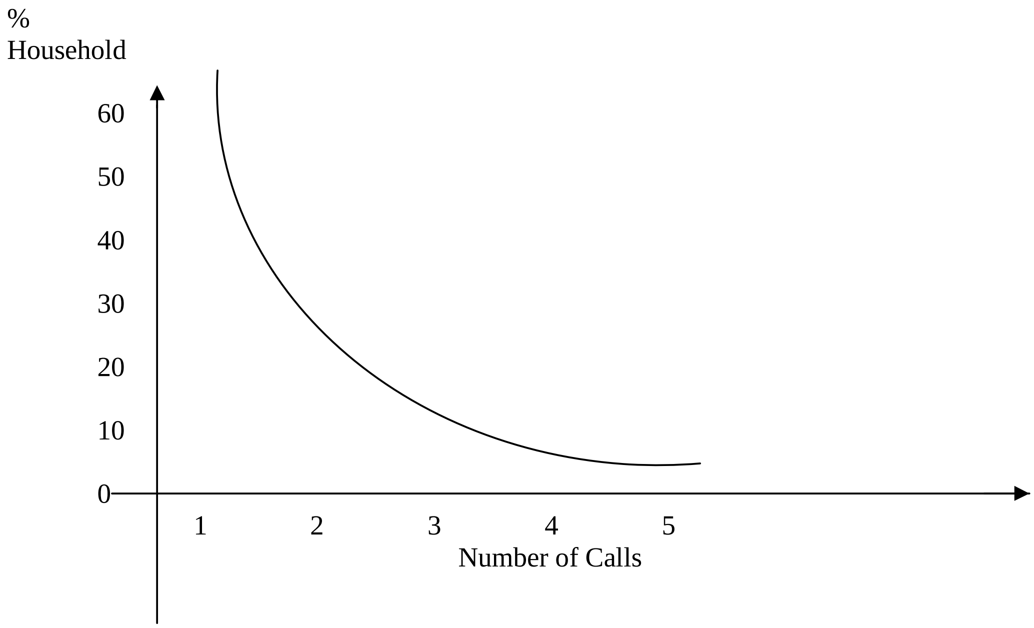


Figure 2: Percentage of eligible sample households by calls to first contact

It may be noted that non-contact non-response may be independent of the purpose of the survey. That is, the sample unit is not difficult to contact because of the topic of the survey but rather because of the set of influences that would be present for any survey request. Clearly, non-response error would arise only for statistics related to those influences.

1.4 Unit Non-Response due to Refusals.

Success in surveys requires the willingness of persons to respond to a complete stranger who calls them on the telephone, mails them a request, or visits their home. The sample persons must have little fear of financial harm from the interviewer, of reputational damage from the interaction or of psychological distress caused by the interviewer. The respondent must believe the pledge of confidentiality that the interviewer proffers; they must believe that they can speak their minds and report intimate details without recrimination or harm. Graves and Kahn (1979) argued that the essential societal ingredients for surveys to gain cooperation of sample persons are rare in human history. Research has shown that non-response involves influences that arise as a result of the following levels:

(a) The social environment [e.g., urban areas tend to generate more refusals in household surveys; households with more than one members generate fewer refusals than single person households (Groves and Couper, 1998].

(b) The person level [e.g., males tend to generate more refusals than females (Smith, 1983)].

(c) The interviewer level [e.g., more-experienced interviewers obtain higher cooperation rates than less-experienced interviewers (Groves and Couper, 1998)].

(d) The survey design level (e.g., incentives offered to sample persons tend to increase cooperation).

The first two influences are out of control of the researcher. For example, there are events that have nothing to do with a survey request that affect how people react to the request. The last two influences, the interviewer level and the survey design level are features that the researcher can manipulate to increase response rates.

The theoretical perspectives that have been applied to survey participation include:

(a) **Opportunity Cost –** this is based on the notion that busy persons disproportionately refuse to be interviewed because the cost of spending time away from other pursuits is more burdensome for them than for others.

(b) **Social Isolation** – this is based on the notion that persons at the high and low ends of the socioeconomic spectrum live isolated life, and consequently, have a tendency to refuse survey requests.
(c) **Topic Interest -** Those who are not interested in the topic of the survey have a tendency to refuse survey requests.
(d) **Over surveying –** This suggests fatigue from survey requests.

A theory known as Leverage – Salience (Groves, Singer, and Corning, 2000) attempts to describe the underpinnings of these behaviours. It claims that different individuals place different importance on features of the survey request (e.g., the sponsor of the survey, topic of the survey, how long the interview would take, what the data will be used for). While some individuals may positively value some attributes, others may negatively value them. As would be expected, these differences in individuals are generally unknown to the statistician. When the sample person is approached for survey requests, one or more of these attributes would be made salient in the interaction with the interviewer. Depending on what is made salient and how much the individual positively or negatively values the attributes would determine a response or refusal outcome. It follows that the value that a sample individual places on a specific attribute of the request, called the leverage of the request is very important in determining an outcome. Another determining factor is how important the specific attributes become in the description of the request, known as salient.

1.5 Unit Non-Response due to the Inability to Provide the Requested Data

Sometimes, sample persons are successfully contacted and would be willing to be respondents, but cannot. Their inability stems from several sources, including:
(a) They are mentally incapable of understanding the questions
(b) They are incapable of retrieving from memory the information requested
(c) Sometimes in business surveys, establishments do not have the necessary information available in the format, or time frame required by the survey

Since the reasons for their inability to comply with the survey request are diverse, statistics affected by non-response are diverse as well

2.0 Item Non-Response

Item non-response occurs when a response to a single question is missing. The impacts of item non-response on a statistic are exactly the same as that for unit non-response, but the damage is limited to statistics produced using data from the affected items.
The causes of item non-response are different from those of unit non-response. Whereas unit non-response arises from a decision based on a brief description of the survey, item non-response occurs after the measurement has been fully revealed. The causes of item non-response include:
(a) inadequate comprehension of the intent of the question, judged failure to retrieve adequate information, and
(b) lack of willingness or motivation to disclose the information, (Beatty and Herrmann, 2002; Krosnick, 2002).

Beatty and Herrmann (2002) posited a model of the response process which distinguishes four levels of cognitive states regarding the information sought by the survey question. These include:
(a) Available (information can be retrieved with minimal effort)
(b) Accessible (information can be retrieved with efforts or prompts)
(c) Generatable (information is not exactly known but can be estimated), and
(d) Inestimable (information is not known and no basis for estimating it)

The above four states are ordered by level of retrieved knowledge suitable for a question response. They posit both errors of commission (reporting an answer without sufficient knowledge) and errors of omission (failing to report an answer when the knowledge exists). Sometimes, social influence prompts sample persons to give an answer which may produce data with measurement errors. Item-missing data can arise legitimately (for those in an "inestimable" cognitive state) or as a response error (for those with the knowledge available). The latter

situation might arise when social desirability influences a respondent to refuse to answer a question (or answer, "do not know") instead of revealing a socially unacceptable attribute.
It follows that item non-response may be reduced by the reduction of the burden of any single question, the reduction of psychological threat or increase in privacy (e.g. , self-administration), and interviewer actions to clarify or probe responses.

The strategies used to compensate for item non-response are often quite different from those for unit non-response, as in the former case the analyst usually has sufficient vector of other responses with which to adjust. Hence, imputation is most often used for item-missing data, whereas weighting class adjustments are common for unit non-response.

3.0 Effect of Non-Response on the Quality of Survey Statistics.

Sometimes, non-response introduces systematic distortion in survey estimates; sometimes, it does not. The principles that determine when non-response distort survey estimates and when it does not are clear, but, in practice, researchers cannot know which situation they are facing.

Bias flows from non-response when the causes of the non-response are linked to the survey statistics measured. For example, if one mounts a survey whose key statistic is the average number of persons per household, $\hat{R}$, an item non-response like "household income" would not affect $\hat{R}$. However, empirical studies have shown that non-response may substantially distort estimates, that is, introduce bias. To give a numerical illustration of the possible effect of non-response on survey statistic, we consider a survey mounted to estimate the percentage P of deaf people in a city (Dalenius, 1985). A simple random sample of $n = 10{,}000$ people was selected and a questionnaire mailed to the 10,000 people, asking if they were deaf. Of these people, n_r returned the questionnaire with the answer (Yes or No) to the question. Among these n_r respondents, P_r^* percent responded that they were deaf. The question is: how close is P_r^* to the corresponding P^* for all 10,000? In order to answer the question, the following computations were considered.
Given the non-response, two quantities were computed, namely:

$Max\ P^*$, corresponding to the assumption that all non-respondents belong to the category of deaf people; and

$Min\ P^*$, corresponding to the assumption that none of the non-respondents belong to the deaf category

The table 1 below presents the two quantities for the case where 30% of the 10,000 were non-respondents.

P_r^*	$Max\ P^*$	$Min\ P^*$	$Max\ P^* - Min\ P^*$
0	30	0	30
10	37	7	30
50	65	35	30
90	93	63	30
100	100	70	30

It is no coincidence that $Max\ P^* - Min\ P^*$ is equal to 30, the percent non-response.

4.0 Design Features to Reduce Unit Non-Response

It is well known that the different modes of data collection tend to have different average response rates. The typical finding is that face-to-face surveys have higher response rates than telephone surveys. Telephone surveys have higher response than self-administered paper surveys, other things being equal. It is also a common finding that the use of interviewers in face-to-face surveys increases response rates, both because of higher success at delivering the survey request and because of their effectiveness in addressing any concerns about participation that sample persons may have.

Figure 3 presents several features that address interviewer actions. First, leverage-salience theory of survey participation offers several deductions about interviewer behaviour. It may be noted that different sample persons are likely to vary in how they evaluate the survey request (assigning different "leverages" to different attributes). Since these are unknown to the interviewer, the interviewer must discern them in order to gain their cooperation.

One further deduction from leverage-salience theory is that training interviewers to recite the same introductory description to each sample person will not be effective (see Morton – Williams, 1993). Groves and Coaper (1998) propose two principles of interviewer behaviour that may underlie the Morton-Williams experimental findings. The principles are maintaining interaction and tailoring. Expert interviewers appear to engage the sample persons in extended conversations (whether or not they are pertinent to the survey request). The interviewers "maintaining interaction" in such a way to attempt to gain information about the concerns of the survey person. Effective interviewers then "tailor" their remarks to the perceived concerns of the sample person. This tailoring appears to explain some of the tendency for experienced interviewers to achieve higher cooperation rates than novice interviewers. They carefully observe the verbal and non-verbal behaviour of the persons in order to discern their concerns. When they form hypotheses about those concerns, the interviewers "tailor" their behaviour to the concerns. They customize their description of the survey to those concerns.
Figure 3 also indicates that if the initial decision of the sample person does not yield an interview, further efforts to bring the person into the respondent pool involve switching interviewers, changing to a different mode or sending persuasion letters. Other methods to increase response rate include.

(a) Making the Public "Survey-Minded"

If the public has a positive appreciation of statistics, it will cooperate as respondents in surveys to a large extent than what else would be the case.

(b) Training the Statisticians

If the statisticians have a good understanding of the problem of non-response, they will address this problem, but without such an understanding, they may just disregard it.

(c) Call-Backs and Reminders

In an interview survey, a respondent may not be at home, at the time when the interviewer pays a visit to make the interview. This may happen, even if the time for that visit has been chosen so as to increase the likelihood that the respondent is at home. If contact is not established, it may be desirable and efficient to make call-backs. By the same token, in a mail survey, those who do not respond to the initial mailing may be sent a reminder (and a new copy of the questionnaire).

(d) Sub-sampling the Non-Respondents

This procedure was developed by Hansen and Hurwitz in 1946 and is widely used in surveys by mail or inter-net.

We will consider a specific case in order to estimate the percentage P of people who are deaf. A sample of $n = 10{,}000$ people is selected and a questionnaire is sent to them. 7,000 people fill in and return the questionnaire; thus the initial number of non-response is 3,000. A reminder is sent to the 3,000. Assume that 1,000 fill in and return the questionnaire; thus, there are 8,000 respondents (corresponding to a response rate of 0.80), and 2,000 non-response (corresponding to a non-response of 0.20). A second step calls for selecting simple random sample of say $n = 400$ of those non-respondents and having them interviewed. Assume all 400 cooperate.

In order to estimate , the following estimate is used;

$$\hat{P} = 0.8p_1 + 0.2p_2$$

where p_1 is the estimate applied to the data collected by mail, and p_2 is the estimate applied to the data collected by interview. The p_1 would have been the estimate if no interview were carried out.

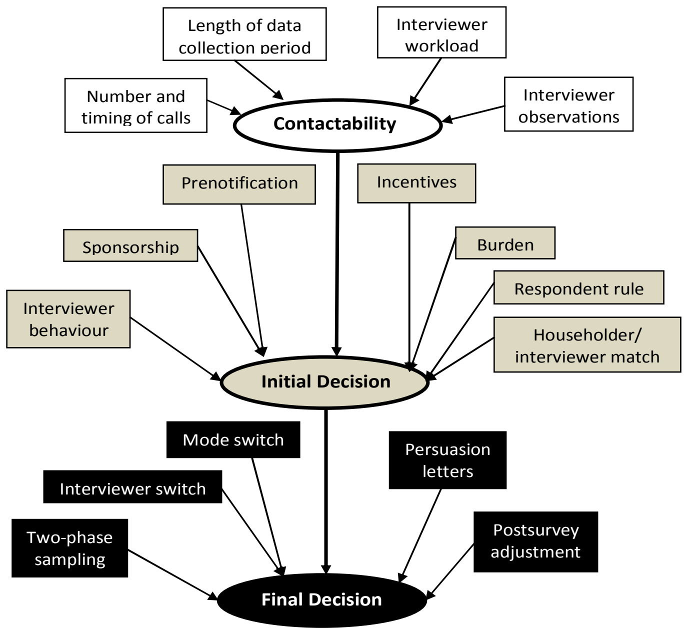


Figure 3: Tools for reducing unit non-response rates

5.0 Reducing the Effect of the Non-Response

The measures discussed in section 4 may greatly reduce the non-response, they may not eliminate it. To this end, measures to reduce the effect of non-response should be considered. These measures are in the nature of "adjustments" of the estimates based on the data available; the term "correction" is sometimes used but should be avoided, as it implies removal of the effects. We will consider two measures; weighting adjustments (used for unit non-responses) and imputations (used for item non-response).

The assumption underlying the weighting and imputation procedures is that once the auxiliary variables used have been taken into account, the missing values are missing at random. To this end, the non-respondents are assumed to be like the respondents within the weighting and imputation classes. Greenless et al (1982) has shown that this assumption can be avoided by using stochastic censoring models. However, as Little (1986) observes, the models are highly sensitive to the distributional assumptions made.

5.1 Weighting Adjustments

Surveys with complex sample designs, often also have unequal probabilities of selection, variation in response rates across important subgroups, and departures from distributions on key variables that are known from outside sources for the population. It is now common practice to generate adjustment weights to compensate for each of these features in analysis.
Weighting adjustments are primarily used to compensate for unit non-response. These procedures increase the weights of the specified respondents so that they represent the non-respondents. They require auxiliary information on either the non-respondents or the total population. There are five types of weighting adjustments; namely post stratification weighting adjustments, population weighting adjustments, sample weighting adjustments, ranking ratio adjustments, and weight based on response probabilities (details are provided by Kalton, 1983).

6.0 Post Stratification Adjustment

Post stratification uses the adjustment cells that are formed in the same way as strata sample selection. They are, however, defined by variables not available at the time the original data were selected. The cells are also mutually exclusive and exhaustive and it is expected that the values of the study variables, Y, in each cell be more similar than among all values in the sample. The best post stratification variables are those strongly correlated to the Y variable. To this end, they are often correlated with individual response probabilities.
Kovar and Poe (1985) used post stratification adjustment in the National Health Interview Survey (NHIS) conducted by the National Centre for Health Statistics. In this survey of the civilian, non-institutionalized population in the United States, each respondent was assigned to one of 60 age – race – sex cross-classification cells for which reliable current population figures $\Delta_h = N_1/N$ were available independent of the survey. A post stratification adjustment is computed for the $h(th)$ cell $(h = 1, 2, 3, \dots, 60)$ as

$$a_h^* = \Delta_h \sum_{h=1}^{H} \frac{\sum_{i=1}^{n_{1h}} W_{hi}^{(2)}}{\sum_{i=1}^{n_{1h}} W_{hi}^{(2)}} \tag{6.1}$$

where $W_{hi}^{(2)}$ is the raw sample weight $\left(W_{hi} = \pi_{hi}^{-1}\right)$ times a weighting class adjustment. The final adjustment sample weight is given by

$$W_{hi}^{(3)} = a_h^* W_{hi}^{(2)} \tag{6.2}$$

From (6.2), it follows that

$$\sum_{i=1}^{n_{1h}} W_{hi}^{(3)} / \sum_{h=1}^{H} \sum_{i=1}^{n_{1h}} W_{hi}^{(3)} = \Delta_h = N_h/N \tag{6.3}$$

This shows that post stratification attempts to make the weighted relative frequency distribution among cells to correspond to the relative distribution among those same cells in the population. By using this adjustment the NHIS sample weights were finally adjusted to bring the sample in line with the U.S. population, at least, with respect to the joint distribution by age, race, and sex as defined in the 60 cells. This means that a sample distorted by non-response, poor sample coverage, and sample variation now has weights allowing the weighted data more accurately to estimate parameters whose measurement of the response variable is correlated with the three post stratification variables.

For the special case where the initial sample is chose by simple random sampling, the same adjustment cells are used for the weighting class and stratification adjustments. Kalton (1983) presents statistical properties of the corresponding estimator of the population mean that uses the weighting class and post stratification adjustments, namely,

$$Est.\bar{Y}_{ps} = \sum_{h=1}^{H} \Delta_h \bar{y}_{1h} = \sum_{h=1}^{H} \frac{N_h}{N} \frac{y_{1h}}{n_{1h}} \tag{6.4}$$

where $\bar{y}_{ih} = \sum_{i=1}^{n_{ih}} Y_{hi}/n_{ih} = y_{ih}/n_{ih}$, and $W_{hi}^{(3)} = N_h/N$, with bias given by

$$Bias\left(Est.\bar{Y}_{ps}\right) = \sum_{h=1}^{H} \Delta_h \lambda_{0h} (\bar{Y}_{ih} - \bar{Y}_{oh}) \tag{6.5}$$

where $\lambda_1 = N_1/N$, expected response rate, and λ_{0h} = expected non-response in the hth cell. The result in (6.5) implies that the amount of non-response bias can be reduced to the extent that cells with equal respondent and non-respondent means are formed. The variance of $Est.\bar{Y}_{ps}$ is expressed as

$$Var\left(\bar{Y}_{ps}\right) \triangleq n \sum_{h=1}^{H} \frac{\Delta_h \lambda_{ih} S_{1h}^2 + n^2 \sum_{h=1}^{H} \lambda_{0h} S_{1h}^2 - \sum_{h=1}^{H} (1 - \lambda_{1h}) S_{1h}^2}{n^2 \lambda_{1h}^2} \tag{6.6}$$

where S_{1h}^2 is the element variance among all respondents in the h^{th} cell.

7.0 Imputation Methods

Despite the researcher's best practices to minimize item non-response through preventive methods, some missing items almost always appear in survey data, thus requiring the researcher to find other ways to deal with the remaining non-response. A wide variety of imputation methods has been developed for assigning values for missing item responses. These methods range from simple ad hoc procedures used to ensure complete records in data entry to sophisticated hot-deck and regression techniques. The following are some common imputation procedures:

Mean-Value Imputation, Regression Methods, Deductive Imputation, Class Mean Imputation, Hot-Deck Methods, Distance Function Matching, Exact Match Imputation and Model-based Methods.

8.0 Choosing Among Methods

The methods for dealing with non-response are basically of two types, preventive and compensatory. Preventive methods are designed to reduce non-response rate, while compensatory methods serve to reduce the effect of remaining non-response, after suitable combination of preventive measures had been applied. In deciding on a suitable preventive strategy for survey non-response, one should take into consideration, the social – environmental attributes, socio-demographic attributes, and the culture of the target population. Based on our prior experience, a combination of incentives, multiple call-backs and endorsements will likely be most effective in many situations. The kind of incentive given would depend on whether the respondent is head of a household or an establishment. Advantages of incentive are more than the disadvantages: it enables timely response to questionnaire, motivates respondents to fill questionnaire or grant an interview, and breaks the resistance of respondents, and promotes propensity to fill questionnaire or grant an interview. For example, the distribution of CBN publications to respondents in establishment surveys will aid the respondent to understand the use of the data supplied, and would likely increase their willingness to co-operate in future surveys.

Assessing the utility of non-preventive methods in deciding on a strategy for dealing with non-response may involve:

(a) finding that method which allows the researcher make statistical inference he had intended while minimizing the effect of non-response on inference,
(b) identifying those methods with the smallest mean square error in evaluating non-preventive strategies,
(c) when investigating relationship (after cross-tabulation) one would like to pick the method that least alters the relationship being studied
(d) when using model-based approach, one may be concerned primarily about finding approaches that minimize the bias and variance arising from the assumed model and whose estimators are most robust to departures from the assumed model. Fast rates of convergence for iterative methods would also be desirable
(e) the cost effectiveness issue must be considered in choosing among approaches to dealing with non-response. Also, the complexity of implementing the methods must be considered. For example, sophisticated approaches such as multiple imputation applied to the hot-deck method may not be practical when staff are unavailable to apply the method and interpret its findings.

The challenge in making the final choice is to recognize the relative strengths and weaknesses of competing alternatives for the survey. The researcher should focus more intently on finding functional and rational basis for choosing among competing methods.

9.0 Conclusion Remarks

Surveys produce data that attempt to describe large populations by measuring and estimating only a sample of those populations. When the designated sample cannot be completely measured and estimates are based only on responding cases, the quality of survey statistics can be threatened. Prevention methods are mandatory for the planning stage of every survey, because no researcher or beneficiary can afford to lose the significance of the collected data. Any survey design should have at the planning stage, the action to be taken when non-response occurs, and appropriate tool for data-collection developed so as to make it possible to obtain maximum information from the sampled units. It seems that the quality of the questionnaire, the training, and experience of the interviewer are the most important aspects that insure the success of a survey.

Not all non-response distort the quality of survey estimates. Non-response produced by causes that are related to key survey statistics is the most harmful kind. Such non-response is termed "non-ignorable non-response. Non-response can harm the quality of both the descriptive and analytic statistics.

There are many tools that survey researchers have to increase the response rates in surveys. These include repeated call backs, small interviewer workloads, advance letters, short-questionnaires, tailoring of interviewer behaviour to the concerns of the sample person, mode and interviewer switches for reluctant respondents. Almost all of these methods require spending more time or effort contacting or interacting with the sample units. This generally increases the costs of surveys.

An important remaining challenge to survey research, regarding non-response is determining when it decreases the quality of survey statistics and when it does not.

References

Beatty & Herrmann (2002): To answer or not to answer: Decision processes related to Survey item Non-response, in Groves, R., Dillman, D., Eltinge, J. and Little, R. (eds.). Survey Non-response. Wiley N.Y.

Dalenius, T. (1985): Elements of Survey Sampling

Durbin, J. (1954): Non-Response and Call Backs in Surveys. Bulletin of International Statistical Institute, 34(2), 72 – 86.

Greenless, W. S., Reece, J. S., and Zieschang, K. D. (1982): Inputation of Missing Values when the Probability of Response Depends on the Variable being Inputed. JASA, 77, 251 – 261.

Groves, R. and Couper M. (1998): Non-response in Household Interview Survey, New York Wiley.

Groves, R., and Kahn, R. (1979): Surveys by Telephone – A National Comparison with Personal Interview. New York; Academic Press.

Groves, R. Singer, E., and Corning, A. (2000): Leverage – Salience Theory of Survey Participation – Description and an Illustration. Public Opinion Quarterly, 64, 299 – 308.

Hansen, M. H. and Hurwitz, W. N. (1946): The Problem of Non-Response in Sample Surveys. JASA, 41, 516 – 529.

Kalton, G. (1983): Compensating for Missing Survey Data. Ann Arbor – Survey Research Centre, University of Michigan.

Kovar, M. G. and Poe, G. S. (1985): The National Health Interview Survey Design, 1973 – 84, and Procedures, 1975 – 83, Vital and Health Statistics, Series 1, Programmes and Collection Procedures, nos. 18, DHHS Publication (PHS), 85 – 1320. Hyattsville, MD: National Centre for Health Statistics, Public Health Service, U. S. Department of Health and Human Services.

Kish, L. (1965): Survey Sampling. New York – Wiley.

Krosnick, J. (2002): The Causes of No Opinion Responses to Attitude Measures in Surveys. Survey Non-Response, 87 – 100, New York – Wiley.

Little, R. J. A. (1986): Survey Non-Response Adjustments for Estimates of Means. International Statistical Review, 54, 139 – 157.

Morton – Williams, J. (1993): Interviewer Approaches. Aldershot – Dartmouth.

Smith, T. (1983): The Hidden 25 Percent – An Analysis of Non-Response on the 1980 General Social Survey. Public Opinion Quarterly, 47, 386 – 404.

Addressing the Problem of Non-response and Response Bias

Fabian C. Okafor[1]

INTRODUCTION

Most of the statistics needed for national planning are derived from large scale sample surveys with households as reporting units both in the developed and developing countries. Examples of these surveys are Income and Expenditure, Employment, Food consumption and Nutrition, Agriculture, Health, Education, Establishment, etc. The required information on these topics is obtained from the selected households or firms in an Establishment survey. The households to be canvassed are usually selected by probability sampling. As it happens in all surveys some selected units may not be contacted, or fail to respond to the interview or when they do provide distorted or inaccurate responses. These introduce some element of bias in the estimates.

Survey planners and analysts in Nigeria have devoted much more attention to sampling errors at the expense of non-sampling errors (non-response and response errors). Sampling error is the degree to which the sample estimate differs from the average value of the characteristic due to chance. The present discussion will be centered on non-sampling error, which may present serious deficiencies in the statistics and render the survey useless. According to Platek and Gray (1986), "Non-response has been generally recognized as important measure of the quality of data since it affects the estimates by introducing a possible bias in the estimates and an increase in sampling variance because of reduced sample." They continued by saying that "in a practical way, the size of non-response may indicate the operational problems and provide an insight into the reliability of survey data." There is need therefore to study the nature and effect of non-response in the surveys conducted in the country.

NON-SAMPLING ERROR

Non-sampling error is due to other causes apart from chance factor, i.e. inductive process of inference from a probability sample. This type of error is found both in sample survey and complete enumeration. Non-sampling error can occur anywhere from the planning to analysis stage of the survey. At the planning stage, we have such factors like ambiguous definitions and concepts in designing a questionnaire, length of a questionnaire, omission or duplication of some units due to use of obsolete frame. At executive stage we have such factors as inability to locate some of the units of enquiry probably due to civil disturbances, flood disaster, security problem, etc; memory lapse on the part of the respondent; inadequate training of enumerators, etc. The factors at the analysis stage include carelessness in editing of the completed questionnaire, tabulation and printing of final results among others. Non-sampling error can be classified into non-response and response errors.

Non-response Error

We define non-response as failure to collect data from a sample unit in the target population. It may occur because of refusal of some units in the sample to return the completed questionnaire or to grant interview in the case of face-to-face interview or through non-contact. The first kind is called unit (total) non-response. Non-response also occurs when a unit provides information to some but not all questions in the questionnaire. This is called item non-response. Item non-response may be as a result of irrelevant or sensitive questions in the questionnaire; question not understood or through fatigue or lack of knowledge.

The size of non-response is an indication of how reliable the survey data are.

[1] Dept of Statistics, University of Nigeria, Nsukka; e-mail: fc_okafor@yahoo.com

Response Error

Response error is the difference between the true value of a characteristic and the actual value supplied by the respondent or recorded by the interviewer. Response error occurs as a result of faulty instrument, the respondent or the interviewer, or as a result of interplay of the three.

More details on response error can be found in Moser and Kalton (1979, P 378).

PROBLEM OF NON-RESPONSE AND RESPONSE ERRORS

Non-response has effect on the survey estimates because of potential bias introduced in the estimates. However, “the response bias may not be nearly as serious relative to the sampling error for small samples as it is for large samples” (Platek and Gray, 1986). Let ϕ be the population proportion of the respondents and 1-ϕ that of the non-respondents. Suppose the interest is in the estimation of the population mean, μ. If only the respondents’ values are used to estimate the mean, the estimate will be □$_{1.}$ Let the average value of this estimate be μ_1; hence the bias introduced due to non-response is given by

$$\mu_1 - \mu = \mu_1 - \{\phi\mu_1 + (1-\phi)\mu_2\} = (1-\phi)(\mu_1 - \mu_2) \qquad (1)$$

This is the measure of non-response bias.

Let y_i be the survey value and x_i the true value of the variable of interest; then $y_i = x_i + e_i$, and $y_i - x_i = e_i$ is the individual response error. The average of all individual response errors given by $\overline{Y} - \overline{X} = \overline{e}$ is called the response bias.

As we have seen, non-response and response errors introduce bias in the estimates and create distortion in the survey results. Furthermore, non-response rate is an essential factor in assessing the reliability of survey results. When the non-response rate is high and vary from one area to the other, the bias may still remain even after adjusting for non-response. Therefore, it is very important to include the response rate in every survey publication. Actually, estimates from the analysis of only the respondents’ records do not apply to the target population, but are representative of the population of individuals who would respond to the survey, especially when the characteristics of the respondents differ much from those of the non-respondents.

DETECTION AND CONTROL OF REPONSE ERROR

There is yet no satisfactory method of ascertaining the size of response errors in the individual sample survey or census. But what is whispered among those who have dealt with surveys in Nigeria is that the greatest cause of response error is deliberate falsification and fabrication of survey data by the respondents and interviewers respectively. It is not an easy task to ascertain whether a respondent has given a true answer or not during interview. The usual practice is to use some checks which mainly determine the discrepancy in responses. These checks include the following.

- Record Checks: We can detect errors in response through the use of records or documents, e.g. birth records, marriage directories. In Nigeria we have not fully imbibed the culture of record keeping and official records are not readily available because of official secrecy and unnecessary bottleneck. Even if the individual records are available, it will not be easy to match sample respondents with official records. In some cases some names may not be available in the records. Moreover, records may not indicate the true value unless one is sure that the information in the record is more accurate than the answers supplied by the respondent during the survey.
- Consistency Checks: Consistency checks are normally inbuilt in the questionnaire. In consistency check, information sought from the individual in a particular issue is obtained by asking the relevant question in two or more different ways located at different points in the questionnaire. This will serve as a check on

the quality of data collected. This notwithstanding, respondents determined to give false answers may still do so.

- Re-interviewing: This is also known as quality check or post enumeration survey in census. It involves re-interviewing a sub-sample of individuals in order to ascertain the extent the first response is consistent with that of the second interview. In re-interview one could repeat the same survey questions or use more detailed questions and better trained and experienced interviewers.

For the control of response error, the procedure to be adopted must be carefully considered before the survey. One of the methods available to the sample survey expert for reducing response error is the careful selection, training and supervision of interviewers. Another is the use of well designed questionnaire with precise and clear instructions.

In Nigeria, most survey bodies pay more attention to sample size and sampling error neglecting non-sampling error, which in most cases may be serious as to distort the estimates from the survey as was earlier pointed out.

ASSESSING THE EXTENT OF NON-RESPONSE BIAS

The question is how do we assess or determine the extent of non-response bias? I have not come across any research carried out in Nigeria to assess the extent and nature of response and non-response bias. We rely on results from other countries, which should not the case. We have to bear in mind that decreasing non-response rate may not necessarily always reduce the non-response bias; but low response rate may yield statistics with large non-response bias. The methods for assessing the extent on non-response bias are found in http://www2.chass.ncsu.edu/garson/pa765/sampling.htm and the UN Handbook of Household Surveys, Part One and summarized below. These are:

- **Population comparison:** Here survey averages are compared with those from other sources like recent population census. This method is useful in comparing variables like gender, age distribution, race, income and occupation. The idea is to identify variables where the sample mean differs from the population mean. The setback in this method is that the deviation may be due also to the sample units selected and canvassed in addition to non-response bias.
- **Intensive post-sampling:** A sample of non-respondents is selected and intensive effort is made to obtain interview on this sub-sample. The difference between the estimates from the respondents and the sub-sample of non-respondents is a measure of the non-response bias, i.e.

$$\frac{n_r}{n}(\bar{y}_r - \bar{y}_{nr}) \qquad (2)$$

- **Matching:** A sub-sample of the respondent and non-respondent sample cases is matched with the current population census. An indication of the non-response bias is given by the difference between census data for respondent and non-respondent sample cases.
- **Call backs:** A comparison of the survey results for sample units interviewed after considerable call backs (e.g. after the third visit) with those interviewed at the first attempt provides some measure of non-response bias.
- **Adjustment:** Another method of assessing non-response bias is to compare the unadjusted estimates based on the respondents with those estimates obtained after adjusting for non-respondents (Ekholm and Laaksonen, 1991).

HANDLING NON-RESPONSE

Non Theoretical Approach

Some of these approaches discussed here came out at the interactive session during the workshop on non-response in surveys organized by Central Bank of Nigeria in September 2010. One of such approaches is developing a rapport with the community or the respondents through social engagements, like attending naming ceremonies, traditional marriages, funeral condolences and sending season's greetings, etc. For example, it was reported of a case of an interviewer who obtained cooperation by merely giving attention to a child in the family; asking about his welfare. Response rate could be increased through advocacy using radio jingles, print media and television to sensitize people about an oncoming survey. Also community heads or chiefs when contacted and educated on the nature and benefit of a survey can help to get cooperation and thereby improve on the response rate.

The manner of dressing determines whether one can obtain interview or not. One who is going to interview a company director must be if possible dress like one and come into the premises with a good car. In short and interviewer should dress well if possible better than the respondents. There was a case where a female interviewer went to a community in a pair of trousers; she never succeeded in obtaining interview from any one. But when she dressed like a typical traditional woman with a wrapper and a matching blouse the situation changed. The view of the community was that she came to corrupt their girls with her type of dressing.

Another way of increasing cooperation is through the use of incentives either material or financial. People are of the opinion hat incentives could be given but only if the respondent fails to respond after two or three visits. Incentives may be in the form of company calendar, diary or reports of past surveys, especially in an establishment survey. Corporate award may be given to an establishment that cooperates regularly and effectively in surveys. Financial incentive may help in some cases but may be counter productive. There is a case of a gateman who refused an interviewer entry to an establishment; but immediately money exchanged hands the gate flew wide open. However, when people get used to financial gift, they fail to respond unless some amount of money is given. This happened in a community where the interviewer who used to give monetary incentive was changed with someone who does not give. At the next round of survey the interviewer met brick wall; response rate dropped drastically.

Finally, use of locals with knowledge of the terrain and good command of the local language and with necessary academic qualification will definitely improve cooperation on the part of the respondents. For lack of space and time we shall not mention all other methods of increasing the response rate.

Theoretical Approaches

There are several methods of handling non-response depending on whether it is total or item non-response. Total non-response is handled by using weighting adjustments whereas item non-response is taken care of by imputation.

Weighting Adjustment

"The aim of any weighting adjustment is to increase the weights of the respondents so as to compensate for the non-respondents" (Okafor, 2002, p358). Proponents of weighting adjustment argue that it is useful for adjusting for biases in the sample. Those against are of the opinion that it makes little or no difference to conclusions. According to Bourque and Clark (1992, p60) "the use of weights does not substantially change estimates of the sample mean unless non-respondents are appreciably different from respondents and there is a substantial proportion of non-respondents." We shall now summarize the methods used for weighting adjustments. For more details the reader is referred to Lohr (1999, p266) and Kalton and Kasprzyk (1986).

1. **Weighting adjustments overall:** Let π_i be the probability of selecting unit i in the sample. The sample weight (raising factor) is $w_i = \dfrac{1}{\pi_i}$, $i = 1,2,\ldots, n$. Now the estimate of response probability is $\varphi = \sum_{i=1}^{n} \alpha_i w_i \Big/ \sum_{i=1}^{n} w_i$; α_i takes the value one if the i^{th} unit respond and zero otherwise. The new weight for the respondents will be $w_i^* = \dfrac{1}{\pi_i \varphi}$ and the estimator of the population total of the variable of interest y will be $\hat{t} = \sum_{i=1}^{n} \alpha_i w_i^* y_i$. (3)
2. **Weighting- class adjustment:** The sample units are grouped into classes using auxiliary variable available for all sample units like sex, race, and geographical area. Each class response probability is estimated and used to adjust for class non-response.
3. **Population weighting adjustment:** The sample respondents are classified into strata using known auxiliary variable. Each stratum has a known population distribution (proportion), W_h from past census result, which is

used as weight for each respondent. The advantage of the population weighting adjustment is that it reduces the non-response bias as well as non-coverage error.

4. **Sample weighting adjustment:** The total sample units are grouped into strata using information available for both respondents and non-respondents in the sample. The sample distribution (proportion) $w_h = n_h / n$ is used as weight in each stratum. For simple random sampling the estimator of the mean is
$$\bar{y}^* = \sum_{h=1}^{H} w_h \frac{1}{n_{rh}} \sum_{i=1}^{n} \alpha_i y_i \qquad (4)$$
5. **Post-stratification using weights:** As described by Lohr (1999, p268), this method uses the ratio of N_h (population units in stratum h) and its estimator $\sum_{j=1}^{n_h} \alpha_j w_{jh}$. With this ratio, the modified weight is given by
$$w_i^* = \frac{w_i N_h}{\sum_{j=1}^{n_h} \alpha_j w_{jh}}.$$

The assumptions in weighting adjustments are that the respondents and non-respondents are similar and each unit is equally likely to participate in the survey. These assumptions are never always true in practice. Hence, weighting may not completely eliminate all non-response bias.

Imputation

Imputation entails assigning values to missing responses making use of available auxiliary information on the sample units. It is used in handling item non-response in order to reduce non-response bias among other reasons. Imputation method has advantages and disadvantages; for details see Kalton and Kaspryzk (1986). Some of the imputation methods available in the literature include:

- Mean Imputation: This involves imputing to the missing response the mean of the respondents for the particular item.
- Random Imputation: A respondent is randomly chosen from among all the respondents; the value of the response for this individual is assigned to the missing response for a given item. In order to preserve the multivariate nature of the data, values from the same donor are used for all missing values of a non-respondent. Sometimes more than one value is imputed for every missing item by carrying out k greater or equal to two imputations. This is called multiple imputation proposed in Rubin (1978) and illustrated in Rubin (1986) and Iwebo (2008). The multiple imputations help one to estimate the additional variance due to imputation.
- Hot-deck Imputation: In hot-deck imputation the value of one of the respondents is assigned to the missing response. The problem with this is that the value of one donor can be used for many missing values in a particular item. Versions of hot-deck imputation are sequential and nearest-neighbour hot-deck imputations. In nearest-neighbour hot-deck imputation a distance function (for the auxiliary variables) is used to impute a value of a respondent to the non-respondent. Chen and Shao (2000) explained this for a bivariate sample values. Let all x-values be observed for all sample units. Suppose $y_1,\ldots,y_r$ are observed values for the respondents and the remaining n-r values are missing. Then the nearest-neighbour hot-deck imputes a missing y_j by y_i, where $1 \le i \le r$ and i is the nearest neighbour of j measured by satisfying the condition $\left|x_i - x_j\right| = \min_{1 \le i \le r}\left|x_i - x_j\right|$.
- Regression Imputation: Logistic regression forms part of this technique. Using the data from the respondents, the values of the variable to be imputed are regressed on some auxiliary variables available for all the sample units. The regression equation so obtained is used to predict for the missing item responses. Residuals could be added to the predicted values, if desired, to introduce randomness. Logistic

regression is used in imputing for dichotomous variable, that takes the value 1 or 0, based on the logistic regression function given by $\pi(x) = \dfrac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}$.

Its logit is $\ln\left[\dfrac{\pi(x)}{1-\pi(x)}\right] = \alpha + \beta x$. The estimates of α and β, and known x value are used to obtain the estimated probability $\hat{\pi}(x)$. A random number is drawn from a uniform distribution [0,1], say τ. If $\hat{\pi}(x_i) > \tau$, then y_i takes the value 1, and 0 otherwise.

NON-RESPONSE BIAS AND NON-RESPONSE RATE

"The relationship between the bias and the size of non-response while perhaps more important is less obvious since it depends on both the magnitude of non-response and the differences in the characteristics between respondents and non-respondents" (Platek and Gray, 1986). In fact there is no clear-cut simple relationship between non-response rate and non-response bias. However, in survey involving face-to-face interview a response rate of between 70 – 85% is regarded as acceptable. But below 70% there is a chance of non-response bias in the estimates. According to Groves (2006), the non-response bias is a function of how correlated the survey variable is to the response propensity (likelihood, probability of being a respondent). It has also been argued that decreasing non-response rates may not necessarily lead to decrease in non-response bias; and that the non-response rate is not a very good measure of the size of the non-response bias. In a review of literature on the study of non-response rate and non-response bias, Groves (2006) concluded that, for a particular survey "if the non-response rate were reduced by methods more attractive to the higher-income persons, then the relative non-response bias might decrease dramatically. If the non-response rate were reduced by methods equally attractive to higher-and lower-income persons, then the bias might be reduced as a simple function of how the non-response rate declined. However, if the non-response rate were reduced by methods more attractive to the lower-income persons, then the non-response bias might actually increase".

There is yet no conclusive study on the linkage between non-response rate and non-response bias. It depends on the variables of interest and statistics obtained and the response propensity of an individual. Non-response rate, however, is indirectly related to non-response bias in some estimates. "In short, non-response bias is a phenomenon much more complex than mere non-response rates" Groves (2006). Finally, in conclusion, "a rational approach to the problem of controlling non-sampling errors will, therefore be to try to reduce them as much as possible to levels at which the results will be usable for the purpose in view, but not to such extent as will render the efforts and costs to become incommensurate with the improvements achieved" (Murthy 1967, p467).

REFERENCES

Bourgue, L.B. and Clark, V.A. (1992): Processing Data: The Survey Example. *Thousand Oaks CA Sage Publications.*

Chen, J. and Shao, J. (2000): Nearest-Neighbour Imputation for Survey Data. *Journal of Official Statistics 16, 583 - 599.*

Ekholm, A. and Laaksonen, S. (1991): Weighting via Response Modeling in the Finnish Household Budget Survey. *Journal of Official Statistics 7: 1325-37*

Groves, R.M. (2006): Non-response Rates and Non-response Bias in Household Surveys. *Public Opinion Quarterly. 70 (5) 646 – 675.*

Iwebo, E. J. (2008): Imputing for Missing Values in the Estimation of the Population Mean of Post University Matriculation Score. *Unpublished MSc project, Dept of Statistics, University of Nigeria, Nsukka.*

Kalton, G and Kasprzyk, D. (1982): Imputing for Missing Survey Responses. *Proceedings of the Section on Survey Research Methods, American Statistical Association, 22 – 31*

---------------------------------- (1986): The Treatment of Missing Survey Data. *Survey Methodology, 12, 1, 1 – 16.*

Lohr, S.L. (1999): Sampling: Design and Analysis. *Duxbery Press, Pacific Grove, CA, USA.*

Moser, C.A. and Kalton, G. (1979): Survey Methods in Social Investigation. 2nd Edn. Heinemann Educational Books, London

Murthy M.N. (1967): Sampling Theory and Methods. *Statistical Publishing Society, Calcutta, India.*

Okafor, F.C. (2002): Sample Survey Theory with Applications. *Afor-Orbis Publications Ltd, Nsukka, Nigeria*

Platek, R. and Gray, G.B. (1986): On the Definitions of Response Rate. *Survey Methodology, 12, 1, 17 – 27.*

Rubin, D.B. (1978): Multiple Imputations in Sample Surveys – a Phenomenological Bayesian Approach to Non-response. *Proceedings of the Survey Research Methods Section of the American Statistical Asso. 20 -34.*

…………….(1986): Basic Ideas of Multiple Imputation for Non-response. *Survey Methodology, 12, 1, 37 – 47.*

Sampling: *http://www2.chass.ncsu.edu/garson/pa765/sampling.htm*

Statistics for National Development[1]

Sani I. Doguwa[2]

Good statistics that has been collected according to agreed good practices are crucial as a tool for development. Gross domestic product (GDP) and other measures of economic activity such as Gross National Income (GNI) together with their individual components, show how the economy is responding to government policy and other influences. The balance of payments can demonstrate the requirement for policy adjustments and is also one of the indicators scrutinised by potential foreign investors in the country. Agricultural statistics clearly have implications for longer-term planning, particularly if they show a move away from the land into urban areas or a change in traditional farming methods. Population statistics often indicate the need for government intervention – for example, if the population is expanding so rapidly that there are major policy implications in the fields of education, housing, etc. Health-care statistics might indicate the need for a change of policy in the provision of medical services or the preponderance of particular diseases. Growth in Monetary aggregates might indicate the stance or anticipated direction of monetary policy.

1. Introduction

Official statistics, that is, statistical data collected and disseminated by governments about different aspects of life in a country, are needed for a number of important purposes. They provide the information, the evidence, needed for the business of government – both day to day administrations as well as for policy analysis. At the same time, statistics are also important for users outside government. They provide information needed for business decision making and also help to keep individual citizens informed about what their government is doing. In a world where national economies are becoming increasingly inter-dependent, official statistics collected, compiled and disseminated by the National Bureau of Statistics (NBS) in the case of Nigeria, provide a basis for understanding how Nigeria interacts with others and how conditions compare with those elsewhere.

Until 1959, the banking industry in Nigeria was largely unregulated. There were, therefore, no reliable and organized data on the monetary and financial sub-sector. However, the phenomenal growth in the number of financial institutions and financial instruments in Nigeria subsequently led to the greater use of monetary policy for economic stabilization. Consequently, the need to monitor the events in the monetary and financial sector calls for more timely, accurate and reliable data. This involves the collection, compilation and dissemination of balance of payments, monetary and banking statistics on the Nigerian economy by the Central Bank of Nigeria (CBN).

Good statistics that has been collected according to agreed good practices - using appropriate methods for data collection, processing and dissemination, are crucial as a tool for development. They provide an objective and replicable picture of the state of a country; enable comparisons, both over time and space; and set benchmarks for measuring progress in the future. The use of well-established data processes provides data and statistics that can be analyzed using the powerful tools of statistics and econometrics.

Every country needs good statistics, therefore, almost all countries have established specialized agencies (such as NBS and CBN in the case of Nigeria) whose job is to collect, process and disseminate good statistics. The problem is that in many countries, especially those in some developing world, the work of these agencies is under-appreciated and under-valued. Many statistical systems are caught in a vicious cycle of under-funding which ultimately leads to under-performance.

Official statistics are public goods. The use of them by one person or agency does not detract from their use by another. While they are often costly to produce, they are readily disseminated and once they are publicly available, it is difficult to exclude other users. The value of statistics depends upon their quality. However, since it is not easy to ascertain the quality of statistics directly, users must have confidence in the producer and in the processes (methods and standards) employed in the production of the statistics. For all these reasons, it is difficult to establish functioning markets for statistics. This leaves national governments, institutions and international agencies to produce and disseminate statistical information.

---

[1] A paper presented at the Public Seminar: Statistics for National Development, Organized by the Nigerian Economic Society, Held at Reiz Continental Hotel Abuja on Tuesday 16th June, 2009.

[2] Statistics Department, Central Bank of Nigeria

Governments have a duty, therefore, to compile and publish official statistics. The kinds of statistics that are collected and disseminated will depend on needs and have to be determined in consultation with users and providers. With many other competing uses of government revenue in developing countries, not all statistics that are needed can be provided. As a result, vary difficult decisions have to be made on priorities. It is the job of statisticians to make the case for adequate resources to provide the most important sets of statistics. At the same time, though they have to demonstrate that they are using the scarce resources provided as efficiently and effectively as possible and the data that are being collected are of use and value. They also have a duty to ensure that the data they provide are of good quality and that users are able to have confidence in the accuracy, reliability and integrity in the data.

The rest of the paper is structured as follows. Section two discusses the development and maintenance of a good statistical system. The uses of statistics for national development are discussed in section three, while section four summarizes and concludes the paper.

2. Developing and Maintaining a Good Statistical System

2.1 Historical Background

The 19th century British Prime Minister Benjamin Disraeli observed that there are three kinds of lies: - *lies, damned lies* and *statistics*. His observation, while not calculated to displease statisticians, was an early indication of the political importance of statistics. Collection of facts and figures by government about its citizens is as old as the most ancient of civilizations. The Romans did it and there is evidence of local censuses in ancient Babylon and Egypt. A complete count of citizens was planned in China in AD 1370. The Domesday Book was an exhaustive and defining survey of England conducted by its Norman conquerors in the 11th century. However, such early population counts were very limited, the two main purposes being information for conscription and taxation.
The word 'statistics' derives from the Latin 'status' or political state and the German 'statistik' – facts and figures for use of the state. Despite their earlier beginnings, statistics didn't become properly organized until the scientific revolution in Europe in the 17th century and the growth of European nationalism. Every nation then determined to gather as many facts as possible about its population, trade, finances, taxes and armed forces. But it wasn't until the 20th century that official collection of statistics developed into the exhaustive process we see today, with data available on nearly every aspect of people's lives and on country's activities. Official statistical systems in developed countries collect and disseminate a very wide range of data about individual nations. Increasingly, as is happening in the European Union[3], for example, consistent and comparable statistics are being collected at the regional level.

2.2 Production and Dissemination of Good Statistics

The Statistics Act, 2007 (Act No. 9 of 2007) established for Nigeria a "National Statistical System" (NSS) with objectives to: raise public awareness about the importance of statistics; collect, process, analyze and disseminate quality data; promote the use of statistics; and build capacity for the production and use of data. Achievement of these objectives could institute a culture of conscious use of statistical data by both the public and private sectors for planning and decision processes.

The Act has in addition established a national statistics office, the "National Bureau of Statistics" (NBS) which plays the role of coordinator of the NSS with powers to collect, request and be provided with data throughout the country on a wide range of matters. The Statistician-General is responsible to the Presidency and is therefore politically in a position to influence decisions of government. In addition, the Act has provisions that ensure professional development. For instance, there is professional requirements for appointment to the Board of NBS, MDAs are to established statistics units with staff to be appointed by the NBS, the NBS employee service scheme is aligned with educational/research institutions in the country, and there is provision for a "National Consultative Committee on Statistics" (NCCS) that examines statistical programmes and develop statistics strategies for the country.

[3] See Eurostat website - http://www.europa.eu.int/comm./eurostat/

In addition to other powers granted the CBN under the Act No. 7 of 2007 (The Central Bank of Nigeria Act, 2007), it has powers to collect and share information relating to or touching or concerning matters affecting the economy of Nigeria. The CBN has therefore been provided with the essential legal instrument to play its role in the NSS. To this end, the Bank established its Statistics Department with the mandate to "collect, analyze and manage data on all sectors of the economy, in order to provide statistical support to the Bank, the government, international organizations and other stakeholders".

Fellegi (1975) noted that:

> *"The objective of a national statistical system is to provide relevant, comprehensive, accurate and objective statistical information. Generally, statistics are invaluable for monitoring the country's economic and social conditions, the planning and evaluation of government and private sector programmes and investment, policy debates and advocacy, and the creation and maintenance of an informed public."*

It is, therefore, useful here to emphasise some key points on how to produce and disseminate good statistics:

a) International standards should be adhered to in collecting and processing data, in order to facilitate comparison and analysis between countries and over time.
b) Those in charge of official statistics must enjoy some degree of autonomy from government in producing and disseminating data.
c) Data dissemination must be undertaken in a way that is credible both inside the country and internationally.

The Statistics Department of the CBN has recognized these key points as encapsulated in its mission statement – to provide comprehensive, timely, accurate and reliable data for policy design, analysis and evaluation through the use of *standard statistical procedures*.

At its special session of 11-15 April, 1994, the United Nations[4] have developed some fundamental principles of official statistics that were adopted unanimously by the United Nations Statistical Commission. The International Monetary Fund (IMF) has also been concerned with the quality and reliability of official statistics. The IMF's *Dissemination Standards Bulletin Board*[5] provides access to three sites:

a) The Special Data Dissemination Standard Site (SDDS) guides countries that have or might seek access to capital markets.
b) The General Data Dissemination System Site (GDDS) guides countries in public provision of timely, accessible and reliable economic, financial and demographic data.
c) The Data Quality Reference Site (DQRS) has been created to foster common understanding of data quality.

2.3 Censuses, Surveys and Administrative Data

In ancient times, collection of official statistics was little more than a simple head count. Today it is a process of considerable sophistication. The quality of the published statistics depends entirely on the ways in which the source data were collected in the first place. In developing countries especially, a key data source is the national population census, conducted in most countries once every 10 years. Generally the population census is the only complete count of the whole population ever carried out. Census data, therefore, provide a valuable source of statistical information about the current state of a country and its people. By comparing one set of results with another, significant long-term changes in the population and in key aspects of national life can be monitored. In the United Kingdom[6], for example, apart from the war year 1941, a census has been taken in Britain every 10 years since 1801, enabling comparisons over two centuries.

Full counts of the whole population, however, are both expensive and very disruptive and consequently are only held at fairly infrequent intervals. Instead, almost all statistical systems make use of statistical theory to select samples so as to be representative of the population they are selected from. Sampling theory provides the basis for selecting samples, for generating estimates of important statistics for the whole population and even for providing measures of the reliability of these estimates.

Sample surveys, therefore, of many different populations such as households, businesses, farms, or areas of the country are a key data source in most countries. In many countries, regular and continuous surveys are a key source

---

[4] See United Nations website – http://www.unstats.un.org/unsd/goodprac/bpabout.asp
[5] See IMF website - http://www.dsbb.imf.org
[6] See the United Kingdom website - http://www.statistics.gov.uk/census2001/census_news.asp.

of statistics essential for monitoring social and economic trends. They may monitor such things as the size and structure of the labour force and the activities of households including: spending patterns, family structure, housing conditions, education, health and so on. In developing countries household surveys are a key source of data to monitor welfare indicators such as those required for poverty reduction strategies and for monitoring progress towards the Millennium Development Goals.

Data derived from or collected as an adjunct to administrative processes are also a major source of official statistics. Depending on the sophistication of government, there are potentially many such sources of statistical data including population registers, tax records the documentation needed for international trade and school enrolment. In some developed societies, for example the Scandinavian countries, these sources provide the majority of official statistics. In developing countries, however, rather less use is made of administrative data, in part because of limitations of coverage, but also because of bureaucratic problems.

2.4 Producer User Relationships

The successful collection of official statistics is often described as a *circular process*, in which government (producer), those who supply data to government (respondents), those who use or benefit from the statistics (users - enterprises and citizens), and the media are all interlinked and interdependent.
The government (the official statisticians) depend on the goodwill of respondents for an adequate and efficient supply of data, even if the respondents are sometimes obliged to cooperate by statute. One way in which such goodwill can be fostered is by making the statistics compiled as a result of such cooperation available in a user-friendly way to respondents, as well as the public in general.

Official statistics are collected on behalf of the public and at its expense, using income from taxation. They provide an indication of the state of the nation, both good and bad. Access to statistical information is a source of political power. To some extent, therefore politicians may wish to “control” the flow of information or perhaps manipulate the data to show their efforts in a more favourable light. But good governance requires that official statistics are made readily available to all citizens, even if some statistics may indicate poor performance.

Ease of access to official statistics is a hallmark of an open and democratic society. Politicians need to be persuaded that providing reliable statistical data that is trusted by the user has long-run benefits in helping to promote economic growth and reduce poverty. Conversely, countries where official statistics are tightly controlled and are subject to political manipulation may well find it more difficult to attract investment and develop markets.

To complete the circle, it follows that the news media play a critical role in the process. They are a key channel through which government informs the public on statistical matters; and they act as a *watchdog* in case of any attempt by government to manipulate official statistics. The best national statistical offices put a high priority on their relationship with the media.

2.5 Role of Technology

Official statisticians manage data collection exercises, process the data to identify and remove errors and to calculate different indicators and statistics, and then disseminate the results. Typically, even in small countries they deal with very large data sets, containing thousands and even millions of records. In the past, data processing and dissemination used to be painstaking operations requiring large teams of people. Often, results didn’t appear for three, four or even five years after the actual data collection. The results that were eventually published were quite limited, special tables to obtain specific information, for example, how many people of a certain age group lived in a particular province, might take months of work. By that time, the data are out-of-date and no longer useful for decision-making and other purposes.

Information and computing technology (ICT) has changed all this, even in poor developing countries. The cost of computing power has now come down so low that it is within reach of even the smallest statistical agency. Some of the changes and opportunities that ICT has made possible include:

a) The possibility of manipulating very large data sets involving millions of figures has offered statisticians the opportunity to study micro data corresponding to very small areas, such as a city block or a rural settlement with a handful of inhabitants.

b) Such micro data can be easily assembled to study whole new geographical entities – areas of poverty, zones of industrial development, suburban sprawls in great metropolitan areas etc.
c) Thanks to new data-warehousing techniques, this process can be completed in seconds: the data are always available, even on the Internet, for compilation in tables that meet exactly the needs of the user.
d) Geographical Information Systems (GIS) have been described by UN as a computer-based tool for the input, storage, management, retrieval, update analysis and output information. The information in a GIS relates to data characteristics that are geographically referenced. GIS allows statisticians to answer questions about *where* things are or about *what* is at a given locations (Handbook on Geographic Information System and Digital Mapping by the United Nations, N.Y., 2000).
e) The Internet provides whole new opportunities in disseminating information and substantially reduces the cost of publishing data. Data can be made accessible to users without having to incur the costs of publishing large and expensive printed reports. Database management programs allow users to access data in new ways, making links between data sets that were not possible previously.

2.6 Cost and benefits of updating the statistical system

A country under development undergoes rapid change. Statistics quickly become dated, as processes such as globalisation and the use of ICT rapidly modify the picture they once portrayed. For such countries, developing a statistical system can't be simply a 'one-off' exercise. It might be possible to obtain help in developing the system and in building capacity, but then a country probably must rely on its own resources to maintain it. Even when resources are scarce, investment in good statistics must be a priority.

Because needs for statistics change, as countries develop and as users become more experienced in analysis and use of data, statistical systems need to ensure that they are continually monitoring needs and adapting their operations to reflect changing demand. But some statistical processes such as large-scale surveys and censuses are complex exercises that require considerable time to plan and execute properly. It may take two years or more to plan and carry out a new large-scale household survey, for example; a population census can take even longer. An effective and efficient statistical system, therefore, is not only one that meets current needs, but which actively anticipates future needs.

The managers of official statistical systems, therefore, need to be forward looking and to be continually arguing for an appropriate allocation of financial and other resources. At the same time they need to ensure that the infrastructure that supports statistical operations is also adequate and is being upgrade as resources permit. Such infrastructure includes physical facilities such as buildings, vehicles, computers and communications equipment as well as the building blocks of effective statistical operations, such as a business register, maps and sampling frames.

3. Using statistics for National Development

3.1 Data for Managing Government

Governments need statistics to run a country efficiently, both for day-to-day administration and for policy making in the longer term. They need statistics to manage the economy, to 'balance the books' – maintaining a balance between revenue and expenditure and ensuring macro-economic stability. Statistical data also have a crucial role to play in resource allocation, in deciding where scarce resources can best be targeted so as to achieve agreed goals and targets.

While governments finance official statistical agencies to generate the data needed for the business of government, it is now widely accepted that these agencies also have a duty to provide information to others outside government to support decision making generally. Because official statistics are public goods, it is usually not possible for the private sector to provide the wide range of data needed by businesses and individuals. Statistics can be a powerful instrument in building a national consensus on a whole range of actions – and, as such, are a cornerstone of the democratic process.

Governments also have international responsibilities through their membership and participation in both regional and international arrangements. The provision of accurate and up to date statistical information is often an important part of this. Good statistics enable a country to fulfil its international responsibilities and send a positive message to the rest of the world.

3.2 Data for Managing the Economy

One of the most important tasks all governments have is to manage the domestic economy and its interactions with the rest of the world. The actions governments take vary from country to country, but include maintaining an appropriate balance between supply and demand in the domestic economy and externally and creating the right environment for investment, economic growth and poverty reduction. The economic statistics that official statisticians collect are crucial for this process.

Growth and prosperity depend to a large extent on controlling the rate of inflation, i.e. the rate at which the retail prices of goods and services are changing (increasing or decreasing). The retail (or consumer) price index is among the most high profile and keenly anticipated statistics issued by governments, usually on a monthly or quarterly basis. One reason is that it affects nearly everyone and most people have some understanding of it – because it indicates the value of their money. Inflation is closely linked to government policy – for example, to interest rates set by the central bank – and its accurate measurement is considered a key criterion of short-term economic management.

Similarly, labour market statistics – levels of employment and unemployment and earnings – are a key short-term indicator of economic health and whether or not the government needs to make adjustments in economic policy. In developed countries these are collected and published monthly. In the developing world the data are no less important, but are more complicated to collect because many people work in what is known as the informal sector where statistics are not so readily available.

Balance of payment statistics, while not so readily understood by the man or woman in the street, are closely monitored by the government and the business world because of the economic penalties that can swiftly accrue from trade and other imbalances. Again, these are usually compiled monthly.
Statistics on industrial output are also a pointer to possible imbalances in the economy - whether or not a particular sector (for example, manufacturing industry or agriculture) is in decline and might need to be the target for government intervention.

3.3 Data for Longer-term Policy-Making

Gross domestic product (GDP) and other measures of economic activity such as Gross National Income (GNI) are key indicators for governments. Together with their individual components, they show how the economy is responding to government policy and other influences. GDP per person is a readily-understood indicator of the nation's economic well-being and one often used when comparing one country with another. The balance of payments can demonstrate the requirement for policy adjustments and is also one of the indicators scrutinised by potential foreign investors in the country.

Agricultural statistics clearly have implications for longer-term planning – particularly if they show a move away from the land into urban areas or a change in traditional farming methods. Population statistics often indicate the need for government intervention – for example, if the population is expanding so rapidly that there are major policy implications in the fields of education, housing, healthcare etc. Health-care statistics might indicate the need for a change of policy in the provision of medical services or the preponderance of particular diseases.

Money and Banking statistics are very important for the purposes of formulating monetary policy and monitoring its implementation. The major users of money and banking statistics in Nigeria are the policy makers of the CBN, Federal Ministry of Finance, National Planning Commission, the Presidency, the financial sub-sector (banks and other non-bank financial institutions), research institutes, private researchers and universities.

A major function of the CBN is the overall supervision of Nigeria's financial system. Based on the statutory returns the banks and other non-bank financial institutions render to the Bank on periodic basis, financial statistics are compiled and used to monitor compliance with policy targets. Also in the early 1990s when the banks experienced varying degrees of distress, Doguwa (1999) noted that the CBN refocused its attention on efforts to identify problem banks and to predict failures with sufficient lead time for remedial actions to be instituted to save the banks from deteriorating further.

There are many other examples: tourism statistics might be a key indicator, especially in a developing economy; transport statistics could be vital for infrastructure planning; crime statistics might suggest problems of criminality that demand positive policy intervention; environmental statistics could highlight looming problems of pollution, especially when industry is expanding rapidly.

3.4 Statistics for Business Growth

It would be a foolish person who started a new business ‘blind’, without any knowledge of the general economic climate or the particular circumstances of the market sector he or she had entered. Here again, official statistics are an invaluable ally. Economic and financial statistics give the background to a country’s economic health. Import/export data will offer clues on the international dimension of your chosen market sector. Figures on household consumption and spending patters might indicate levels of demand for goods and services, while those on retail or consumer prices (inflation) and retail sales (‘factory gate’ prices) will also yield useful business intelligence. Statistics on bankruptcies and company liquidations can also provide useful and cautionary note background information. At a more specialised level, there might be data that focus on particular sectors of the economy, or particular groups of consumers.

The retail or consumer price index (RPI/CPI) is also of great significance for government and business. Business contracts can often contain clauses stipulating that agreed payments are subject to rises in line with the RPI or CPI. Businesses operating in a climate of high inflation could invite ruin if they did not take such a precaution in drawing up contracts. Governments often have ‘contracts’ with their citizens based on the RPI/CPI – for example, the level of state benefits or pensions might rise in line with inflation.

3.5 Using data to Improve People’s Lives

In addition to managing the economy, official statistics are also needed to monitor the welfare, or well-being, of people. All governments have a concern on the status of their citizen’s health, education and other areas of welfare. All countries have signed up to the United Nations’ Millennium Declaration that requires actions to improve welfare and sets out specific indicators to be tracked between now and 2015. At the same time, many governments have developed specific poverty reduction strategies that also set out targets for improvements in welfare and identify specific indicators to be measured on a regular basis.

Some indicators of well-being focus such as infant mortality and life expectancy at birth are concerned with survival and life experiences. Others focus on specific aspects of welfare such as health or education. Increasingly countries are accepting that they have an obligation not only to monitor welfare, but also to provide regular reports on progress and the information needed to monitor what is happening regionally and at the global level.

The World Health Organization[7] is probably the most important international organization which directs and coordinates authority on international health work that strives to bring the highest level of health to all peoples. The UNESCO[8] Institute of Statistics (UIS) carries out similar tasks on education statistics.

3.6 Statistics to Attract Foreign Investment

Statistics are a crucial guide for firms considering investment in other countries, and for international organisations providing development assistance. The IMF's General Data Dissemination System Site was established in 1997 to provide a framework for evaluating needs for data improvement and setting priorities in this respect; and to guide member countries in the dissemination to the public of comprehensive, timely, accessible, and reliable economic, financial, and socio-demographic statistics. The website provides information on data produced and disseminated by member countries that participate in the GDDS[9].

---

[7] See WHO website – http://www.who.int
[8] See UNESCO website – http:// www.uis.unesco.org/
[9] See IMF website for GDDS - http://dsbb.imf.org/Applications/web/gddshome/

Data are essential for world financial operations. International organisations, financial institutions and banks rely on statistics to evaluate their investments. But statistics can be unreliable, prone to different interpretations or too dated to be useful. Or they could be reliable but unusable if international investors do not trust them. To avoid such risks and with a view to becoming players on the global financial stage, by the end of 2000, 47 countries had agreed to submit their statistical systems to the rules of the IMF's Special Data Dissemination Standard. A list of the data the IMF maintains (and that are considered particularly important for investors) is provided. The website provides information about economic and financial data disseminated by member countries that subscribe to the SDDS[10].

4. Summary and Conclusion

The word "Statistics" has different meaning to people of diverse backgrounds and interest. Some people view statistics as a field of hocus-pocus whereby a person in the known overwhelms the amateur. To other people it is a way of collecting and displaying large amounts of numerical information. And still to another group it is a way of making decision in the face of uncertainty. Generally speaking, Statistics is concerned with collecting, collating, summarizing, presenting and analyzing data as well as drawing valid conclusions and making reasonable decisions on the basis of this analysis. Careful use of standard statistical methods enables us to accurately describe the finding of a scientific or socio-economic research.

Statistics is use in almost all areas of the physical and social sciences. The professional fields of engineering, education and business all employ statistics in setting standard, establishing policies and planning. The Civil Engineer can use statistics to determine the properties of various materials, the school superintendent may use statistics to mould curriculum, and the Business manager may employ statistics to forecast sales, design products and produce goods more efficiently. The role of statistics in the social sciences, especially in psychology, sociology and economics is a critical one. Here the behavior of individuals and organizations often must be monitored through numerical data to lend credence to models and theories that cannot be supported by theoretical arguments alone.

Despite the need of statistics as public goods for national development, it is difficult to establish functioning markets in statistics. This leaves national governments, institutions and international agencies to produce and disseminate statistical information. These national governments, institutions and international organizations have a duty, therefore, to compile and publish balance of payments, monetary, banking, other financial and official statistics by providing adequate funds and other support to their Statistical Agencies.

References

Fellegi, I. P. (1975). **Characteristics of an Effective Statistical System**. Presented at the Washington Statistical Society – Morris Hansen Lectures and also published in the International Statistical Review in 1996.

Doguwa, S. I. (1999). An Overview of Nigeria's Central Bank Statistics within the National Statistical Information System. **Irvin Fisher Committee Bulletin** No 5 pp 89-96.

[10] See IMF website for SDDS - http://dsbb.imf.org/Applications/web/sddshome/