THE INTERNATIONAL JOURNAL OF
SCIENCE & TECHNOLEDGE

Optimal Funds Flow: A Case Study of the Sustainability of the Ghana National Health Insurance Scheme

Mutawakilu Yahaya
Department of Mathematics/ICT Education, Tamale College of Education, Ghana

Abstract:
The study was carried out to determine the state of National Health Insurance Scheme, the long term stability behaviour and the optimal investment policy for the fund. As a result, quarterly data for both inflows and outflows were modeled as vector autoregressive of order one (VAR (1)) process. Both the system and control matrix was identified and stability analysis was performed. The present state of the system however is not sustainable. It is therefore recommended that an optimal investment policy be adopted.

Keywords: *National health insurance scheme, sustainability*

1. Introduction

The health sector is central to Ghana's developmental agenda, while improving health is intrinsically desirable, it is broadly recognized that health is a necessary pre-requisite for socio-economic development. Improving health will improve human capital, productivity and wealth (MOH, 2004). As a result, many health care financing options have been explored and experimented by governments since independence. Among these includes the Cash and Carry system which existed since 1980's and the current Health Insurance Scheme which was piloted in 2001 and finally enacted into law in 2003(NHI Act 2003; Act 650).

According to Atim, Grey and Apoya (2001), NHI Schemes based on the European Social Health Insurance (SHI) model, have tended to be unsuccessful. Conventional, Social Health Insurance models depend largely on the ability of government to enforce compulsory membership through the deduction of payroll taxes. They are therefore most suited to the context in which there are high levels of well paid, well regulated formal employment (Cichon, Newbrander, Yamabana,& Preker, 2003). This stands in stark contrast to Africa where labour markets tend to be dominated by poorly paid and unregulated informal employment. The health financing reforms that are taking place in Africa aim largely to address the problems of affordability and access to health care. Most African countries have positive expectations forsocial/national health insurance and this view has been largely propagated by the World Bank and other multilateral and bilateral organizations. However, thereis some evidence that, large scale implementation of health insurance scheme is a complex task, and in some instances, perpetuates the very iniquities it is intended to correct (Nguyen &Akal, 2003). There is equally anecdotal evidence to suggest that poorly designed schemes can have very negative consequences.

As noted by Bennett, Creeseand Monasch (1998) and cited in Atim Grey and Apoya (2001) have expressed a similar view and are even less optimistic of Community Health Insurance. They argued that, their risks pools are often too small, adverse selection problems are frequent and the schemes are heavily dependent on subsidies, which are most often infrequent and unreliable. Jutting (2003) noted that, the scheme that experiences managerial and financial difficulties most are those in the environment of rural and remote areas where unit transaction costs of contracts are often too high.

The financial viability of national health insurance scheme is also a matter of concern. For example, Mossialos, Dixon, Figueras, and Kutzin (2002) report that, France social insurance contributions reached an untenable 55% of wages costs and government had to propose a gradual shift to taxation, which is being implemented. It is worthy to note that there is no universal model for the design of national health insurance schemes and that explains why countries like Germany, France and the Netherlands that started social health insurance schemes almost a century ago are still evolving design paradigms that would ensure equity as well as financial viability of the schemes.

1.1. Brief History

National Health Insurance became a law in August 2003. The law puts in place a national health insurance system that aims at enabling residents of Ghana to obtain essential health care services without having to pay for services on demand. The passing of the insurance law was against the background that the government funded about eighty percent(80%) of the public health services bill through taxation and donor funding and the rest from user fees, which was unsustainable as levels of poverty dwindled tax base. This was also compounded with diminished foreign aid which further reduces the capacity of vulnerable in the society to have access to health care services (GOG, 2003).

1.2. Structure and Functions

The directorate of administration which was responsible for projects, procurement and human resource requirements of the Authority has had carved out of it directorate for projects and procurement, headed by a well-qualified professional in the field and a human resource directorate headed by a Deputy Director. Another unit which has been strengthened to place it in good stead to perform its functions better is the Internal Audit Directorate. It is headed by a well-qualified Chartered Accountant as Chief Internal Auditor of the Authority. The unit has generally been strengthened with qualified accounting professionals in the regions and the schemes across the country. The lessons learnt from snap audits under- taken across the country revealed horrendous fraudulent practices in the schemes leading to large financial bleeding of the schemes and great loss to the Ghanaian tax payer (2011 Audit Report).

The newest of the creations to strengthen the Authority in its right against fraud is the Clinical Audit Unit (CAU) which is headed by a qualified medical doctor.Clinicians are being recruited to staff the unit in order to properly vet claims submitted by healthcare service providers. The National Health Insurance Scheme has come to be accepted as a welcome relief for Ghanaians in the area of health care financing. The challenge now is how to sustain the system and make it viable into the future. That is the challenge facing the new management of the authority and the current Government.

1.3. Inflows

The National Health Insurance Scheme has one of the most stable sources of funds. Inflows into the funds are of two main sources.

- Contributions from members
- Revenue generated from tax

1.3.1. Contributions

Contributions are recovered directly from pension fund, that is, two and one half per cent of each person's contribution to the Social Security and Pensions Scheme Fund of some members levy every month by the finance officer and paid into the accounts of the scheme. These contributions are made with respect to each at a rate of two and one half per cent fixed by government and authorities of the scheme. All persons in the informal sector will have to make direct contributions into their District Health Insurance Fund to either register or renew their subscription. Other contribution includes;

- Such monies as may be allocated to the Fund by Parliament.
- Grants, donations, gifts and any other voluntary contribution made to the Fund and Money that may accrue to the Fund from investments made by the National Health Insurance Council.

1.3.2. Tax Revenue

The National Health Insurance Levy (NHIL) is a levy on goods and services supplied in or imported into Ghana. All goods and services are subject to the levy unless they are exempt. The Levy is charged at a rate of two and one half percent on the VAT-excusive selling price of the goods supplied or services rendered. The levy is collected by the Domestic Tax Revenue Division of the Ghana Revenue Authority through VAT-registered persons in the same way that VAT is collected. The Ghana Revenue Authority (GRA) will pay the collected levy directly into the National Health Insurance Fund within thirty days of collection. The levy is to partly finance the National Health Insurance Scheme (NHIS). More than sixty percent (60%) of the resources of the Fund are expected to be obtained from the National Health Insurance Levy.

1.3.3. Communications Service Tax

The Communications Service Tax (CST) is a tax levied on charges for the use of communications services that are provided by communications service operators. CST is imposed under Section 1 of the Communications Service Tax Act 2008, (Act 754). It is paid by consumers to the communications service providers, who in turn pay all CST collected to the Domestic Tax Revenue Division of the Ghana Revenue Authority on a monthly basis. The GRA is required under the law, to pay the CST collected into the Consolidated Fund. It is important to note that at least twenty one (20%) of the Revenue generated from the CST is used by government to finance the National Youth Employment Programme (NYEP) now GYEEDA in particular and support the national development agenda of the country in general.

1.3.4. VAT and NHIL Account

To simplify record keeping, VAT and NHIL are to be combined into one account. This account should be treated in exactly the same way as the previous VAT account which it replaces.

1.3.5. Submission of Returns

VAT registered persons are expected to submit a monthly VAT and NHIL return on the modified VAT return on the same basis as the previous VAT return was being submitted.

1.3.6. Other Sources

The other sources includes; interest on bank accounts, dividend from investment and donations made to the fund.

1.4. Outflows

Outflows from the scheme are in the form of payment to service providers, bank charges, and management expenses. Management expenses can be predictable and anticipated well in advance of the date payment must be made from the scheme, however, payment to service providers are less predictable, but these demands normally represent a large percentage of the fund outflow.

1.5. Statement of the Problem

Under Act 650, the Ghana Health Service (GHS) has been mandated to provide accessible, affordable and quality services to the clients of the National Health Insurance Scheme. Providers are therefore an essential part of the insurance program. The law provides for provider- purchaser split, where payers and providers of health service are dependent on each other. This calls for the payer to enter into contract with accredited providers to provide the agreed minimum benefit package.

However, nearly thirty five percent of Ghanaians fall below the stipulated poverty line (WHO, 2000), while the national health insurance policy framework indicated that hundred percent collection premium cannot be realized from forty percent of the seventy percent of the population of the informal sector (MOH, 2004). Such situation affects sustainability and increases dependency burden on the scheme. To improve on the health service delivery is a collective responsibility of both the government and the governed. In the light of this, failure of the free medical system under Nkrumah's government was partly blamed on the low-participation by the end-user. In recognition of that fact and its consequences on the health service delivery system, the end-user was required to pay money at point of service delivery which came to be known as Cash and Carry system. This payment system compounded the utilization problem by creating a financial barrier to health care access especially for the poor.

From the point of observation, the NHIS in Ghana is bedeviled with several challenges and this development has affected various health providers' ability to increase client's patronage as well as provide quality health service in the country. In order to contribute to understanding and solving these problems an empirical research is necessary to assess the optimal funds flow of the Scheme.

1.6. Objectives of the Study

The purpose of the study was to identifyan appropriate measure for the financial sustainability of the Health Insurance Scheme.

Specifically, the study intends to examine:

- The state of the National Health Insurance fund by considering a mathematical model for the fund flow.
- The stability and sustainability of the funds flow.

2. **Literature Review**

2.1. Theoretical Framework of Time Series Model

A time-series model uses past data as the basis for estimating future results. The models that fall into this category include decomposition, moving average, exponential smoothing, and Box-Jenkins. The premise of a causal model is that a particular outcome is directly influenced by some other predictable factor. These techniques include regression models. Judgmental techniques are often called subjective because they rely on intuition, opinions, and probability to derive the forecast. These techniques include expert opinion, Delphi, sales force composite, customer expectations (customer surveys), and simulation (Kress, 1985; Wilson and Keating, 1994) among others.

2.2. Model Based Prognostics

Autoregressive moving average (ARMA) models are used in time series analysis to describe time series data. Once the ARMA model is determined, it is used to estimate and predict future values of time series with the past values.

ARMA model consists of two parts: an autoregressive (AR) part and a moving average (MA) part. The model is usually referred to as the ARMA (p, q) model, where p is the order of the autoregressive part and q is the order of the moving average part. The notation ARMA (p, q) refers to the model with p autoregressive terms and q moving average terms and is formulated as follows:

$$X_t = c + \varepsilon_t + \sum_{i-1}^{p} \varphi_i X_{t-1} + \sum_{i-1}^{q} \theta_i \varepsilon_{t-1}$$

The summation expression with p number of terms with white noise (ε_t) comes from AR (p) model while the summation expression with q number terms and a constant (c) comes from MA (q) model.

Generally, ARMA models are commonly used when modeling the conditional mean of a (strictly) stationary time series. In conventional terminology (Brock-well and Davis (2006)), an ARMA process is called causal if, at each point in time, its components can be expressed as a weighted sum of present and past error terms.

On the other hand, it is called invertible if these error terms can be represented as a weighted sum of the present and past components of theprocess. Stationarity and invertibility are typically expressed by requiring the autoregressive and moving average polynomials to have their roots outside the unit circle. If causality (invertibility) does not hold, the model is called noncausal or noninvertible (Rosenblatt 2000). Much of the literature on ARMA models considers only the conventional stationary and invertible case. A reason for this is that if the error terms are independent and identically distributed (IID) with a Gaussian distribution, the observed process forms a Gaussian sequence, and in this case a noncausal or noninvertible ARMA model will be statistically indistinguishable from a particular causal and invertible ARMA model (Rosenblatt 2000). Therefore in the Gaussian case, causality and invertibility are often imposed to ensure identification.

However, in many applications, it seems more reasonable to allow the observed process to be potentially non-Gaussian. Alternatively, after fitting a causal and invertible ARMA model to an observed time series one may find that the residuals appear non-Gaussian. In such cases, noncausal or noninvertible ARMA models may be more appropriate and can be distinguished from their conventional causal and invertible counterparts (Lanne& Saikkonen 2009), allowing for the possibility of non-causality or noninvertibility can in such cases also lead to better fit and increased forecast accuracy (Breidt 2003). Noncausal or noninvertible ARMA models have found applications in various fields. Many of the early applications were in natural sciences or engineering, but recently there have also been applications to economic and financial time series (Huang and Pawitan 2000, Breidt, Davis&Trindade,2001).

Maximum likelihood (ML) estimation in noncausal or noninvertible ARMA models has been studied in a number of papers. Brockwell and Davis (1996) discussed the case of noncausal AR models whiles Lii and Rosenblatt (1996) discussed noninvertible MA models, and Lii and Rosenblatt (1996) looked at noncausal or noninvertible ARMA models. However, Andrews, Davis and Breidt (2007) consider the “all-pass models”, which are noncausal or noninvertible ARMA models, in which all the roots of the autoregressive polynomial are reciprocals of the roots of the moving average polynomial and vice versa. Estimation of all-pass models based on the least absolute deviation criterion and rank-based methods are considered in Andrews, Davis and Breidt (2007). All of the above-mentioned literature on noncausal or noninvertible ARMA models considers the case in which the errors are IID. Unlike in the causal and invertible case, the observed process will nevertheless be conditionally heteroskedastic.

Indeed, Breidt, Davis, and Trindade (2001) motivate all-pass models as alternatives to nonlinear models with time varying conditional variances, such as Autoregressive Conditionally Heteroskedastic (ARCH) models. However, they note that “while all-pass models can generate examples of linear time series with nonlinear behavior, their dependence structure is highly constrained, limiting their ability to compete with ARCH”. It is therefore of interest to consider noncausal or non-invertible ARMA models with errors that are not IID but themselves conditionally heteroskedastic, such models may be more appropriate in many applications, especially those in economics and finance.

This paper is in an attempt in combining noncausal or noninvertible ARMA models and Savings/investment-type models. The work considered a particular noninvertible ARMA model with errors that are not IID, but dependent, following a standard MA model. As discussed above, such models may be particularly appealing in economic and financial applications.There are a number of methods (Box, Jenkins& Reinsel, 1994) for estimating the parameters of an ARMA model. Although these methods are equivalent asymptotically, in the sense that estimates tend to the same normal distribution, there are large differences in finite sample properties. In a comparative study of software packages, Huang & Pawitan (2000)and cited in Andrews, Davis and Breidt (2006)showed that this difference can be quite substantial and, as a consequence, may influence forecasts. They recommended the use of full maximum likelihood.

Landsman (1989) presented evidence that the James-Stein ARIMA parameter estimator improves forecast accuracy relative to other methods. If a time series is known to follow a univariate ARIMA model, forecasts using disaggregated observations are at least as good as forecasts using aggregated observations. However, in practical applications, there are other factors to be considered, such as missing values in disaggregated series.

The problem of incorporating external (prior) information in the univariate ARIMA forecasts has been considered (Rosenblatt, 2000; Lanne&Saikkonen, 2009). As an alternative to the univariate ARIMA methodology, Parzen (1982) proposed the ARARMA methodology. The key idea is that a time series is transformed from a long-memory AR filter to a short memory filter, thus avoiding the bharsher Q differencing operator. In addition, a different approach to theconventional T Box Jenkins identification step is used. In the M-competition Makridakis, Andersen, Carbone, Fildes, Hibon, Lewandowski, Newton, Parzen& Winkler, 1982), the ARARMA models achieved the lowest MAPE for longer forecast horizons. Hence, it is surprising to find that, apart from the paper by Meade and Smith (1991), the ARARMA methodology has not really taken off in applied work. Its ultimate value is better judged by assessing the study of Meade (2000) who compared the forecasting performance of an automated and non-automated ARARMA method. Automatic univariate ARIMA modelling has been shown to produce one-step-ahead forecasts as accurate as those produced by competent modellers (Parzen, 1981). Several software vendors have implemented automated time series forecasting methods including multivariate methods (Brockwell&Davis, 1996). Often these methods act as black boxes.

The technology of expert systems (Mefilard & Pasteels, 2000) can be used to avoid this problem. Instead of adopting a single AR model for all forecast horizons, Kang (2003) empirically investigated the case of using a multi-step-ahead forecasting AR model selected separately for each horizon. The forecasting performance of the multi-step-ahead procedure appears to depend on, among other things, optimal order selection criteria, forecast periods, forecast horizons, and the time series to be forecast.

3. Methodology

3.1. Introduction

In this chapter we look at autoregressive time series model. This was extended to vector autoregressive time series model of order one and income and expenditure model together with linear programming model. The vector autoregressive time series model of order one [VAR (1)] is given as

$$X_t = AX_{t-1} + \varepsilon_t$$

Where X_t is a vector time series, A is a constant matrix and ε_t is an error term. The deterministic part of the equation leads us to a system of linear discrete time dynamical autonomous system and vector difference equations. Stability and solutions to discrete time dynamical autonomous system for the homogenous and non-homogenous cases was given much attention.

3.2. Time Series Analysis

Given that the time series values observed are the realizations of random variables $y_1,\ y_2,\ y_3, \dots\ y_t$ which are in turn part of a stochastic process$y_t: t\ \in Z$. Although it is best to distinguish the observed time series from the underlying stochastic process, the distinction is usually blurred and the term time series is used to refer to both the observations and the underlying process that generates them.

<u>3.2.1. Mean and Auto covariance Function</u>

The mean function of time series is defined as $\mu_t = E(Y_t)$and the autocovariance function is defined to be $\gamma(s,t) = Cov(Y_s, Y_t)$.The mean and autocovariance functions are fundamental parameters and it would be useful to obtained sample estimates for them. For general time series there are $2T + T(T-1)/2$parameters associated with $y_1,\ y_2,\ y_3, \dots\ y_t$and it is not possible to estimate all these parametersfrom T data values. In this case we must impose constraints on the time series.The most common constraint is the stationarity.

<u>3.2.2. Stationarity</u>

A time series X_t is strictly stationary if $X_t,\ X_{t+s} = X_{t+r},\ X_{t+r+s}$for all r and s. In practice, we often relax the requirement of stationarity by considering only the weak stationarity. A time series X_t is weakly stationary if

- $E(X_t) = \mu$, a constant
- $Cov(X_t, X_{t+k}) = \gamma_k$, a function depending only on k.

<u>3.2.3. Autocorrelation Function</u>

For the given observed time series $X_1,\ X_2, X_3, \dots\ X_n$ the sample autocorrelation function (ACF) was defined as;

$$r_j = \frac{\sum_{t=1}^{n-j}(x_i - x)(x_{i+j} - x)}{\sum_{t=1}^{n-j}[(x_i - x)]^2}$$

Where $x = \frac{\sum_{i-1}^{n} x_i}{n}$is the sample mean of the series,$x_t$are the observations in the time series and $1 \le r_j \ge 1$; where $r_j$is the autocorrelation function of time lag j; $r_j = 0$indicates a purely random time series showing that the observations in the series are completely independent of each other.

The ACF is an important guide to the properties of a time series. It measures the correlation between observations at different distances apart. This behavior is a powerful tool to identify a preliminary model for the time series. Bartlett (1946) as cited in Gabriel (2002) indicates that if there is no correlation among observations that are more than q steps apart $\rho_j = 0\ for\ j > q$, the variance of $r_j$can be approximated.

3.3. Autoregressive Models

An autoregressive model of order p [AR(p)] states that y_i is the linear function of the previous p values of the series plus an error term

$$y_i = \varphi_1 y_{i-1} + \varphi_2 y_{i-2} + \varphi_3 y_{i-3} + \dots + \varphi_p y_{i-p} + \varepsilon_i$$

Where $\varphi_1, \varphi_2, \dots, \varphi_p$ are weights that determined and $\varepsilon_i$are normally distributed with zero mean. Thus the model says that y_i is normally distributed with a mean $\sum_{j-1}^{p} \varphi_j y_{i-j}$and variance σ^2

3.3.1. Autoregressive Series of Order One; AR (1)

The generating process of the simplest model is, AR (1) where

$$y_i = \varphi_1 y_{i-1} + \varepsilon_i$$

For fitting an AR (1) model, we have the observations $y_1,\ y_2,\ \dots\ y_n$ that defines thelinear system of n-1 equations:

$$\begin{aligned} y_2 &= \varphi_1 y_1 + \varepsilon_2 \\ y_3 &= \varphi_1 y_2 + \varepsilon_3 \\ \vdots &= \cdots \quad \vdots \\ y_n &= \varphi_1 y_{n-1} + \varepsilon_n \end{aligned}$$

For a lagged series defined by $x_i = y_{i-1}$ for $i = 2, \dots, n$. The AR(1) model is then $y_i = \varphi_1 x_i + \varepsilon_i$. We say that the generating process of the AR(1) model is regressedon x. This is just the linear lagged series that will give us the best values for theparameter φ_1 and an estimate

$$\sigma^2 = \frac{1}{n\,(1+2)\sum_{i=1}^{inf} {\rho_i}^2}\ \forall\ i > q.$$

If all the observations in the series are uncorrelated, then $\rho^2 = 0 \forall\ i > 0$. The above equation reduces to $\sigma^2 \approx \frac{1}{n}$

Since y_{i-1} and ε_i are uncorrelated, the variance of the series is:

$$Var\ (y_i) = \varphi_2\ Var(y_{i-1} + {\sigma_\varepsilon}^2$$

If y_i is stationary, then

var$(y_i) = var(y_{i-1}) = {\sigma_y}^2$ and so ${\sigma_y}^2 = \varphi^2 {\sigma_y}^2 + {\sigma_\varepsilon}^2$

Implying that

$$1 > \varphi^2$$

Using the alternative formulation of equation 3.2, that is,

$$(1 - \varphi L) y_i = \varepsilon_i$$

It is possible to invert the autoregressive operator to obtained

$$(1 - \varphi L)^{-1} = \sum_{u=n+1}^{\infty} \varphi^u\ L^u.$$

Applying this to the series ε_t produces

$$y = \sum_{u=n+1}^{\infty} \varphi^u\ \varepsilon_{t-u} \qquad \text{..........................3.3}$$

If $\varphi < 1$,the series converges in mean square because $\sum \varphi^u < \infty$ .The limit series is stationary and satisfies equation 3.2.

3.3.2. Theorem

If $\varphi < 1$, there is a stationary solution to $y_i = \varphi_1 y_{i-1} + \varphi_2 y_{i-2} + \varphi_3 y_{i-3} + \dots + \varphi_p y_{i-p} + \varepsilon_i$.

However, if $\varphi > 1$, the series does not converge and if $\varphi = 1$, there is no stationary solution.

For the purpose of modeling and forecasting stationary time series, we restricted our attention to series for which $\varphi < 1$.

3.4. Vector Autoregressive Process

Unlike the univariate case, the vector autoregressive processes have a column vector with n variables that is lagged by, say p, previous values. A vector autoregressive process of order p [VAR (p)] for a system of H variables $x_t = (x_{1t}, x_{2t}, x_{3t}, \dots H_T)$ is

$$x_t = \alpha + \sum_{i=1}^{p} \beta_i\, x_{t-1} + \varepsilon_t$$

In this system of H equations, $\alpha = \alpha_1, \alpha_2, \dots \alpha_H$ is an H dimensional vector,

$$\beta_i = \begin{pmatrix} \beta_{11,i}\beta_{12,i}\beta_{13,i} \cdots & \beta_{1H,i} \\ \beta_{21,i}\beta_{22,i}\ \beta_{23,i} \cdots & \beta_{2H,i} \\ \vdots \quad \vdots \quad \vdots \quad \cdots & \vdots \\ \beta_{H1,i}\beta_{H2,i}\beta_{H3,i} \cdots & \beta_{HH,i} \end{pmatrix}$$

are H × H coefficient matrices and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots \varepsilon_{Ht})$same properties as the shock term in the scalar case. The parameters $\alpha, \beta_1, \beta_2, \dots \beta_p$ were unknown and were estimated by the ordinary least squares method from the available data before the process in equation 3.1 was used for the analysis. The stationarity of the VAR were useful in deriving the asymptotic properties of the parameter estimators.

3.4.1. Income and Expenditure Model Using Vector Autoregressive Process

In many instances a single equation model cannot describe the relationship between the variable in a dynamic system. Instead, a system of equations may be required to describe the data generation process adequately. For instance, using monthly variables, the consumption expenditure X_t was assumed to depend on income y_t on the consumption expenditure of the previous period x_{t-1}. Thus, we specify the consumption function as

$$X_t = \alpha_1 + \beta_1 y_t + \beta_2 x_{t-1} + \tau_{1t}$$

Since increased consumption will stimulate economic growth and thereby generate an increase in future income, an income equation could be expressed as

$$y_t = \alpha_2 + \tau_1 x_{t-1} + \tau_2 y_{t-1} + \varepsilon_{2t}$$

Income is assumed to depend on lagged consumption and in addition on the income of the previous period. The inter temporal relationship between the consumption and income variables is dynamic and yield

$$X_t = \alpha_1 + \beta_1 \propto_2 + (\beta_1\tau_1 + \beta_2)x_{t-1} + \beta_1\tau_2 y_{t-1} + (\varepsilon_{1t} + \beta_1\tau_{2t})$$

Setting

$y_1 = \alpha_1 + \beta_1 \propto_2, \theta_1 = \beta_1\tau_1 + \beta_2$ and $\theta_1 = \beta_1\tau_2$ the system was express more compactly as

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \alpha_1 \end{bmatrix} + \begin{bmatrix} \theta_1 & \theta_2 \\ \tau_1 & \tau_2 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$ and is said to be Bivariate VAR (1).

4. Results

4.1. Data Collection

Secondary data was used in respect with accounting variables from NHIA office covering a period of 2009 to 2017. The variables selected for the study of the NHIA system was classified into inflows and outflows as shown in table 6.

4.2. Classification of Variables

The National Health Insurance fund can be view as a system of inflows and outflows. The study identified six variables for both inflows and outflows. However, we chose the net inflows and outflows as state variables (see Appendix 1).

4.3. Statistical Identification of the System

We assume a model and fit the categorized data into the model. We also estimate the parameters in the model and the adequacy of the model was then checked using ordinary least squares method.

4.3.1. Construction of Model

We first identify the system by fitting both the inflow and outflow data into the homogenous model

$$X_t = AX_{t-1} + \varepsilon_t \quad \text{...4.1}$$

The system matrix A was then analyze for stability issues. However, the model constructed for the NHIA system was of the form:

$$X_t = AX_{t-1} + BU_{t-1} + \varepsilon_t \quad \text{...4.2}$$

Where X_t is state vector of the i^{th} year and U_t is control variable of NHIA system. A is the system matrix and B a control matrix. Since the NHIA system has both inflow and outflow as state variables. Since the NHIA system has both inflow and outflow as state variables, the state vector X_t is given by

$$X_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$$

Where x_t is inflow at the i^{th} year and y_t is corresponding outflow at the i^{th} year. We also considered type of investment as the control variables. Thus, investment in Money Market Funds (Treasury Bills), investment in Growth or Equity Funds (Shares) and investment in Fixed Income Funds. The control vector for the i^{th} year is

$$U_t = \begin{bmatrix} M_t \\ E_t \\ F_t \end{bmatrix}$$

Where M_t is investment in Money Market Funds, E_t is investment in Equity Funds and F_t is investment in fixed income funds. Next we fitted the data into the forced system in 4.2 and the equation becomes;

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \rho_{1,1} & \rho_{1,2} & \rho_{1,3} \\ \rho_{2,1} & \rho_{2,2} & \rho_{2,3} \end{bmatrix} \begin{bmatrix} M_t \\ E_t \\ F_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \varepsilon_t \end{bmatrix}$$

And the parameters were estimated and the following results were obtained.

Parameter	Estimate
$\beta_{1,1}$	0.8100
$\beta_{1,2}$	0.7946
$\beta_{2,1}$	0.8388
$\beta_{2,2}$	0.8242

Table 1: Parameters of the homogenous model
Source: Authors Construct with Data from NHIA, (2017)

The matrix representation of the forced system is:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0.8100 & 0.7946 \\ 0.8388 & 0.8242 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} \quad \text{...4.1.1}$$

The error term being zero, which is a homogenous vector difference equation representing a deterministic dynamic model.

Parameter	Estimate
$\beta_{1,1}$	0.6219
$\beta_{1,2}$	0.9204
$\beta_{2,1}$	0.4926
$\beta_{2,2}$	-0.7280
$\rho_{1,1}$	0.3246
$\rho_{1,2}$	0.4988
$\rho_{1,3}$	0.8746
$\rho_{2,1}$	0.5932
$\rho_{2,2}$	0.0053
$\rho_{2,3}$	0.9382

Table 2: Parameters of the Unforced System
Source: Authors Construct with Data from NHIA, (2017)

Similarly, the matrix representation of the unforced system is:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0.6219 & 0.9204 \\ 0.4926 & -0.7280 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 0.3246 & 0.4988 & 0.8746 \\ 0.5932 & 0.0053 & 0.9382 \end{bmatrix} \begin{bmatrix} 5.3782569 \times 10^7 \\ 7.6074112 \times 10^7 \\ 9.6680057 \times 10^7 \end{bmatrix}$$

Where $\begin{bmatrix} 5.3782569 \times 10^7 \\ 7.6074112 \times 10^7 \\ 9.6680057 \times 10^7 \end{bmatrix}$ is the mean of each of the investment used as the control vector.

The system was written as:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0.6219 & 0.9204 \\ 0.4926 & -0.7280 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 1.3996 \times 10^8 \\ 1.2301 \times 10^8 \end{bmatrix} \quad \text{...4.2.1}$$

4.4. Stability of the System

We examined the stability of both the forced and the unforced systems by determining the eigen values of the system matrix. A particular stability behavior depends upon the existence of real and imaginary components of the Eigen values, along with the signs of the real components and the distinctness of their values. We examine each of the possible cases in the model.

4.4.1. The Unforced System

The Eigenvalues of the system matrix were of the form $a + b_i$(i.e. -0.5132 and 2.1474). With the real part being positive, the system is unstable and behaves as an unstable oscillator. (That is, since a > 0, the trajectories spiral away from the origin, and (0,0) is an Unstable Equilibrium). This can be visualized as a vector tracing a spiral away from the fixed point. The plot of response with time of this situation would look sinusoidal with ever increasing amplitude, as shown below.

This situation is usually undesirable when attempting to control a process or unit. If there is a change in the process, arising from the process itself or from an external disturbance, the system itself will not go back to steady state.

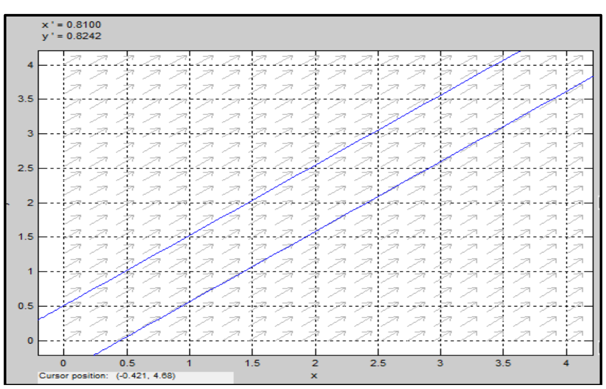

Figure 1: Phase portrait of the unforced system
Source: Authors Construct with Data from NHIA, (2017)

4.4.2. The Forced System

The eigenvalues of the force system matrix were of the form $-a + b_i$ with λ_1 = 0.9003 and λ_2 = -1.0064. Since the real part is negative, then the system is stable and behaves as a damped oscillator. This can be visualized as a vector tracing a spiral toward the fixed point. The plot of response with time of this situation looks sinusoidal with ever-decreasing amplitude, as shown below.

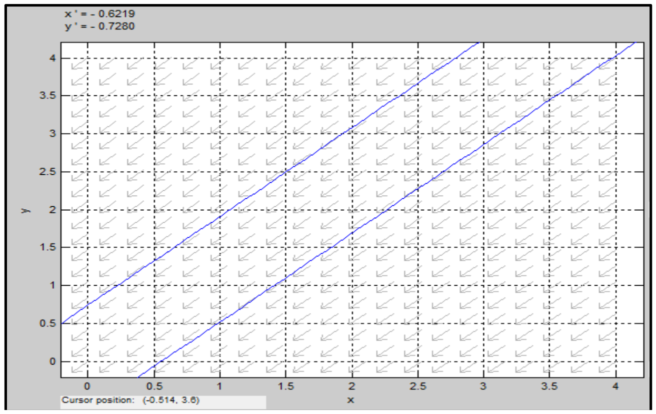

Figure 2: Phase portrait of the force system
Source: Authors Construct with Data from NHIA, (2017)

4.5. Model Diagnostics

A run sequence plot of the response variables indicates the presence of outliers and seasonal patterns as shown below.

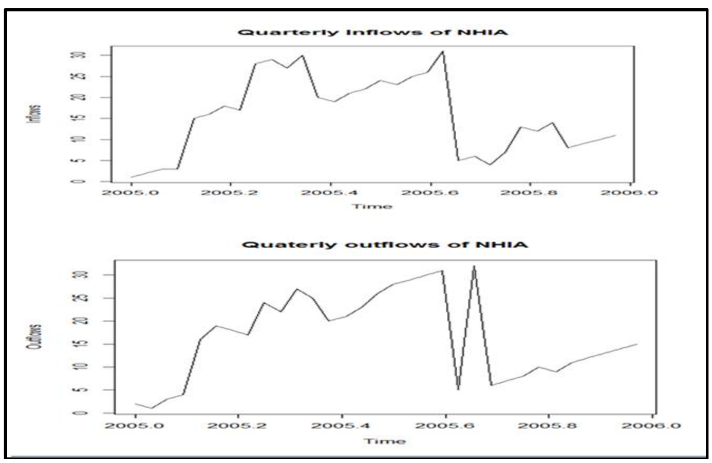

Figure 2 & Figure 3: Run sequence plot of inflows and outflows
Source: Authors Construct with Data from NHIA (2017)

Figure 1 shows a time series plot of the original data. It is clear from thisfigure that there is an increasing trend in this time series, so we conclude that the original time series is not trend-stationary. That is, there does not seem to be a significant trend or any obvious seasonal pattern in the data. So a run sequence plot of differenced data was examined. That is, we consider difference-stationarity. An Augmented Dickey-Fuller (ADF) test was applied to test for the presence of a unit root in the time series.

The ADF test on the original data provided evidence of the presence of a unit root (p-value0.5203; where the null hypothesis assumes that the data is non-stationary i.e. there is a unit root present in the data). We difference the data to counteract the effect of this unit root. Again we apply the ADF test to formally test for the presence of a unit root. We also test for consistency in the joint distribution of samples of the first-differenced time series. The ADF test (p-value: 0.0835) failed to reject the null hypothesis of stationarity at the 5% level (though it is a borderline rejection). These tests, alongside the graphical evidence seen in the time series plot of the first-difference of the time series provide sufficient evidence that the first-difference of the time series is non-stationary (See Appendix). Therefore the data is differenced for a second time.

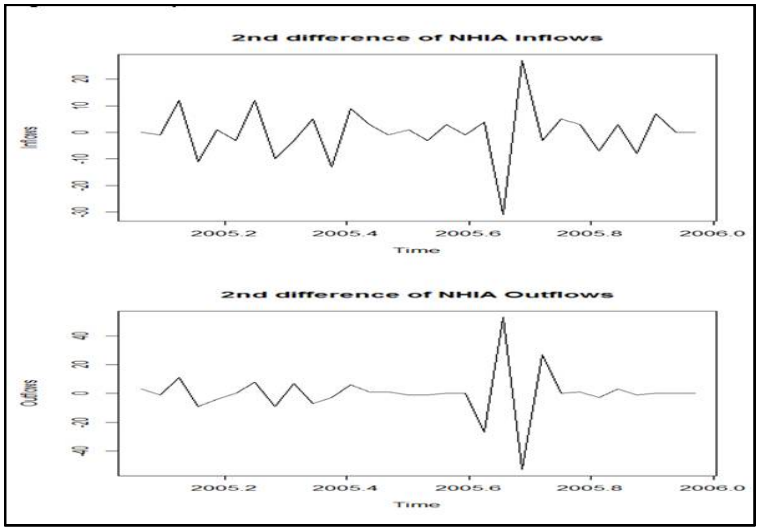

Figure 4: Stationary time series
Source: Authors Construct with Data from NHIA, (2017)

The run sequence plot of the differenced data shows that the mean of the differenced data is around zero, with the differenced data less auto correlated than the original data. The next step was to examined the sample autocorrelations of the differenced data.

A Plots of Autocorrelation Function of Residuals of both Inflows and OutflowsA plot of autocorrelation of the two set of series (inflows and outflows) were examined for stationarity and the order of the series respectively. The autocorrelation plot together with run sequence of the differenced data suggests that the differenced data are stationary. Based on the autocorrelation plot, an MA (1) model was suggested for the differenced data.

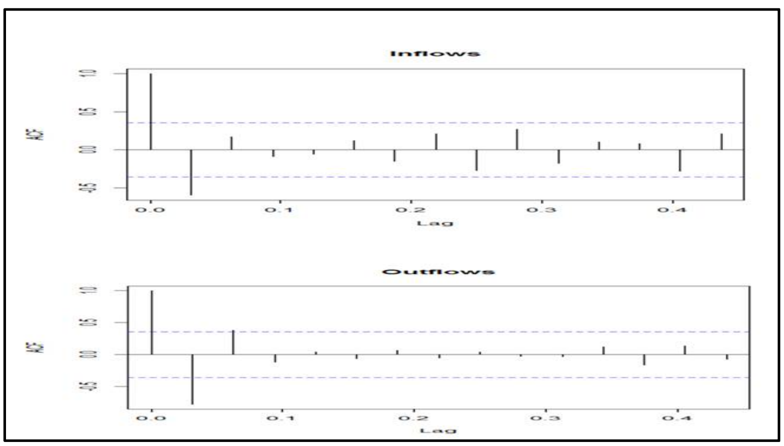

Figure 5: Autocorrelation plot of the difference data
Source: Authors Construct with Data from NHIA(2017)

4.5.1. Partial Autocorrelation Plots of Residuals of both Inflows and Outflows

To examine other possible models, we produce the partial autocorrelation plot of the differenced data. The time series, ACF and PACF plots of the second-differenced time series indicates that the second-differenced time series is stationary (constant mean and approximately constant variance). The ADF test also confirm this (p-values = 0.01). This situation is what is generally desired when attempting to control a process or unit. This system is stable since steady state will be reached even after a disturbance to the system. The oscillation will quickly bring the system back to the set point, but will over shoot, so if overshooting is a large concern, increased damping would be needed.

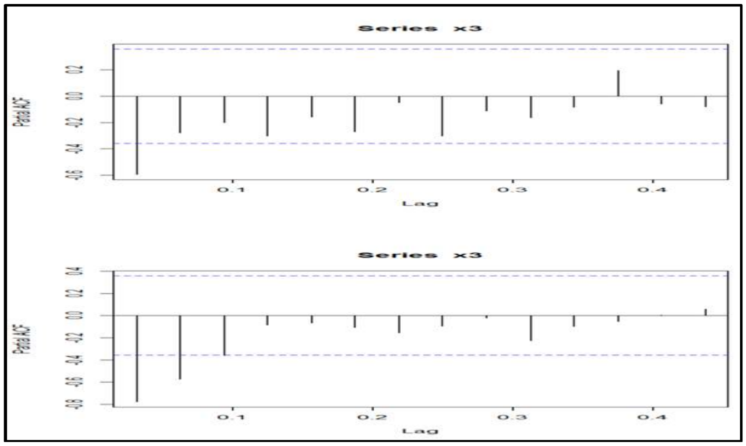

Figure 6: Partial Autocorrelation plot of the difference data
Source: Authors Construct with Data from NHIA, (2017)

4.6. Model Identification

Having determined the stationarity characteristics of the original data (i.e. the original data is an I (2) time series), we than apply graphical techniques to determine the initial ARIMA model for the original time series.

The ACF and PACF plots (Figure 3 and 4) of the stationary data (the second- differenced data) were used here. There is evidence of decay on the PACF plot. The ACF plot show two non-zero values. (The first value represents ρ_0 and is identically equal to one. It is therefore ignored within the analysis.) These plots indicate that an MA (1) model would be appropriate for the second-differenced data. This is equivalent to fitting an ARIMA(0, 2, 1) to the original time series. An ARIMA(0, 2, 1) model may be written as follows:

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1}$$

Where Y_t represents the quarterly inflows time series. Following this, an ARIMA(0,2,1) model was fitted to the original time series. The estimate of the coefficient was reported in Table 3 alongside the standard error and AIC of the estimate.

ARIMA(0,2,1)	Estimate	SE	AIC
θ_1	39.83	0.3214	203.11

Table 3: ARIMA (0,2,1) model - Estimated Model
Source: Authors Construct with Data from NHIA,(2017)

However, residual diagnostic tests and the over fitting process were used to determine the goodness-of-fit of the ARIMA (0, 2, 1) model to the original time series.

4.6.1. Residual Diagnostics

We first consider the residuals of the model. Figure 7 shows three diagnostic plots that are useful here. Firstly, the time series plot of the model residuals allow us to look for trends and heteroscedasticity in the residuals. It is clear from the time series plot shown in Figure 7 that the series of residuals are a stationary series (with constant mean and variance). The ACF of the residuals is also used as a diagnostic tool. We see that the ACF values are all within the 95% zero-bound, indicating that there is no correlation amongst the residuals. This plot is used as an indicator of the independence of the residual terms.

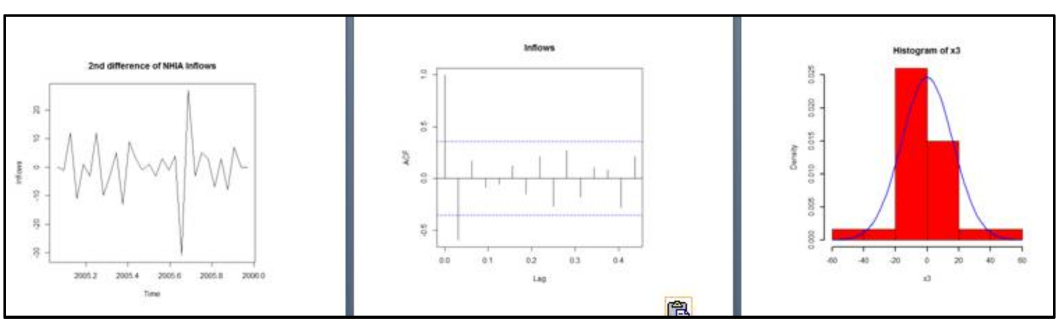

Figure 7: Residual Diagnostic plots - Time Series plot, ACF plot, and Histogram of the ARIMA (0,2,1) residuals
Source: Authors Construct with Data from NHIA (2017)

The histogram plot tests the normality. Since it shows a bell-shaped distribution, this is good indicator of Normality within the residuals. These three plots (and associated tests) confirm that the residuals of the ARIMA (0, 2, 1) model are distributed as white noise indicating that the model fits the data well.

4.7. Over Fitting

The over fitting process is a technique used to test for other possible significant terms (that may have been missed at the model identification stage of the analysis). We begin by over fitting an MA term on the model (i.e. we fit an ARIMA(0, 2, 2) to the original time series). This additional MA term is not significantly different from zero (Table 3). We therefore revert to our original model and over fit an AR term (i.e. we fit an ARIMA(1, 2, 1) to the original time series). Again, the additional term is not significant.

ARIMA(2,2,1)	Estimate	SE	Residual Variance	AIC
AR 1	0.1098	0.2067	39.65	206.49
AR 2	0.0726	0.2017		
MA 1	0.9780	0.3592		
ARIMA(2,2,0)				
AR 1	0.7382	0.1722	54.61	211.7
AR 2	0.2623	0.1692		
ARIMA (1,2,2)				
AR 1	0.1360	0.1925	40.51	204.62
MA 1	0.9433	0.1530		
ARIMA(1,2,2)				
AR 1	0.4752	0.6439	40.18	206.49
MA 1	0.6070	0.6786		
MA 2	0.3307	0.6650		

Table 4: Over fitting ARIMA(p,q,d) model –Estimated Model
Source: Authors Construct with Data from NHIA, (2017)

4.8. Final Model

We therefore conclude that the ARIMA (0, 2, 1) model is the best-fit ARIMA model for the original time series being analysed here. The final model is of the following form:

$$Y_t = \epsilon_t - 39.83\epsilon_{t-1}$$

Where the estimated MA coefficient θ_1 has a standard error of 0.3214.

<u>4.8.1. Forecasting</u>

Forecasts of future income(inflow) for NHIA are of particular interest to the researcher. We used the final form of the best-fit ARIMA model for the time series to estimate future income values. The forecasted income for the next three quarters is displayed in Table 5 (alongside the standard errors of the estimates).

Month	Estimate	Standard Error
1	11.32258	6.411748
2	11.64516	9.208172
3	11.96774	11.447259
4	12.29032	13.411133
5	12.61290	15.206795
6	2.93548	16.887988
7	13.25806	18.485970
8	13.58065	20.020676
9	13.90323	21.505656
10	14.22581	22.950564
11	14.54839	24.362532
12	14.87097	25.746979

Table 5: Forecasted values using ARIMA(0, 2, 1) model
Source: Authors Construct with Data from NHIA (2017)

It is clear from these forecasts that the monthly inflow of NHIA is expected to follow the positive trend visible in the time series plot of the original data (Figure 1).

5. Conclusion

The study found that the monthly inflows were non-stationary. In particular the original data is I (2). Overall it was determined that the ARIMA (0, 2, 1) was the most appropriate ARIMA model for the data. Future inflows values for the next twelve (12) months were estimated using this model. We found that the forecasted values followed the upward trend present in the data. This may be used as an indicator of the short- term future of the NHIA(where the estimated standard errors of the forecasts are taken into account).

However, the sustainability of the fund depends on the excess cash flow and the expenditure pattern of the NHIA system. The excess cash flow for the system which is the inflow minus outflow of the state vectors is (GHs2, 722,660.35).Comparing with the expenditure pattern over the period which lies around GHs297, 221,142.67 suggest a non-sustainability of the scheme.

Considering the fact that the outflow has two facets (expenditure and investment made), it suffix to reason that much attention should be given to the type of investment the system consider in order to have sufficient returns to boost up the inflow of the system. For now, little can be done about client contributions considering the economic situation of the country. Hence much attention should be placed on the type of investment options available to the scheme.

6. Acknowledgements

This study would not have been a reality without the support of others in terms of material, intellectual and emotional support. I therefore express my gratitude to God almighty for His grace. To Him are glory, honour, and praise for the needed strength and enablement to carry out this study. I also owe a great debt of gratitude and appreciation to my supervisor, Dr. Gabriel Okyere Asare for the rich suggestions, corrections, and direction given to me in the process of this work. He has been meticulous, patient, and helpful. I am really very grateful to him. Again I thank the following for their emotional, material and prayer support. In addition, my family especially my wife, son and daughter, had been of great help to me by keeping me going even in times of difficulties. I am really grateful to them. Finally, I am grateful to all those who gave me various kinds of support in relation to this work. God richly bless you all.

7. Appendix

Description	Variable
Contributions from members	Inflow
Revenue generated from Tax	Inflow
Interest on Treasury Bills	Inflow
Dividend from Shares	Inflow
Interest on Bank account	Inflow
Donations	Inflow
Bank Charges	Outflow
Audit Fees	Outflow
Management Expenses	Outflow
Investment in Treasury bills	Outflow
Shares Purchase	Outflow
Payment to service providers	Outflow

Source: Authors construct with data from NHIA, December 2013.

Table 6: Classification of Variables

8. References

i. Andrews, B. R.A. Davis, R.A & Breidt F.J. (2006). Maximum likelihood estimation for all-pass time series models. Journal of Multivariate Analysis 97, 1638. 1659.

ii. Atim, C., Grey, S., & Apoya P. (2001). A Survey of Health Financing Schemes in Ghana. Bethesda, MD: Partners for Health Reform plus, Abt Associates Inc.

iii. Bennett, S. Creese, A. & Monasch, R. (1998). Health Insurance Schemes for People Outside Formal Sector Employment, ARA Paper No. 16.

iv. Breidt, F. J. (2003). Maximum Likelihood Estimation for All-Pass Time Series Models. (Appear in J. Multivariate Analysis).

v. Breidt, F.J., Davis, R.A. Lii, K.S &Rosenblatt M. (2003). Maximum likelihood estimation for noncausal autoregressive processes. Journal of Multivariate Analysis 36, 175.198.

vi. Breidt, F.J., Davis, R.A. &Trindade A.A (2001). Least absolute deviation estimation for all-pass time series models. The Annals of Statistics 29, 919.946.

vii. Breidt, F.J. & Hsu N. J. (2005). Best mean square prediction for moving averages.Statistical Sinica, 15, 427-446.

viii. Brockwell, P. J. & Davis, R. A. (1996). Introduction to Time Series and Forecasting, Second Edition Springer, Berlin.

ix. Box, G. E. P. Jenkins, G. M. &Reinsel, G. C. (1994). Time SeriesAnalysis, Forecasting and Control, Volume 2, P 734, Prentice-Hall, Englewood Cliffs.

x. Cichon, M. Newbrander, W. Yamabana, H. & Preker, A. (2003). Modelling in Health Care Finance: A Compendium of Qualitative Techniques for Health Care Financing. ISBN 92-2-110862 p.85.

xi. GOG(2003). National Health Insurance Act(Act 650), Government of Ghana. Ghana Publishing Corporation (Assembly Press), Accra-Ghana.

xii. Huang, J. & Pawitan Y. (2000). Quasi-likelihood estimation of noninvertible moving average processes. Scandinavian Journal of Statistics 27, 689.710.

xiii. Jutting, J.(2003).Health Insurance for the Poor: Determinants of Participation in Community-Based Health Insurance Schemes in Rural Senegal. OECD Development Centre, Working 204, OECD, Paris.

xiv. Kim J.(2013). Estimating Item Response Theory Models Using Markov Chain Monte Carlo Methods. An NCME Instructional Module. University of Wisconsin, Madison.

xv. Lanne, M., & Saikkonen P. (2009). Noncausal vector autoregression. Bank of Finland Discussion Paper 18/2009.

xvi. Landsman , D. (1989), Asymptotically optimal shrinkage estimates for non-normal data. Journal of Statistical Computation and Simulation Volume 81,Issue 12.

xvii. Lii, K.-S. & M. Rosenblatt (1996). Maximum likelihood estimation for non- Gaussian nonminimum phase ARMA sequences. Statistica Sinica 6, 1-22.

xviii. Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E.& Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition, *Journal of Forecasting*, 1, 111-153. (all)

xix. Meade, C.D. & Smith, C.F. (1991) Readability formulas: cautions and criteria. Patient Education and Counselling, 17, 151–158.

xx. Ministry of Health (2004). National Health Insurance Scheme: What You Must Know, Accra, Ghana.

xxi. Mossialos, E., Dixon, A., Figueras, J., Kutzin, J.(2002).Funding Health Care: Options for Europe. UK: Open University Press.

xxii. Nguyen T.K.P. &Akal A. (2003). Recent Advances in Social Health in Vietnam: A comprehensive Review of Recent Health Insurance Regulations, Health Financing Master Plan Technical Paper Series-1.

xxiii. Rosenblatt, M. (2000). Gaussian and non-Gaussian linear time series and random yields. Springer Verlag, New York.

xxiv. Rosenblatt, M. (2006). Pile-up Probabilities for the Laplace Likelihood Estimator of a Non-invertible First Order Moving Average IMS Lecture Notes Monograph Series, Vol 52, pp.1-19.

xxv. Parzen, E., (1982). AR–ARMA models for time series analysis casting, Journal of Time Series Analysis, Journal of Business and and forecasting. *Journal of Forecasting*, 1, 67–82.

xxvi. Parzen, E. (1981). Time series model identification and prediction variance horizon. Findley, D. (ed.) *Applied Time Series Analysis*, 11, New York: Academic Press, pp. 415-447. (EP)

xxvii. WHO (2000).The World Health Report. Health Systems, Improving Performance, Geneva. World Health Organisation.