ON GENERATING MECHANISMS AND DETECTION OF OUTLIERS IN MULTIVARIATE TIME SERIES

BY

***OLUFOLABO*, OLUSESAN OLUYOMI**
PDS, B.Sc., M.Sc., (Statistics) (Ibadan)

A DISSERTATION IN THE DEPARTMENT OF STATISTICS SUBMITTED TO THE FACULTY OF SCIENCE IN PARTIAL FULFILLMENT OF THE REQUIREMENT OF THE AWARD OF THE MASTER OF PHILOSOPHY, UNIVERSITY OF IBADAN, IBADAN, NIGERIA.

October, 2014

ABSTRACT

Outliers are aberrant observations that adversely affect parameter estimation and predictive capability of a given model. The problem of outlier detection in time series has gained much attention in the literature and various methods of detection have been developed, but are limited to univariate time series with its attendant swamping effect. This work is focused on developing outlier Generating Mechanisms (GMs) for the detection of outliers in the Multivariate Time Series (MTS) setting that is capable of ameliorating the swamping effect.

Two-variable Vector Autoregressive (VAR) models $X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t}$ and $X_{2t} = \phi_{21} X_{2t-1} + \phi_{22} X_{1t-1} + a_{2t}$ were considered, where X_{it} and X_{jt-1}, i,j=1, 2 were the current and lagged values of the response and explanatory variables respectively, ϕ_{ij} , i,j=1, 2, were coefficients, t is the time and a_{1t} and a_{2t} were distributed as N(0,Σ). Each series was assumed to have been generated by the model $f(Z_t, \varphi_t(\beta), \omega\xi_t^T)$ where Z_t is an outlier free time series, $\xi_t^{(T)}$ is a time indicator where $\xi_t^{(T)} = 1$ for all $t = T$ and $\xi_t^{(T)} = 0$ otherwise, $\varphi_t(\beta) = 1 - \Theta_1 B - \cdots - \Theta_p B^p$ were polynomials of order p and $\omega = (\omega_1, \ldots, \omega_k)'$ were the magnitude of outliers. The nature of effect of outlier on uncontaminated series determines the model which could be Innovative (IO), Additive (AO), Multiplicative (MO), and Convolution (CO) which is the combination of IO and AO effects. These models were used to develop four GMs for detection of outliers in multivariate time series. The magnitudes of outliers and their variances with the test statistics were derived for the four generating mechanisms. Simulation data of sample sizes of 10, 50, and100 were used to establish the validity of the developed models. Data on Nigerian Gross Domestic Product (GDP) and Consumer Price Index (CPI), commercial bank deposits and loans were also used. Estimates of the magnitude and residual variance of outliers were obtained using method of least squares. The percentages of outliers detected for simulated data and the number of detected outliers in data sets were observed. The relative efficiency of the models was evaluated in determining the best outlier generating mechanism.

The developed generating mechanisms were: $X_{it} = \Phi_{ii} X_{jt-1} + \Phi_{ij} \boldsymbol{\xi}_t^T \varphi(\beta)(1+\omega)$, $X_{it} = \Phi_{ii} X_{jt-1} + \Phi_{ij} \boldsymbol{\xi}_{t-1}(\omega+\varphi(\beta))$, $X_{it} = \Phi_{ii} X_{j,t-1} + \Phi_{ij}\omega\boldsymbol{\xi}_t^{T^2} \varphi(\beta)$ and $X_{it} = \Phi_{ii} X_{jt-1} + \Phi_{ij} \boldsymbol{\xi}_t^T [2\varphi(\beta)+\omega(1+\varphi(\beta))]$ for IO, AO, MO and CO respectively. The performance of the generating mechanisms based on simulations showed that the percentages of outliers detected using IO, AO, MO, and CO were 21%, 71%, 86%, and 100% respectively. For GDP and CPI, 30 outliers were detected by CO; 29 each by IO and AO while MO was unable to detect any outlier because it did not exhibit any multiplicative effect on the data. For deposit and loan, 6 outliers each were detected by all the GMs except MO. The CO gave a high precision with low percentage of variation compared with other generating mechanisms. It was observed that whenever the explanatory variable was infested with outlier, the response variable is also contaminated.

The derived outlier generating mechanisms were able to detect potential outlier independently in multivariate time series with the swamping effect ameliorated. The pairwise relative efficiency of the variances indicated that convolution model was the best. It is therefore recommended for outlier detection in multivariate time series setting.

Keywords: Outlier generating mechanism, Vector autoregressive models, Gross domestic product, Consumer price index.

Word count: 498

CERTIFICATION

I certify that this work was carried out by ***Olufolabo, Olusesan Oluyomi*** in the Department of Statistics, University of Ibadan, Ibadan.

……………………………………..

Supervisor
Shittu, I.O.
PDS, B.Sc., M.Sc., Mphil., Ph.D. (Statistics) (Ibadan)
Department of Statistics,
University of Ibadan,
Ibadan.

DEDICATION

To the Glory of Almighty God

The One who Was, who Is and Always.

ACKNOWLEDGMENTS

I give thanks to the Almighty Father for the grace, giving to me to carry out this work, without Him nothing is impossible.

I will like to acknowledge the good gestures, assistances, and encouragements given by men and women of good will who have provided me with crucial help and support while carrying out the research for this dissertation.

Firstly, I must acknowledge my Supervisor Dr. I.O. Shittu who made himself available for consultation, advice and has spent countless months in discussion of this project and provided me with ideas while keeping me on the right track despite his tight schedule. The good atmosphere he provided cemented my hope and confidence.

I will like to acknowledge Dr. Adepoju (Mrs), the Acting Head of the Department, Dr. Osowole, Dr. Yaya, and Mr. Ayoola for their assistance, not only through this work, but also in and out of academic work most especially during my time of health challenge.

I also need to appreciate Dr. Ojo, the post graduate coordinator for his encouragement and contributions.

I will also want to appreciate every academic staff of the department for their contributions.

My appreciation also goes to Messers Adepoju Kazim and Ohikhuemeh Stanley for sharing ideas and technical help during the course of this project.

I highly recognize the help and encouragement I got from the following; Mr. Balogun of Federal of School of Statistics, Mr. Babajide, Dr. Ayannuga, Mrs. Akanbi, Mrs. Talabi and Eng. Adeyemi (Mrs.) of Yaba College of Technology, Mr. Bolarinwa (AJecrehin) of Federal Polytechnic Bida , Mr. Nuga Kehinde and Mr. Olugbode of National Bureau of Statistics..

I will also appreciate Niyi for his painstaking in typing this project.

Finally, I am thankful to my wife and children for their love, patience, endurance, understanding, motivation, and support throughout this passage.

Olufolabo, 'Sesan Oluyomi

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.0 General Introduction

Generally, time series is defined as a collection of observations made sequentially over time or data that are collected at regular interval of time. Although the ordering is usually through time, particularly in terms of some equally spaced time intervals, the ordering may also be taken through other dimensions such as space known as frequency domain. Time series occur in a variety of fields. These observations are stochastic and are known to follow patterns based on time series theory. It is an important aspect of statistics that is well known for its descriptive capability, analysis, identification, and determination of stochastic models for the existing dynamic system as well as its uses in forecasting and monitoring of events. Among the components of series are the trend, seasonal movement, cyclical movement, irregular movement and outliers.

Real data and databases may often include some erroneous parts. These situations, which damage the characteristics of data, are called “abnormal conditions,” and the values, which cause these “abnormal conditions,” are called outliers, Kaya (2010). The outliers, which are really independent, are the situations that cause the parameter estimation values in modelling to be subjective, they damage the processes even though they are set properly, and it is an obligation to destroy or to eliminate the effects.

A commonly used definition of outliers is that they are minority of observations in a datasets that have different patterns from that of the majority of observations in the dataset or are observations, which deviate so much from other observations as to arouse suspicious that they were generated by a different mechanism, Harkins (1980).

Outlier can also be defined as observations that appear to be inconsistent with the remainder of the data set, Betnett and Lewis (1994). Another definition is that outliers are minority of observations in a datasets that have different patterns from that of the majority of observations in the dataset. The assumption here is that there is a core of at least 50% of observations in a dataset that are homogenous (that is, represented by a common pattern) and the remaining observations (hopefully few) have patterns that are inconsistent with this common pattern.

Identification of outlying data points is often by itself the primary goal, without any intention of fitting a statistical model. The outliers themselves are points of primary interest, drawing attention to unknown aspects of data, or especially if unexpected, leading to new discoveries.

On human angle, in the September 11, 2001 attacks on World Trade Centers in New York, United State of America, 5 out of the 80 passengers on one of the flights displayed unusual characteristics. These five passengers (outliers) were not U.S. citizens but had lived in the USA for some periods of time, were citizens of a particular foreign country, had all purchased one-way tickets, had purchased these tickets at the gate with cash rather than credit cards, and did not have any checked luggage. One or two of these characteristics might not be very unusual, but taken together, could be seen as markedly different from the majority of airline passengers. Also unauthorized computer network intrusions could also be seen as outliers, whereby the intruder exhibits a combination of characteristics that jointly considered, are different from typical network users. Perpetrators of credit card fraud provide yet another example where identification of outliers is critical and where the transaction database needs to be analyzed with the specific purpose of identifying unusual transactions. These examples demonstrate the need for outlier identification on every kind of datasets.

The essence of outlier detection is to discover the unusual data, whose behaviour is very exceptional when compared to the rest of the data set. Examining the extraordinary behaviour of outliers helps to uncover the valuable knowledge hidden behind them and to help the decision makers to improve on the quality of data.

Detection methods are divided into two parts: univariate and multivariate methods. In univariate methods, observations are examined individually while in multivariate methods, associations between variables in the same dataset are taken into account.

Different types of outliers such as additive and innovation outliers were studied by Tsay et. al (2000). A graphical method was explored by Khattree and Naik (1987). Grossi (1999) proposed a leave-*k*-out diagnostic procedure while Bayesian analysis was performed by Barnett (1978).

The problem of outlier detection in time series has gained much attention as far back as early 1970s and various methods are available. For this reason, several outlier detection, and robust estimation procedures have been proposed in the literature for time series analysis.

Fox (1972) concluded that the importance of outlier detection in time series lead to:

- Better understanding of the series under study
- Better modelling and estimation

- Improved intervention analysis and
- Better forecasting performance

1.1 Justification for the Study

Several methods of outlier detection in time series are available for univariate time series but very limited ones for multivariate case. Although outliers could be easily identified in univariate through graphical examination of the data, visual inspection does not work for more than one dimension. Examining each dimension by itself or in pairs does not work because it is quite possible for data to be outliers in multivariate space, but not outlying in any of the original univariate dimensions. Thus, graphical inspection is insufficient for outlier detection. In many cases, multivariate observation cannot be detected as outlier when each variable is considered independently, but when multivariable analysis is performed and the interactions among different variables are compared within the class of data, outlier detection is better done. Thus there is need for multivariate outlier detection due to the well-known swamping effects.

1.2 Specific objectives of the study

In addition to the review of the existing techniques for identification and labelling outliers in both univariate and multivariate time series, attention in this study is focused on:

1. Outlier identification and estimation of the magnitude of outliers under classical rules.
2. Models for discovering occurrences of outlier's in multivariate time series data are proposed under Gaussian assumption; the theoretical basis for the new model is developed in order to determine their relative efficiency.
3. Application of the models to both simulated and real life data sets is presented to justify or check the effectiveness of the proposed models.

CHAPTER TWO

REVIEW OF LITERATURE

2.0 Introduction

Before proceeding, we review the literature on outlier detection in both univariate and multivariate time series.

Outlier detection has gained much attention in the 1980s and various methods are available especially in univariate time series. In time series analysis, outliers are known to cause biases in parameter estimation as well as model misspecification, resulting in misleading conclusion. For this reason, several outlier detection and robust estimation procedures have been proposed in the literature for time series analysis.

2.1. Outlier Detection in Univariate Time Series

According to Fox (1972), there are two types of outliers in a time series. The first type of outlier is called the "additive outlier (AO)" which assumes disturbances are committed by addition of an unknown magnitude of outlier to a particular observation. Mathematically, the observed time series is

$$Y_t = Z_t + \omega\xi_t^T \quad , \quad t = 1,2,\ldots\ldots\ldots n \qquad (2.1)$$
$$T = 1,2,\ldots\ldots\ldots n$$

Where Z_t is an outlier-free time series, assumed to follow the autoregressive time series as $Z_t = \frac{\theta(\beta)}{\phi(\beta)} a_t$; while $a_t = \frac{\phi(\beta)}{\theta(\beta)} Y_t$, ω denotes the magnitude of the disturbance and

ξ_t^T is the indicator variable defined by:

$$\xi_t^T = \begin{cases} 1 & if \;\; t = T \\ 0 & if \;\; t \neq T \end{cases} \; and \;\; Y_t \text{ is the outlier contaminated series} \qquad (2.2)$$

In other words, for an AO model

$$Y_t = Z_t \; if \; t \neq T \text{ and } Y_t = Z_t + \omega \qquad (2.3)$$

A typical practical example of an AO is a typographical or a recording error in a time series.

Another type of outlier is called "Innovative Outlier (IO)", a weighted function of the disturbance term innovative outlier may affect every subsequent observation of the series.

Mathematically, an IO model is defined as:

$$Y_t \;=\; \frac{\theta(B)}{\phi(B)}(a_t \;+\; \omega\xi_t^T) \qquad (2.4)$$

Where ξ_t^T is defined as before and ω denotes the magnitude of outlier while a_t is the disturbance term. Rewriting the model as

$$Y_t \;=\; Z_t + \frac{\theta(B)}{\phi(B)}\omega\xi_t^T \qquad (2.5)$$

where $Z_t = \frac{\theta(B)}{\phi(B)}\; a_t$

We see that an IO affects the series through its own dynamic weights $\frac{\theta(B)}{\phi(B)}$ and, in effect, becomes part of the system thereafter. In practice, an IO often indicates an onset of certain changes in the system. For instance, in a manufacturing process, changing an operator or a measurement instrument may result in an IO.

Of course, many other types of disturbances can occur in a time series. The AO and IO models are two of many possibilities.

In Chen et.al (1988) and Tsay (1988) papers, two other types of disturbances were introduced. They are the level shift and temporary change in level. Mathematically, a level shift (LS) which was described by

$$Y_t \;=\; Z_t + \frac{\omega_s}{(1-B)}\xi_t^T \qquad (2.6)$$

Where ω_s is the amount of shift in the level Z_t. Writing

$$\frac{1}{(1\;-\;B)} \;=\; 1\;+\;B\;+\;B^2\;+\;\cdots \qquad (2.7)$$

We see that for the above model

$$Y_t \;=\; \begin{cases} Z_t & for\;\; t < T \\ Z_t + w_s & for\;\; t \geq T \end{cases} \qquad (2.8)$$

Where w_s is the initial shift

Thus, the fixed constants w_s i.e.(the magnitudes of outlier for level shift) is added to every observation one after the other . Such a level shift is permanent.

In some cases, the effect of a level shift is only temporary.

A mathematical model capable of describing such a shift is

$$Y_t = Z_t + \frac{\omega_c}{(1-\delta B)} \xi_t^T, \; 0 < \delta < 1 \qquad (2.9)$$

Since

$$\frac{1}{1-\delta B} = 1 + \delta B + \delta^2 B^2 + \delta^3 B^3 + \cdots \qquad (2.10)$$

the magnitudes of level shift at times d, d + 1, d + 2, … are ω_c, $\delta\omega_s$, ….

where ω_c and ω_s are magnitudes of outliers for temporary change and level shift respectively and d is the rate at which subsequence shifts are discounted.

Thus, the initial shift is ω_s and the subsequent shifts are discounted at the rate δ. With $0 < \delta < 1$, the shift decays exponentially to zero. We refer to such a temporary level shift as a transient change (TC) model.

In practice, outliers can occur at any time point is a series. Thus, to detect an outlier, we need to estimate the parameters $\omega_a, \omega_v, \omega_s, \omega_c$ and check the significance of these estimates.

For simplicity, Chen et.al (1988) assumed that time series parameters are known and can be estimated and an iterative procedure was employed to detect outliers. The four outlier models discussed were then put in the general form as

$$Y_t = Z_t + \omega_0 \frac{\omega(B)}{\delta(B)} I_t^{(d)} \qquad (2.11)$$

Where Y_t is the outlier contaminated series and

$$\omega_0 = \begin{bmatrix} \omega_a & AO\ case \\ \omega_I & IO\ case \\ \omega_s & LS\ case \\ \omega_c & TC\ case \end{bmatrix}; \quad \frac{\omega(B)}{\delta(B)} = \begin{bmatrix} 1 & AO\ case \\ \frac{\theta(B)}{\phi(B)} & IO\ case \\ \frac{1}{(1-B)} & LS\ case \\ \frac{1}{1-\delta B} & TC\ case \end{bmatrix}$$

Given $\theta(B)$ and $\phi(B)$, define as $\left(1-\theta_1\beta-\theta_2\beta^2-\ldots\theta_p\beta^p\right)$ *and* $\left(1-\theta_1\beta-\theta_2\beta^2-\ldots\theta_q\beta^q\right)$ as the polynomial function of the autoregressive $AR_{(p)}$ process and moving average $MA_{(q)}$ respectively.

$$y_t = \frac{\theta(B)}{\phi(B)}\frac{\omega(B)}{\delta(B)}\xi_t^T \tag{2.12}$$

Then,

$$y_t = \omega_0 y x_t + a_t \tag{2.13}$$

which is precisely a simple linear regression equation. Therefore,

$$\hat{\omega}_0 = \frac{\sum_{t=1}^{n} y_t x_t}{\sum_{t=1}^{n} y_t x_t^2} \tag{2.14}$$

and

$$\text{Var}(\hat{\omega}_0) = \frac{\sigma_a^2}{\sum_{t=1}^{n} x_t^2}, \tag{2.15}$$

where n is the sample size. Using simple least squares technique, the following were obtained by Chen et.al (1988)

- IO case: $\hat{\omega}_{v,d} = y_d$ and $\text{Var}(\hat{\omega}_{v,d}) = \sigma_a^2$
- AO case: $\hat{\omega}_{a,d} = \rho_{a,d}^2\left(y_d - \sum_{i=1}^{n-d} \pi_i\, y_{d+i}\right)$ and $\text{Var}(\hat{\omega}_{a,d}) = \rho_{a,d}^2\,\sigma_a^2$ where $\pi's$ are the π-weights of Z_t and $\rho_{a,d}^2 = (1 + \pi_1^2 + \cdots + \pi_{n-d}^2)^{-1}$
- LS case: $\hat{\omega}_{s,d} = \rho_{s,d}^2\left(y_d - \sum_{i=1}^{n-d} \eta_i\, y_{d+i}\right)$ and $\text{Var}(\hat{\omega}_{s,d}) = \rho_{s,d}^2\,\sigma_a^2$ where $\eta_i's$ are the coefficient of B^i in the polynomial $\eta(B) = \eta_0 - \eta_1 B - \eta_2 B^2 - \cdots = \frac{\pi(B)}{1-B}$ and $\rho_{s,d}^2 = (1 + \eta_1^2 + \cdots + \eta_{n-d}^2)^{-1}$
- TC case: $\hat{\omega}_{c,d} = \rho_{c,d}^2\left(y_d - \sum_{i=1}^{n-d} \beta_i\, y_{d+i}\right)$ and Var $(\hat{\omega}_{c,d}) = \rho_{c,d}^2\,\sigma_a^2$ where $\beta_i's$ are the coefficient of B^i in the polynomial $\beta(B) = \beta_0 - \beta_1 B - \cdots = \frac{\pi(B)}{1-\delta B}$ and $\rho_{c,d}^2 = (1 + \beta_1^2 + \cdots + \beta_{n-d}^2)^{-1}$.

Based on the above results, these test statistics were employed

- Existence of an IO at d: $\lambda_{v,d} = \frac{\hat{\omega}_{v,d}}{\sigma_a}$

- Existence of an AO at d: $\lambda_{a,d} = \frac{\widehat{\omega}_{a,d}}{\rho_{a,d}\sigma_a}$

- Existence of an LS at d: $\lambda_{s,d} = \frac{\widehat{\omega}_{s,d}}{\rho_{s,d}\sigma_a}$

- Existence of an TC at d: $\lambda_{c,d} = \frac{\widehat{\omega}_{c,d}}{\rho_{c,d}\sigma_a}$

Under the null hypothesis of normality, no disturbance at *d* and knowing the time series parameters and *d*, all of the above four statistics are distributed as *N (0, 1)*. In practice, the parameters can be replaced by the MLEs. However, since *d* is unknown, there is need to apply the tests to all possible values of *d*. Consequently, in other words, there is need to consider the maximum of test statistics over *d*. The resulting statistics are no longer normal. However, one can obtain certain percentiles via simulation or using distributions of certain extreme-value statistics. Experience suggests that using a critical value of 3.0 or 3.5 works reasonably well in practice.

Iterative procedure for time series analysis in the presence of outliers, level-shifts, and temporary changes was considered by Tsay (1988) and Chang, et al. (1988).

The procedure considered by them is the very basic one as follows:

a. Identify an ARMA model for y_t, estimate the associated parameters and assuming that there are no outliers in y_t.

b. Based on the model of step (a) above, compute the four test statistics for each time point and identify

$$\lambda_{v,max} = \max_d \{|\lambda_{v,d}|\},$$

$$\lambda_{a,max} = \max_d \{|\lambda_{a,d}|\},$$

$$\lambda_{s,max} = \max_d \{|\lambda_{s,d}|\} \text{ and}$$

$$\lambda_{c,max} = \max_d \{|\lambda_{c,d}|\}$$

Where λ_v, λ_a, λ_s, and λ_c are the test statistics for innovative, additive, level shift and temporary change models respectively.

c. Let $\lambda = max\{\lambda_{v,max}, \lambda_{a,max}, , \lambda_{s,max}, \lambda_{c,max}\}$ and compare λ with the pre-specified critical value C. if $\lambda < C$, there is no outlier and stop. If $\lambda \geq C$, continue to the next step.

d. Compute a modified series Y_t^* by removing the effect of the identified outlier and go to step a with y_t replaced by Y_t^* .

Chen and Liu (1993) proposed iterative procedure to reduce masking effects by estimating jointly the model parameters and the magnitudes of outlier effects.

Pankratz (1993) considers AO and IO in a dynamic regression model with a single input and a single output. He classifies outliers in the input series as passed and non-passed outliers and uses a weighted average of least squares estimators to estimate non-passed outliers. The approach however becomes complicated when there are multiple inputs or multiple output series.

As a result of outlier masking effect of both Additive and Innovative on the estimates of parameters and the multiplicative effect on parameters estimated, Shittu (2000) introduced two other types of outliers which are Convolution outlier (CO) and Multiplicative outlier (MO).

The MO and CO were derived as follows:

Assuming that outlier free series Z_t follows the $ARMA_{(p,q)}$, process and can be written as

$$Z_t = \frac{\theta(B)a_t}{\Phi(B)} \qquad (2.16)$$

Where $\theta(B)$and $\phi(B)$ are define as : $(1-\theta_1\beta - \theta_2\beta^2 - ...\theta_p\beta^p)$ *and* $(1-\theta_1\beta - \theta_2\beta^2 - ...\theta_q\beta^q)$ as the polynomial function of the autoregressive ($AR_{(p)}$) process and moving average $MA_{(q)}$ respectively.

Letting $\theta(B) / \theta(B) = \pi(B)^{-1}$

Then multiplicative outlier model was defined as

$$X_t = Z_t \, \omega \xi_t^{(T)} \qquad (2.17)$$

and using the least square theory, the estimate of the magnitude of outlier ω and residual variance were arrived at as

$\varpi_{(M)\,T} = \pi(B)^{-1} e_t$ and

$\pi(B)^{-2}\sigma^2_a$) respectively.

The corresponding test statistic was defined as

$$\lambda_{MT} = \frac{\varpi(M)T}{\pi(B)^{-1}\sigma_a} \tag{2.18}$$

The testing criteria $\lambda_M = \text{Max}_{(T:\ 1 \leq T \leq n)} |\lambda_M|$ was employed for outlier detection using the critical values suggested by Chang et.al (1988).

For non-parametric approach, Ljung,(1993) and Battaglia and Baragona, (2007) have proposed specific procedures based on the relationship between the additive outliers and the linear interpolator, while Baragona et.al., (2001) used a genetic algorithm.

Shittu and Shangodoyin (2008) considered the identification and detection of outliers in frequency domain using the spectral method. By assuming both the additive and multiplicative effect of outliers on a series, the parameters of the model were estimated using the maximum likelihood method with a view to measuring the effect of the suspected outlier on the parameter of the series. They concluded that the occurrence of outliers has led to a shift in the phase, amplitude of the Fourier series thus affected the periodogram estimates, and detection of aberrant observations is more exact in the frequency domain than in the time, domain.

2.2 Reviews on Outlier Detection in Multivariate Time Series

As earlier noted, not much work has been done on outlier detection in multivariate time series. Among the existing ones, was the projection pursuit techniques used by Tsay et.al (2004) in order to find the linear combination of a multivariate time series that maximizes kurtosis with the purpose of best reproducing the outlying signal. Then, detection of time points and estimating the magnitudes of multiple outliers were accomplished by employing univariate searching methods.

Baragona and Battaglia (2007) proposed the Independent Component Analysis (ICA) as a tool capable of identifying the locations of multiple outliers in multivariate time series. It was believed that outlying components have a very large kurtosis. The ICA was therefore used at identifying a set of independent unobservable variables that are supposed to generate the data set of interest. An unknown mixing matrix was postulated to linearly transform the unobservable variables to produce a set of observable mixed ones. Both unobservable variables and the mixing matrix have to be estimated from the data. The ICA has been applied successfully to a variety of fields such as biomedicine, speech, and radar, signal processing and time series. Suppose that we observe a contaminated multivariate time series obtained by linearly mixing some independent Gaussian signals, and adding, only at some fixed time points, a constant to each observed component. When the series is decomposed by ICA, the most important non-Gaussian components is likely to represent the outlying pattern, while the remaining independent components would be essentially similar to Gaussian linear combinations of the outlier free time series.

In their own work, Cucina, et.al (2008) used meta-heuristic methods to detect additive outliers in multivariate time series. The implemented algorithms were; simulated annealing, threshold accepting and two different versions of genetic algorithm. They used the same objective function, the generalized AIC-like criterion, and in contrast with many of the existing methods, they do not require specifying a vector auto regressive moving averages model for the data and are able to detect any number of potential outliers simultaneously. They concluded that almost all available methods for outlier detection are iterative, but there is a crucial difference with respect to the meta-heuristic algorithms in that it seems to be able to provide more flexibility and adaptation to the outlier detection problem.

In the detection of outliers in multivariate time series model, Helbling and Cleroux, (2009), introduced the coefficient of vector autocorrelation, obtained its influence function together with its distribution, and used it for testing the hypothesis of presence of outliers.

CHAPTER THREE

METHODOLOGY AND THEORETICAL FRAME WORK

3.0 Introduction

This chapter is divided into two sections. The first section deals with a brief review of some basic tools, concepts, and methodology in time series analyses applicable in this research work.

The second section deals with the theoretical framework where, four model-generating mechanisms for the detection of outliers in the multivariate time series are developed by specifying two-variable Vector Auto-Regressive (VAR) models and comparing their relative sensitivity to outlier.

For the four models, estimates of the magnitude of outlier as well as their residual variances are obtained using the method of least squares. The test statistic for each model for testing the existence of outlier will then be constructed.

3.1 Basic Concept of Time Series

3.1.1 Time Series

A time series is a set of observations measured sequentially over time. These measurements may be made continuously through time or be taken at a discrete time points. By convention, these two types of series are called continuous and discrete time series, respectively, according to the nature of time. In other words, for discrete time series, for example, it is the time of occurrence that is discrete. For a continuous time series, the observed variable is typically a continuous variable recorded continuously on time, such as a measure of brain activity recorded from an electronic machine. The usual method of analysing such a series is to sample (or digitize) the series at equal time interval to give a discrete time series. Little or no information is lost by this process provided that the sampling interval is small enough, Chatfield (1980).

3.1.2 Univariate and Multivariate Time Series Models

A univariate time series model for a given variable is based only on past values of that variable, while a multivariate model for a given variable may be based, not only on past values of that variable, but also on present and past values of other (predictor) variables. In the latter case, the variation in one series may help to explain the variation in another series.

3.1.3 Time Plot

This is the plot of the series against the corresponding time period. It is the first step in analyzing a time series. It helps to show up the important features such as trend, seasonality, discontinuity, outliers, and smooth changes in structure, turning points, and sudden discontinuities. The plot of time series graph is vital, both in describing the data, helping to formulate a sensible model and in choosing an appropriate forecasting method. Chatfield (1980).

3.2 Analysis of a Time Series

There are two major types of analysis of time series data.

3.2.1 Deterministic Modelling

An observed realization of a series is believed to be made up of four major components regarded as variations. These are the secular trend, seasonal, cyclical and irregular variations.

The secular trend is the general direction of the movement of a series with time. It indicates the direction of assessment of the behaviour of the series upon studying the time-plot of the series. In order to isolate the contributing effect of the secular trend the moving average analysis of the series is performed or the series is regressed on time variable (t).

The seasonal variation component gives the structure of the series with respect to equal-spaced, defined periods with relatively small time intervals such as daily, monthly, quarterly, annually repetitive patterns. This variation is isolated by the determination of seasonal index corresponding to each defined period, which is used to adjust the original data.

The cyclical variation represents the long term repetitive cycle that may be inherent in a series. It should be long enough to exhibit such cycle; while the irregular variation is the residual, unpredictable, non-structured components of the series. It usually results from unexpected shock, mishaps such as wars, natural disasters or extreme favourable condition such as in the income series of an individual who win a lottery jackpot. Since the irregular variation is expected to be less prominent in determining the true structure of the series it is often merged with the cyclical component which is equally less influential, especially in short series. They both form the remnant after the secular trend and the seasonal variation components must have been eliminated from the series.

The process of decomposition requires that either an additive or multiplicative model be used in the representation of the series depending on, whether the time plot of the series reveals a series with increasing or constant bandwidth.

3.2.2 Dynamic Modeling

Modeling in this sense is about constructing mathematical relation by exploring statistical properties of two or more series. The model thus generates the underlying series or establishes relationship between series. A time series model can be constructed in time or frequency domain.

In time domain modeling, time is the reference parameter. The model that may be as a result of relating present observation with past ones or from evaluating random error terms. The time plot shows the structure of the series in this form of modeling, while the plot of a correlogram is based on values of Auto Correlation Function (ACF) or Partial Autocorrelation Function (PACF) against respective time lag. The correlogram is an important tool employed in the identification and estimation of a model.

On the other hand, it is often postulated that a series is made up of more than one sinusoidal wave curves, which is typical of long series. In the process of splitting such series into its different component waveforms, the series is modeled in the frequency domain. In this respect evaluating periodic functions and relevant statistical properties are required to achieve this goal. The plot of either a Periodogram or Spectral Density Function is synonymous to what plotting the ACF or PACF correlogram achieves in the time domain.

3.3 Fundamental Properties of Dynamic Models

A series is completely described by its expected value, variance, auto-covariance function, autocorrelation function and partial autocorrelation function; irrespective of the model constructed for such series. These measures form the basis upon which some fundamental properties peculiar to model building are considered.

3.3.1 Stationarity

Stationarity is a condition required to infuse certain level of control for congenital variability within a series. Stationarity is a measure of state of equilibrium of a process about a constant mean level. When a series is not stationary, it is difficult to exploit its analysis for predicting future values of the series since such forecast tends to explode with increase in time value. This inhibits control actions required to be taken if the need arises. In addition, a stationary series enables a parsimonious reduction of parameters required to be estimated in a model.

3.3.2 Weak Stationarity

A time series is said to be stationary if its underlying generating process is based on a constant mean and constant variance with its autocorrelation function essentially constant through time. Thus, if different subsets of a realization are considered (time series 'sample') the different subsets will typically have means, variances, and autocorrelation function that do not differ significantly.

A statistical test for stationarity is the most widely used Dickey Fuller test. To carry out the test, estimate by Ordinary Least Squares (OLS) regression model is constructed:

$$y_t' = \phi y_{t-1} + b_1 y_{t-2}' + \cdots + b_p y_{t-p}' \tag{3.1}$$

where y_t' denotes the differenced series ($y_t - y_{t-1}$). The number of terms in the regression, p, is usually set to be about 3. Then if ϕ is nearly zero, then the original series y_t needs differencing and if $\phi < 0$ then y_t is already stationary.

3.3.3 Autocorrelation Functions

Autocorrelation refers to the way the observations in a time series are related to each other and is measured by the simple correlation between current observation (Y_t) and observation from p periods before the current one (Y_{t-k}). That is for a given series Y_t, autocorrelation at lag p, correlation (Y_t, Y_{t-k}) and is given by

$$\rho_k = \frac{Cov(Y_t,\ Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} \tag{3.2}$$

The value of ρ_k ranges from –1 to +1. Box and Jenkins has suggested that maximum number of useful r_p are roughly n/4 where n is the number of periods upon which information on Y_t is available. The plot of ACF on time t is called the correlogram.

3.3.4 Partial Autocorrelation

Partial autocorrelations are used to measure the degree of association between Y_t and Y_{t-p} when the Y-effects at other time lags 1, 2, 3, ..,, p–l are removed.

3.3.5 Model Identification Process

Theoretical ACFs and PACFs (Autocorrelations versus lags) are used to determine the appropriate model for a time series. Thus one can compare the correlogram (plot of sample ACFs versus lags) with these theoretical ACFs/PACFs, to find a reasonable good match and tentatively select one or more ARIMA models.

3.4 Types of Dynamic Models

3.4.1 Autoregressive (AR) Model

Let X_t be any time series and X_t is said to follow an autoregressive process of order *p*, AR(*p*), if it satisfies the equation:

$$X_t = \alpha_0 + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \ldots + \theta_k X_{t-k} + \varepsilon_t \quad (3.3)$$

where ε_t is a sequence of independent and identically distributed Gaussian variables with mean zero and variance σ^2.

An AR process is said to be stationary if the root of the polynomial equations

$(1-\theta(\beta))=0$ lies outside the unit root circle otherwise, the series is said to be non-stationary.

An AR of order ρ i.e. AR(ρ) has the following properties:

$$E(X_t) = \frac{\alpha_0}{1-\theta_1-\theta_2-\ldots\theta_\rho} \quad (3.4)$$

$$Var(X_t) = \gamma_0 = \frac{\alpha^2_{t\varepsilon}}{1-\rho_1\theta_1-\rho_2\theta_2-\ldots\rho_\rho\theta_\rho} \quad (3.5)$$

where $\sigma^2_{\varepsilon_t}$ is the variance of the error terms.

A plot of its correlogram (ACF/PACF vs Time lag) shows an exponentially decaying ACF function and the PACF cut-off after lag k thus identifying an AR model of order k.

3.4.2 Moving Average (MA) Model

This model expresses the current observation X_t for $t \in T$ as the combination of a constant term and current error term with the linear combination of past random error terms up to a lagged period q i.e.

$$X_t = \varepsilon_t + \alpha_0 + \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \ldots + \phi_q\varepsilon_{t-q} \qquad (3.6)$$

Characteristically, a moving average process is always stationary, as it largely depends on q past random errors which are independent and identically distributed depending on the order of the model; however invertibility condition is required to be satisfied in the process of estimating the parameters of the model.

The random error terms are white noise having mean zero and variance σ_ε^2, hence

$$E(X_t) = \mu = \alpha_0 \qquad (3.7)$$

$$\mathrm{Var}(X_t) = \gamma_0 = \sigma_\varepsilon^2\left(1+\phi_1^2+\phi_2^2+\ldots+\phi_q^2\right) \qquad (3.8)$$

and $$\gamma_k = \begin{cases} \sigma_\varepsilon^2 & if \quad k = 1, 2\ldots, j \\ 0 & k > j \end{cases} \qquad (3.9)$$

To identify a MA(q) process, the ACF cut-off after lag q i.e. and the PACF progressively dampens-out with increase in lag period.

3.4.3 Autoregressive Moving Average (ARMA) Model

This model as the name suggest combines the feature of both $AR_{(\rho)}$ and $MA_{(q)}$ models. An ARMA(ρ,q) is represented as:

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \ldots\ldots\ldots + \theta_p X_{t-p} - \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \ldots + \phi q\varepsilon_{t-p} \qquad (3.10)$$

To satisfy the stationarity and the invertibility conditions, the root of polynomial equations θ(B) = 0 and ϕ(B) = 0 must lie outside the unit root circle. In the identification process of an $ARMA_{(p, q)}$ model, it is required that the ACF and PACF tails off as the lag periods k and j increases. The $ARMA_{(p,q)}$ process have the advantage of modelling a time series with fewer parameters estimated as either a pure $AR_{(p)}$ or $MA_{(q)}$ would have done.

Table 1.1: Summary of Model Identification Process

Model	ACF	PACF
AR(p)	Spikes decay towards zero	Spikes cutoff to zero
MA(q)	Spikes cutoff to zero	Spikes decay to zero
ARMA(pq)	Spikes decay to zero	Spikes decay to zero

Pankratz (1983)

3.4.4. Autoregressive Moving Integrated Average (ARIMA) Model

In general, an ARIMA model is characterized by the notation ARIMA (p, d, q) where, p, d and q denote orders of auto-regression, integration (differencing) and moving average respectively. In ARIMA, time series is a linear function of past actual values and random shocks. For instance, given a time series process $\{Y_t\}$, a first order auto-regressive process is denoted by ARIMA (1, 0, 0) or simply AR (1) and is given by

$$Y_t = \mu + \phi_1 Y_{t-1} + \epsilon_t \qquad (3.11)$$

and a first order moving average process is denoted by ARIMA (0, 0, 1) or simply MA (1) and is given by

$$Y_t = \mu - \theta_1 \epsilon_{t-1} + \epsilon_t \qquad (3.12)$$

Alternatively, the model ultimately derived, may be a mixture of these processes and of higher orders as well. Thus a stationary ARIMA (p, q) process is defined by the equation

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots \theta_q \epsilon_{t-q} + \epsilon_t \qquad (3.13)$$

where $\boldsymbol{\varepsilon}_t$'s are independently and normally distributed with zero mean and constant variance σ^2 for t = i, 2, ...n. The values of *p* and *q*, in practice lie between 0 and 3.

3.5 Estimation

At the identification stage, one or more models are tentatively chosen that seem to provide statistically adequate representations of the available data. Then precise estimates of parameters of the model are obtained by least squares as advocated by Box and Jenkins (1994). Standard computer packages like SAS, SPSS, R-programming etc. are available for finding the estimates of relevant parameters using iterative procedures.

3.5.1 Diagnostics

Different models can be obtained for various combinations of autoregressive and moving averages individually and collectively. The best model is obtained with following diagnostics:

3.5.2 Model Identification Criteria

For the model identification criteria Pankratz (1983) gave three alternatives; Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Schwarz-Bayesian Information Criteria (SBC).

AIC is given by AIC = (–2 log L + 2 m) (3.14)

where m = p+ q+ P+ Q and L is the likelihood function. Since –2 log L is approximately equal to {n (l+log 2π) + n log σ^2} where σ^2 is the model mean square error.

AIC can be written as: MC={n (t+log 2 π) + n log σ^2 + 2m} and because first term in this equation is a constant, it is usually omitted while comparing between models.

As an alternative to AIC, sometimes SBC is also used which is given by:

SBC= log σ^2 + (m log n)/n.

After tentative model has been fitted to the data, it is important to perform diagnostic checks to test the adequacy of the model and, if need be, to suggest potential improvements. One way to accomplish this is through the analysis residuals. It has been found that it is effective to measure the overall adequacy of the chosen model by examining a quantity Q known as Box-Pierce statistic (a function of autocorrelations of residuals) whose approximate distribution is chi-square and is computed as follows:

$$Q = n \sum r^2(j) \qquad \mathrm{Q = n} \sum_{i=1}^{k} r^2(j) \qquad (3.15)$$

where summation extends from 1 to k with k as the maximum lag considered, n is the number of observations in the series, r (j) is the estimated autocorrelation at lag j; k can be any positive integer and is usually around 20. Q follows Chi-square with (k-m) degrees of freedom where m is the number of parameters estimated in the model. A modified Q statistic is the Ljung-box statistic which is given by

$$Q = n(n+2)\sum_{i=1}^{k} r^2(j)\,/(n-1) \qquad (3.16)$$

The Q Statistic is compared to critical values from chi-square distribution. If model is correctly specified, residuals should be uncorrelated and Q should be small (the probability value should be large). A significant value indicates that the chosen model does not fit well.

3.6 Theoretical Framework

This work is premised on detection of outliers in the multivariate time series setting using the Vector Autoregressive (VAR) modeling approach.

3.6.1 Vector Autoregression

Vector Autoregression (VAR) is a statistical model used to capture the linear interdependencies among multiple time series. VAR models generalize the univariate autoregression (AR) models. All the variables in a VAR are treated symmetrically; each variable has an equation explaining its evolution based on its own lags and the lags of all the other variables in the model. VAR modelling does not require expert knowledge, which previously had been used in structural models with simultaneous equations.

VAR models were advocated by Sims (1980), who criticized the claims and performance of earlier modelling in macroeconomic econometrics. Sims recommended VAR models, which had previously appeared in time series statistics and system identification a method to estimate economics relationships, thus being an alternative to the "incredible identification restrictions" in structural models.

A VAR model describes the evolution of a set of k variables (called *endogenous variables)* over the same sample period (t = *1, …, T) as a* linear function of only their past evolution. The variables are collected in a $k \times 1$ vector y_t .

A (reduced) p-th order VAR, denoted *VAR (p)*, is

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \ldots\ldots + A_p y_{t-p} + e_t \qquad (3.17)$$

Where c is a $k \times 1$ vector of constants (intercept), A_i is a $k \times k$ matrix (for every $I = 1, \ldots, p$) and e_t is a $k \times 1$ vector of error terms satisfying the following:

1. $\mathrm{E}(e_t) = 0$ every error term has mean zero;
2. $\mathrm{E}(e_t e'_t) = \Omega$ the contemporaneous covariance matrix of error terms is Ω ($\mathrm{k} \times \mathrm{k}$ positive definite matrix);

3. $E(e_t e'_{t-k}) = 0$ for any non-zero k, there is no correlation across time; in particular, no serial correlation in individual error terms.

The l-periods back observation y_{t-1} is called the l-th lag of y, thus, a pth-order VAR is also called a VAR with p lags.

.

In matrix notation, one can write a VAR(p) with a concise matrix notation as:

$$Y = BZ + U \tag{3.18}$$

For a general example of a VAR (p) with k variables,

A VAR (1) in two variables can be written in matrix form (more compact notation) as

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} A_{1,\mathbf{1}} & A_{1,\mathbf{2}} \\ A_{2,\mathbf{1}} & A_{2,\mathbf{2}} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}, \tag{3.19}$$

or, equivalently, as the following system of two equations

$$y_{1,t} = C_1 + A_{1,\mathbf{1}} y_{1,t-1} + A_{1,\mathbf{2}} y_{2,t-1} + e_{1,t} \tag{3.20}$$

$$y_{2,t} = C_2 + A_{2,\mathbf{1}} y_{1,t-1} + A_{2,\mathbf{2}} y_{2,t-1} + e_{2,t} \tag{3.21}$$

Note that there is one equation for each variable in the model. Also note that the current (time *t*) observation of each variable depends on its own lags as well as on the lags of each other variable in the VAR.

Now writing VAR (p) as VAR (1), a VAR with p lags can always be equivalently rewritten as a VAR with only one lag by appropriately redefining the dependent variable. The transformation amounts to merely stacking the lags of the VAR (p) variable in the new VAR (1) dependent variable and appending identities to complete the number of equations.

For example, the VAR (2) model

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + e_t \tag{3.22}$$

Can be recanted as the VAR (1) model

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} A_1 & A_{\mathbf{2}} \\ I & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} e_t \\ 0 \end{bmatrix}, \tag{3.23}$$

Where I is the identity matrix.

The equivalent VAR (1) form is more convenient for analytical derivations and allows more compact statements.

3.6.2 Linear Models

Linear statistical model refers to the fact that the fraction y_i is linear in the unknown parameters, Graybill et.al (1974).

A general linear function is of the form.

$$Y = f(x) \tag{3.24}$$

or equivalently

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + a_n X_n \tag{3.25}$$

where

$\alpha_i, i = 1, 2, \ldots m$ are fixed arbitrary vectors, and

$X_i, i = 1, 2, \ldots m$ are scalars.

A necessary and sufficient condition that (3.25) has a non-trivial solution, that is not all X_i are simultaneously zero, is that $\alpha_1, \alpha_2, \ldots, \alpha_m$ are dependent, Rao (1965).

A special and important case of linear equations is where

$\alpha_i \in \ E_n \mathbb{R}^m$.

If $\alpha_i = (a_{ii}, \ldots a_{im})$, then (3.25) may be written as:

$$\begin{aligned} &Y_{11} + a_{11}X_1 + a_{12}X_2 + \cdots + a_{1m}X_m = 0 \\ &Y_{21} + a_{21}X_1 + a_{22}X_2 + \cdots + a_{2m}X_m = 0 \\ &\vdots \qquad \vdots \qquad \vdots \qquad \qquad \vdots \quad \vdots \quad \vdots \\ &Y_{n1} + a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nm}X_m = 0 \end{aligned} \tag{3.26}$$

are n linear equations in m unknowns.

If we consider uncorrelated observations $(Y_1, Y_2, \ldots Y_n)$, such that

$$Y_i = \beta_0 + \beta_1 X_{11} + \cdots + \beta_m X_{ni} \tag{3.27}$$

where i = 1, 2, 3,…, n

If we assume that

$$E(Y_i) = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_m X_{ni} \tag{3.28}$$

and $V(Y_i) = \sigma^2$

where $(\beta_0, \beta_1 \ldots \ldots \beta_m)$ and σ^2 are unknown parameters and X_{ij} are unknown coefficients.

We can define random errors $\varepsilon_1, \varepsilon_{2,} \ldots \ \ldots \ldots \varepsilon_n$ by

$$\varepsilon_1, = Y_i - \beta_0 - \beta_i X_i \tag{3.29}$$

for i = 1, 2, …m

and the ε_i satisfy

$$E(\varepsilon_i) = 0 \tag{3.30}$$

and

$$V(\varepsilon_i) = \sigma^2 \tag{3.31}$$

then we can write

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_m X_{i_n} + \varepsilon_i \tag{3.32}$$

for i = 1, 2, …, n

3.6.3 Least Squares Method

The theory of least squares is concerned with the estimation of parameters in linear model, Rao (1968) and improvement on it was have made by Rao (1973), (1974).

Equation (3.32) is a complete mathematical model of a multiple regression equation which in matrix form can be written as

$$Y = X\beta + \varepsilon \tag{3.33}$$

where

$Y = (Y_1, Y_2, \ldots, Y_n)$ is the observation vector, β is a m x 1 vector of parameters and X is a matrix (n x m) of explanatory variables (assumed to be of full rank) andε is an (nx1) vector of residuals.

We assume that ε has zero mean vector and variance matrix $V(\varepsilon) = \sigma^2 I_n$ where I is the nxn identity matrix so that the true residuals have common variance and are uncorrelated.

In the absence of any contaminants, requiring modification of the model in (3.33) and possibly revealed as outliers we have the familiar least squares analysis of the linear model.

The least squares estimation of β is

$$\hat{\beta} = (X`X)^{-1} X`Y \tag{3.34}$$

with

$$V(\hat{\beta}) = \mathrm{V}\left[\left(X`X\right)^{-1} X`Y\right] \qquad (3.35)$$

and that of ε is

$$\varepsilon = (h - X\beta) = (I_n - R)Y$$

$$= (I_n - R)\varepsilon \qquad (3.36)$$

where

$$R = X\left(X`X\right)^{-1} X`X \qquad (3.37)$$

and with

$$V(\varepsilon) = (I_n - R)\sigma^2$$

The last term in (3.36) shows how the estimated residual ε relates to the unknown true residual ε but the determination of ϵ must be sought in terms of the known quantities such as $(I_n - R)Y$. The estimated residuals ε have zero means. From (3.36), we see that they are typically correlated and have differing variances.

Explicitly, we can write

$$Var\,(\varepsilon_j) = [1 - X_j\,(X`X)^{-1} X_j]$$

$$= (1 - r_{jj})\sigma^2 \qquad (3.37)$$

Where $X_j`$ is the j[th] row of X.

The error variance is unknown. An unbiased estimate is obtained as

$$\sigma^2 = \frac{\varepsilon'\varepsilon}{n-q}.$$

$$= \frac{\varepsilon^1 \varepsilon\left(I_n - R\right)}{n-q} \qquad (3.38)$$

in view of the idempotency of $(I_n - R)$; ε is termed the residual sum of squares and is denoted by $V(\hat{\in})$ can now be estimated as

$$S^2(\varepsilon) = (I_n - R)\sigma^2 \qquad (3.39)$$

So that the estimated variance of ε_j is

$$S^2 = (1 - r_{jj})\sigma^2$$

$$= \frac{(1 - r_{jj})\varepsilon^1 \varepsilon}{n - q}$$

$$= \frac{\frac{(1 - r_{jj})\varepsilon^1}{s^2}}{n - q} \tag{3.40}$$

$$\varepsilon_j = \frac{\varepsilon_j}{S_j}$$

$$\varepsilon_j = \frac{\varepsilon_j}{S_j} \quad \frac{\sqrt{n-1}}{1-r_{jj}} \tag{3.41}$$

They have an immediate intuitive appeal in that they constitute weighted version of the estimated residual ε_j, where the weights are inversely proportional to the estimates of the standard deviation of the ε_j.

3.6.4 Gaussian Distribution

Gaussian distribution otherwise known as the normal distribution has proved to be the most useful of all distributions for continuous random variables.

The normal distribution function is given by

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}}\, exp^{-\frac{1}{2\sigma^2}(X-\mu)^2} \qquad -\infty < x < \infty \qquad (3.42)$$

$$-\infty < \mu < \infty$$

$$\sigma^2 > 0$$

This shows that a normal distribution is completely determined by specifying its mean μ and standard deviation σ, also the graph of a typical normal curve is symmetrical about the mean μ.

The maximum likelihood estimate of the parameters of a normal distribution can be obtained by the likelihood function.

$$L\,(\theta) = \prod_{i=1}^{n} f(X_i, \theta) \qquad (3.43)$$

where $X_i,\ i = 1, 2, \dots, n$ are random samples from a population X with the probability density function $F(X, \theta)$ and θ is an unknown parameter.

The probability function (3.43) can be expressed as

$$f(X, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\, \ell^{-\frac{1}{2\sigma^2}(X-\mu)^2} \qquad (3.44)$$

The likelihood function is

$$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp[-\frac{1}{2\sigma^2}(X-\mu)^2] \qquad (3.45)$$

then the log likelihood function

$$\text{In } L = -nln\,(\pi\sigma^2) - \sum_{i=1}^{n} \frac{1}{2\sigma^2}(X-\mu)^2 \qquad (3.46)$$

The maximum likelihood estimate of the population mean μ can be obtained by differentiating equation (3.46) with respect to μ and equate to zero.

$$\frac{\partial L(\mu, \sigma^2)}{\partial \mu} = -\sum_{i=1}^{n} \frac{1}{2\sigma^2}(x_i - \mu)^2 = 0 \qquad (3.47)$$

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (3.48)$$

The maximum likelihood estimate for the variance is obtained by differentiating (3.46) with respect to σ^2 and equating the result to zero.

We have,

$$\frac{\partial L(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{1}{2\sigma^2} - \sum \frac{1}{2\sigma^4}(X - \mu)^2 = 0 \qquad (3.49)$$

$$\frac{n}{2\sigma^2} = \sum \frac{(X_i - \mu)^2}{2\sigma^4} \qquad (3.50)$$

$$\text{n}\sigma^2 = \sum (X_i - \bar{X})^2 \qquad (3.51)$$

That is $X \sim N\ (\mu, \sigma^2)$

3.6.5 Test for Normality

To ensure strict stationarity of a time series data, we need to show that the series come from a normal distribution with constant mean and variance. This can easily be achieved by carrying out a normality test on the collected data using the χ^2 (chi-square) goodness of fit test as follows:

(i) Obtain the maximum likelihood estimate of the population parameters as described in (3.48) and (3.51).

(ii) Classify the data into intervals of equal sizes and obtain the observed frequencies $0_i, i = 1, 2, , \dots k$, where k is the number of classes.

(iii) Standardize the class intervals and obtain the probability for each of the intervals.

(iv) Obtain the expected frequencies $e_i, i = 1, 2, \dots k$ by multiplying the observed frequencies by the probability obtained for each of the standardized intervals.

(v) Carry out the χ^2 goodness of fit test using

$$\chi^2 = \sum_{i=1}^{k} \frac{(0_i - e_i)^2}{e_i} \qquad (3.52)$$

When compared with the critical value of χ^2 at specified level of significance.

Therefore, results obtained for multiple regression could be extended to time series model fitting. The method of least squares discussed above is used to develop the theoretical basis for our proposed outlier generating mechanisms in the section.

3.7 Derivation of the Outlier generating Mechanisms

In this section the four model-generating mechanisms i.e. Innovative Outlier Model (IO), Additive Outlier Model (AO), Multiplicative Outlier Model (MO) and Convolution Outlier Model (CO) for the detection of outlier in Multivariate time series are derived by specifying two-variable Vector Autoregressive (VAR) models.

3.7.1 Innovative Outlier Model

An Innovative Outlier (IO) represents an unexpected change in the innovations that drive the vector time series. For instance, suppose that the noise in a bivariate series consisting of oven temperature and a chemical concentration reading is mainly due to the random variability of the, feed rate. Then, a sudden change in the feed rate that happens at just a particular time point, due to some exogenous effect, will produce an IO in the series.

The innovative outlier-generating model is defined as:

$$X_t = Z_t + \varphi(\beta)\omega\xi_t^{(T)} \tag{3.53}$$

with the unobservable outlier free series by $Z_t = \frac{\theta(\beta)}{\phi(\beta)} a_t$

(3.54)

and $$\ell_t = a_t + \varphi(\beta)\omega\xi_t^{(T)}. \tag{3.55}$$

where X_{ti} = (x_{1t}, …,x_{kt}) is a k-dimensional time series, Z_t is an outlier free time series that is assumed to follow the ARMA$_{(pq)}$, $\xi_t^{(T)}$ is a time indictor such that $\xi_t^{(T)} = 1$ for all $t = T$ *and* $\xi_t^{(T)} = 0$ otherwise, $\varphi(B)$ = 1- Θ_1B- Θ_2 B^2 … – Θ_p B^p are polynomials of order p and $\omega = (\omega_1, \ldots, \omega_k)'$ is the size of the magnitude of outliers.

For the general case of IO:

Given a vector model X_{1t} and X_{2t} such that X_{1t} contains outlier and X_{2t} is outlier free, the magnitude of such outlier and its corresponding variance can be obtained by specifying the two variable VAR$_{(2)}$ as:

$$X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t} \quad (3.56)$$

$$X_{2t} = \phi_{21} X_{2t-1} + \phi_{22} X_{1t-1} + a_{2t} \quad (3.57)$$

where X_{1t} is the current value of the response variable

X_{1t-1} is the lag value of the current variable

X_{2t} is the current value of the explanatory variable

X_{2t-1} is the lag value of the explanatory variable

Now considering

$$X_{2t} = \emptyset_{21} X_{2t-1} + \emptyset_{22} X_{1t-1} + a_{1t} \quad (3.58)$$

When X_{2t-1} is contaminated and assumed innovative model, we then have

$$X_{2t} = \emptyset_{21}\left(Z_{t-1} + \varphi(\beta)\omega\xi_t^{(T)}\right) + \emptyset_{22} X_{1t-1} + a_{1t} \quad (3.59)$$

$$X_{2t} = \emptyset_{21}\, Z_{t-1} + \emptyset_{21}\, \varphi(\beta)\omega\xi_t^{(T)}\ \emptyset_{22} X_{1t-1} + a_{1t}$$

$$X_{2t} = \emptyset_{21}\, \varphi(\beta)\, \epsilon_{t-1} + \emptyset_{21}\varphi(\beta)\omega\xi_t^{(T)} + \emptyset_{22} X_{1t-1}$$

$$X_{2t} = \emptyset_{22}\, X_{1t-1} + \emptyset_{21}\varphi(\beta)\,(1+\omega) \quad (3.60)$$

Therefore in general the IO generating mechanism is:

$$Inovative\ model:\ X_{I\,i\,t} = \phi_{i\,i}\, X_{jt-1} + \emptyset_{ij}\, \xi_t^T \varphi(\beta)\,(1+\omega)$$

3.7.1.1 Derivation of the Magnitude of Outlier for IO

Assuming X_{1t} contains an outlier,

Then

$$X_{1t} = Z_{1t} + \varphi(\beta)\omega\xi_t^{(T)} \quad (3.61)$$

$$Z_{1t} + \varphi(\beta)\omega\xi_t^{(T)} = \phi_{11}\left(Z_{1t-1} + \varphi(\beta)\omega\xi_t^{(T)}\right) + \phi_{12} X_{2t-1} + a_{1t} \quad (3.62)$$

According to Tssay (1988) $\ell_t = \pi(\beta)X_t$ *and* $Z_t = \dfrac{\theta(\beta)}{\phi[\beta]}a_t$

As defined in equations (3.54)

$$X_t = \frac{\ell_t}{\pi(\beta)} = \varphi(\beta)\ell_t \tag{3.63}$$

which becomes

$$\frac{\theta(\beta)}{\phi(\beta)}a_t + \varphi(\beta)\omega\xi_t^{(T)} = \phi_{11}\left(\frac{\theta(\beta)}{\phi(\beta)}a_t + \varphi(\beta)\omega\xi_t^{(T)}\right) + \phi_{12}\varphi(\beta)\ell_{t-1} + a_{1t} \tag{3.64}$$

$$\frac{\theta(\beta)}{\phi(\beta)}at - \phi_{11}\frac{\theta(\beta)}{\phi(\beta)}a_t - a_{1t} = \phi_{11}\varphi(\beta)\omega\xi_t^{(T)} + \phi_{12}\varphi(\beta)\ell_{t-1} - \varphi(\beta)\omega\xi_t^{(T)} \tag{3.65}$$

$$a_t\left[\frac{\theta(\beta)}{\phi(\beta)} - \phi_{11}\frac{\theta(\beta)}{\phi(\beta)} - 1\right] = \phi_{11}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{(T)} + \phi_{12}\varphi(\beta)\ell_{t-1} \tag{3.66}$$

Making a_t the subject of the formulae we have

$$a_t = \frac{\phi_{11}\varphi(\beta)\omega_t^{(T)} - \varphi(\beta)\omega_t^{(T)} + \phi_{12}\varphi(\beta)\ell_{t-1}}{\dfrac{\theta(\beta)}{\phi(\beta)}(1-\phi_{11}) - 1} \tag{3.67}$$

Using the least squares theory to differentiate the sum of squares and equating it to zero, thus we have

$$\sum_{t=1}^{n} a_t^2 = \sum_{t=1}^{n}\left[\frac{\phi_{11}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{(T)} + \phi_{12}\varphi(\beta)\ell_{t-1}}{\dfrac{\theta(\beta)}{\phi(\beta)}(1-\phi_{11}) - 1}\right]^2 \tag{3.68}$$

$$\frac{\partial \sum_{a_t^2}}{\partial \omega} = \frac{\left[2\phi_{11}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{(T)}\right]}{\left[\frac{\theta(\beta)}{\phi(\beta)}(1-\phi_{11})-1\right]^2}\sum_{t=1}^{n}\left[\phi_{11}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{T} + \phi_{12}\varphi(\beta)\ell_{t-1}\right] \tag{3.69}$$

which becomes

$$\sum_{t=1}^{n}\left[\phi_{11}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{T} + \phi_{12}\varphi(\beta)\ell_{t-1}\right] = t$$

(3.70)

Since $\xi_t^{(T)}$ is a time indictor where $\xi_t^{(T)} = 1$ for all $t = T$ *and* $\xi_t^{(T)} = 0$ otherwise, we have

$$\sum_{t=1}^{n}\omega(\varphi(\beta)-\varphi(\beta)) + \phi_{12}\varphi(\beta)\sum_{t=1}^{n}\ell_{t-1} = 0 \tag{3.71}$$

$$\sum_{t=1}^{n}\omega(\varphi(\beta)-\phi_{11}\varphi(\beta)) = \phi_{12}\varphi(\beta)\sum_{t=1}^{n}\ell_{t-1} \tag{3.72}$$

$$\omega\varphi(\beta)(1-\phi_{11}) = \phi_{12}\varphi(\beta)\ell_{t-1} \tag{3.78}$$

Therefore, the estimator of the magnitude outlier for IO is

$$\hat{\omega}_I = \frac{\phi_{12}}{1-\phi_{11}}\ell_{t-1} \tag{3.79}$$

Its variance is

$$V(\hat{\omega}_I) = \left(\frac{\phi_{12}}{1-\phi_{11}}\right)^2 V(\ell_{t-1}) \tag{3.80}$$

Therefore

$$= \left(\frac{\phi_{12}}{1-\phi_{11}}\right)^2 \varphi\left(\overset{2}{\beta}\right)\sigma_a^2 \tag{3.81}$$

With the estimate of ω and its corresponding variance, the test statistic for innovative model is defined as:

$$\lambda_i = \frac{\hat{\omega}_I}{S.e(\hat{\omega}_I)} \tag{3.82}$$

$$\lambda_i = \frac{\dfrac{\phi_{12}\, e_{t-1}}{1-\phi_{11}}}{\dfrac{\phi_{12}\sigma_a^2}{1-\phi_{11}}}$$

$$= \frac{\ell_{t-1}}{\sigma_a} \tag{3.83}$$

3.7.2 Additive Outlier (AO) Model

An additive outlier represents an unexpected change in the value of one of the observations. It can appear as a result of a recording or measurement error or other single effect.

The additive outlier model is defined as

$$X_t = Z_t + \omega\xi_t^{(T)} \tag{3.84}$$

recall that $\ell_t = \pi(\beta)X_t$ and $Z_t = \dfrac{\theta(\beta)}{\phi(\beta)} a_t$

as defined in equation (3.54) and

$$\ell_t = a_t + \pi(\beta)\omega\xi_t^{(T)}$$

Now considering

$$X_{2t} = \emptyset_{21}X_{2t-1} + \emptyset_{22}X_{1t-1} + a_{1t} \tag{3.85}$$

When X_{2t-1} is contaminated and assumed additive model, we have

$$X_{2t} = \emptyset_{21}\left[Z_{t-1} + \omega\, \xi_{t-1}^{(T)}\right] + \emptyset_{22}X_{1t-1} + a_{1t} \tag{3.86}$$

$$= \emptyset_{21}\, Z_{t-1} + \phi_{21}\, \omega\, \xi_{t-1}^{(T)} + \emptyset_{22} X_{1t-1} + a_{1t}$$

$$X_{2t} = \emptyset_{22}\, X_{1t-1} + \phi_{21}\, Z_{t-1} + \phi_{21}\, \omega\, \xi_{t-1}^{(T)}$$

$$X_{2t} = \emptyset_{22}\, X_{1t-1} + \phi_{21}\, \varphi(\beta) \xi_{t-1} + \phi_{21}\, \omega\, \xi_{t-1}$$

$$X_{2t} = \emptyset_{22}\, X_{1t-1} + \phi_{21}\, \xi_{t-1}\, (\omega + \varphi(\beta)) \qquad (3.87)$$

Therefore in general the additive model is given as

$$Additive\ model: X_{A\,i\,t} = \phi_{i\,i}\, X_{jt-1} + \emptyset_{ij}\, \epsilon_{t-1}\, (\omega + \varphi(\beta))$$

2.7.2.1 Derivation of the Magnitude of Outlier for AO

With $X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t}$ as defined in equation (3.56)

Then,

$$Z_t + \omega\xi_t^{(T)} = \phi_{11}\left(Z_t + \omega\xi_t^{(T)}\right) + \phi_{12}\, X_{2t-1} + a_{1t} \qquad (3.88)$$

$$\frac{\theta(\beta)}{\phi(\beta)} a_t + \omega\xi_t^{(T)} = \phi_{11}\left[\frac{\theta(\beta)}{\phi(\beta)} + \omega\xi_t^{(T)}\right] + \phi_{12}\, \varphi(\beta)\ell_t + a_{1t} \qquad (3.89)$$

$$a_t\left[\frac{\theta(\beta)}{\phi(\beta)} - 1\right] = \phi_{11}\left[\frac{\theta(\beta)}{\phi(\beta)} + \omega \in_t^{(T)}\right] + \phi_{12}\, \varphi(\beta)\ell_t - \omega\xi_t^{(T)} \qquad (3.90)$$

$$a_t = \frac{\phi_{11}\left[\frac{\phi(\beta)}{\theta(\beta)} + \omega\xi_t^{(T)}\right] + \phi_{12}\, \varphi(\beta)\ell_t - \omega\xi_t^{(T)}}{\frac{\phi(\beta)}{\theta(\beta)} - 1} \qquad (3.91)$$

Summing the square of equation (3.91) over n we have

$$\sum_{t=1}^{n} a_t^2 = \sum_{t=1}^{n}\left[\frac{\phi_{11}\left(\frac{\phi(\beta)}{\theta(\beta)} + \omega\xi_t^{(T)}\right) + \phi_{12}\, \varphi(\beta)\ell_t - \omega\xi_t^{(T)}}{\frac{\phi(\beta)}{\theta(\beta)} - 1}\right] \qquad (3.92)$$

Differentiating equation (3.92) with respect to ω and setting to zero, we obtain the magnitude of outlier in the model as

$$\frac{\partial a_t^2}{\partial \hat{\omega}_A} = \frac{2}{\left(\frac{\phi(\beta)}{\theta(\beta)} - 1\right)^2} \sum_{t=1}^{n} \phi_{11}\left(\frac{\phi(\beta)}{\theta(\beta)} + \omega\xi_t^{(T)}\right) + \phi_{12}\varphi(\beta)\ell_t - \omega\xi_t^{(T)} = 0 \tag{3.93}$$

$$\sum_{t=1}^{n} \phi_{11}\left(\frac{\phi(\beta)}{\theta(\beta)} + \omega\right) + \sum_{t=1}^{n} \phi_{12}\varphi(\beta)\ell_t - \sum_{t=1}^{n} \omega = 0$$

$$\phi_{11}\left(\frac{\phi(\beta)}{\theta(\beta)}\right) + \phi_{11}\omega + \phi_{12}\varphi(\beta)\ell_t - \omega = 0$$

$$\omega(1 - \phi_{11}) = \phi_{11}\frac{\theta(\beta)}{\phi(\beta)} + \phi_{12}\varphi(\beta)\ell_t$$

$$\hat{\omega}_A = \frac{\phi_{11}\frac{\theta(\beta)}{\phi(\beta)} + \phi_{12}\varphi(\beta)\ell_t}{1 - \phi_{11}}$$

$$\hat{\omega}_A = \frac{\varphi(\beta)(1 + \phi_{12}\ell_t)}{1 - \phi_{11}} \tag{3.94}$$

Therefore, the estimate of the variance is

$$V(\hat{\omega}_A) = \frac{\varphi\left(\overset{2}{\beta}\right)\phi_{12}^2\sigma_a^2}{(1 - \phi_{11})^2} \tag{3.95}$$

$$= \left(\frac{\phi_{12}}{1 - \phi_{11}}\right)^2 \varphi\left(\overset{2}{\beta}\right)\sigma_a^2 \tag{3.96}$$

Given the estimates of the mean and its variance of the magnitude, the test statistic for testing the presence of outlier for additive model is constructed as follows

$$\lambda_i = \frac{1+\phi_{12}\ell_t}{1-\phi_{11}} X \frac{1-\phi_{11}}{\phi_{12}\sigma_a}$$

$$\lambda_i = \frac{(1+\phi_{12}\ell_t)}{\phi_{12}\sigma_a} \tag{3.97}$$

3.7.3 Multiplicative Outlier Model

Since outlier may have multiplicative interaction effect on a series, there is a need to develop its generating model.

The multiplicative outlier model is defined as:

$$X_t = Z_t \ \omega \xi_t^{(T)} \tag{3.98}$$

Specifying $X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t}$ as defined in equation (3.56)

with the outlier free series $Z_t = \frac{\theta(\beta)}{\phi(\beta)} dt = \varphi(\beta) dt$

For the general case of MO

If we consider

$$X_{2t} = \emptyset_{21} X_{2t-1} + \emptyset_{22} X_{1t-1} + a_{1t} \tag{3.99}$$

When X_{2t-1} is contaminated and assumed multiplicative model, then

$$X_{2t} = \emptyset_{21}(Z_{t-1}\ \omega\ \xi_{t-1}) + \emptyset_{22} X_{1t-1} + a_{1t} \tag{4.0}$$

$$X_{2t} = \emptyset_{21}(\varphi(\beta)\ \epsilon_{t-1}\ \omega\ \xi_{t-1} + \emptyset_{22} X_{1t-1} + a_{1t}$$

$$X_{2t} = \emptyset_{21}\ \omega\ \xi_{t-1}^2\ \varphi(\beta) + \emptyset_{22} X_{1t-1} + a_{1t}$$

$$X_{2t} = \emptyset_{22}\ X_{1t-1} + \emptyset_{21}\ \omega\ \xi_{t-1}^2\ \varphi(\beta) \tag{4.1}$$

Therefore in general, the MO generating mechanism is

$Multiplicative\ model: X_{M\,i\,t} = \phi_{i\,i}\, X_{jt-1} + \emptyset_{ij}\ \omega\ \xi_{t-1}^{2}\ \varphi(\beta)$

3.7.3.1 Derivation of the Magnitude of outlier for MO

By taking the logarithm of equation (3.98) in other to make it linear and further simplification, we have

$$\log\left(\pi(\beta)^{-1}\right) = \log\left(\pi(\beta)^{-1}\right) a_t + \log \omega \xi_t^{(T)} \tag{4.2}$$

with $a_t = \log\left[\pi(\beta)^{-1}\right]\ell_t$, $\Omega_t^T = \log \omega \xi_t^{(T)}$ and $a_t = \ell_t - \Omega_t^T$

From: $$Z_t\, \omega \xi_t^{(T)} = \phi_{11} Z_{t-1}\, \omega \xi_t^{(T)} + \phi_{12} X_{2t-1} + a_t \tag{4.3}$$

$$\log Z_t + \log \omega \xi_t^{(T)} = \phi_{11}\left[\log Z_{t-1} + long\ \omega \xi_t^{(T)}\right] + \phi_{12}\ X_{2t-1} + a_t \tag{4.4}$$

Since, $Z_t = \dfrac{\theta(\beta)}{\phi(\beta)} a_t$

Then $\log Z_t = \log \varphi(\beta) a_t$

where $\varphi(\beta) = \dfrac{\theta(\beta)}{\phi(\beta)}$

$$\log \varphi(\beta) a_t + \log \omega \xi_t^{(T)} = \phi_{11}\left[\log \varphi(\beta)\right] a_{t-1} + \log \omega \xi_t^{(T)} + \phi_{12}\ \varphi(\beta)\ell_t + a_t \tag{4.5}$$

If we let $\Omega_t^T = \log \omega \xi_t^T$, and $\ell_t = \log \varphi(\beta) a_t$

Then we have

$$\ell_t + \Omega_t^{(T)} = \phi_{11}\left(\ell_{t-1} + \Omega_t^{(T)}\right) + \phi_{12}\varphi(\beta)\ell_t + a_t \tag{4.6}$$

$$\ell_t + \Omega_t^{(T)} = \phi_{11}\ell_{t-1} + \phi_{11}\Omega_t^{(T)} + \phi_{12}\varphi(\beta)\ell_t + a_t$$

$$a_t = \ell_t + \Omega_t^{(T)} - \phi_{11}\ell_{t-1} - \phi_{11}\Omega_t^{(T)} - \phi_{12}\varphi(\beta)\ell_t \tag{4.7}$$

By summing the square of equation (4.7) over n we have

$$\sum_{t=1}^{n} a_t^2 = \sum_{t=1}^{n} \left[\ell_t + \Omega_t^{(T)} - \phi_{11}\ell_{t-1} - \phi_{11}\Omega_t^{(T)} - \phi_{12}\varphi(\beta)\ell_t\right]^2 \tag{4.8}$$

Differentiating equation (4.8) with respect to Ω_t we have

$$\frac{\partial \sum a_t^2}{\partial \omega} \Rightarrow 2(1-\phi_{11})\sum_{t=1}^{n}\left(\ell_t + \Omega_t^{(T)} - \phi_{11}\ell_t + \Omega_t^{(T)} - \phi_{12}\varphi(\beta)\ell_t\right) = 0 \tag{4.9}$$

$$(1-\phi_{11})\Omega_t^{(T)} + \ell_t - \phi_{11}\ell_{t-1} - \phi_{12}\varphi(\beta)\ell_t = 0 \tag{4.10}$$

$$\Omega_t^{(T)} = \frac{\phi_{11}\ell_{t-1} + \phi_{12}\varphi(\beta)\ell_t - \ell_t}{1-\phi_{11}} \tag{4.11}$$

$$\Omega_t^{(T)} \approx \frac{\ell_t[(\phi_{11} + \phi_{12}\varphi(\beta) - 1]}{1-\phi_{11}} \tag{4.12}$$

recall that $\Omega_t^{(T)} = \log \omega \xi_t^{(T)}$ when $\xi_t^{(T)} = 1$ there is presence of outlier we have

$$\log \omega_m \cong \frac{\ell_t[\phi_{11} + \phi_{12}\varphi(\beta) - 1]}{1-\phi_{11}} \tag{4.13}$$

$$\hat{\omega}_m \cong Anti\log \frac{\ell_t[\phi_{11} + \phi_{12}\varphi(\beta) - 1]}{1-\phi_{11}} \tag{4.14}$$

Its variance is

$$V(\hat{\omega}_m) = Anti\log\left[\frac{\phi_{11} + \phi_{12}\varphi(\beta) - 1}{1-\phi_{11}}\right]^2 \sigma_a^2 \tag{4.15}$$

Hence the test statistic is defined as

$$\lambda_i = \frac{\hat{\omega}_m}{S.e(\hat{\omega}_m)}. \tag{4.16}$$

$$= \frac{Anti\log\left[\frac{\ell_t(\phi_{11}+\phi_{12}\,\varphi(\beta)-1)}{1-\phi_{11}}\right]}{\left[\left(\frac{\ell_t(\phi_{11}+\phi_{12}\,\varphi(\beta)-1)}{1-\phi_{11}}\right)\sigma_a\right]}$$

$$= Anti\log\left(\frac{\ell_t}{\sigma_a}\right) \tag{4.17}$$

3.7.4 Convolution Outlier Model

The outlier effect on a given series may be either additive or innovative and the effect may be a combination of the two. By this, we propose the convolution of the additive and innovative outliers for the multivariate setting as follows:

The innovative and additive models as defined earlier respectively are as follows:

$$X_{tI} = Z_t + \varphi(\beta)\omega\xi_t^T \quad \text{for innovative model}$$

$$X_{tA} = Z_t + \omega\xi_t^{(T)} \quad \text{for additive model}$$

The convolution involved adding both innovative and additive models: This gives

$$X_{tC} = 2\, Z_t + \omega\xi_t^T \left(1 + \varphi(\beta)\right) \tag{4.18}$$

For the general case of CO, now considering,

$$X_{2t} = \emptyset_{21} X_{2t-1} + \emptyset_{22} X_{1t-1} + a_{1t}$$

Assuming X_{2t-1} is contaminated, we have

$$X_{2t} = \emptyset_{21}\left(2Z_{t-1} + \omega\,\xi_t^{(T)}(1 + \varphi(\beta))\right) + \emptyset_{22} X_{1t-1} + a_{1t} \tag{4.19}$$

$$X_{2t} = 2\emptyset_{21} Z_{t-1} + \emptyset_{21}\omega\,\xi_t^{(T)}(1 + \varphi(\beta)) + \emptyset_{22} X_{1t-1} + a_{1t}$$

where $Z_t = \varphi(\beta)\xi_t^{(T)}$ and $Z_{t-1} = \varphi(\beta)\xi_t^{(T)}$

we then have

$$X_{2t} = 2\emptyset_{21}\varphi(\beta)\xi_t^{(T)} + \emptyset_{21}\omega\xi_t^{(T)}(1 + \varphi(\beta)) + \emptyset_{22} X_{1t-1} + a_{1t} \tag{4.20}$$

$$X_{2t} = 2\emptyset_{21}\varphi(\beta)\xi_t^{(T)} + \emptyset_{21}\omega\,\xi_{t-1} + \emptyset_{21}\omega\,\xi_t^{(T)}\varphi(\beta) + \emptyset_{22} X_{1t-1} + a_{1t}$$

$$X_{2t} = \emptyset_{22}\, X_{1\,t-1} + 2\emptyset_{21}\varphi(\beta)\xi_t^{(T)} + \emptyset_{21}\omega\,\xi_t^{(T)} + \emptyset_{21}\omega\,\xi_t^{(T)}\varphi(\beta)$$

$$X_{2t} = \emptyset_{22}\, X_{1\,t-1} + 2\emptyset_{21}\varphi(\beta)\xi_t^{(T)} + \emptyset_{21}\omega\,\xi_t^{(T)}\left(1 + \varphi(\beta)\right)$$

$$X_{2t} = \emptyset_{22}\, X_{1\,t-1} + \emptyset_{21}\xi_t^{(T)} \left[2\varphi(\beta) + \omega \left(1 + (\beta)\right)\right) \quad (4.21)$$

Therefore in general, the MO generating mechanism is

$$Convolution\ model: X_{C\,i\,t} = \phi_{i\,i}\, X_{jt-1} + \emptyset_{ij}\ \ \xi_{t-1}^{T}\ \left[2\varphi(\beta) + \omega \left(1 + (\beta)\right)\right)$$

3.7.4.1 Derivation of Magnitude of Outlier for CO

Now, specifying $X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t}$ (4.22)

and substituting X_{tC} in equation (4.22) gives

$$2Z_t + \omega\xi_t^{(T)}(1+\varphi(\beta)) = \phi_{11}(2Z_{t-1} + \omega\xi_t^{(T)}(1+\varphi(\beta))) + \phi_{12} X_{2t-1} + a_{1t} \quad (4.23)$$

$$2\varphi(\beta)a_t + \omega\xi_t^{(T)}(1+\varphi(\beta)) = \phi_{11}[2\varphi(\beta)a_{t-1} + \omega\xi_t^{T}(1+\varphi(\beta))] + \phi_{12}\varphi(\beta)\ell_{t-1} + a_{1t} \quad (4.24)$$

$$2\varphi(\beta)a_t - \phi_{11}\, 2\varphi(\beta)a_{t-1} - a_{1t=}\ \phi_{12}\varphi(\beta)\ell_{t-1} - \omega\xi_t^{(T)}(1+\varphi(\beta)) + \phi_{11}\omega\xi_t^{(T)}(1+\varphi(\beta)) \quad (4.25)$$

$$a_t\,[2\varphi(\beta) - 2\varphi(\beta)\phi_{11} - 1] = \phi_{12}\,\varphi(\beta)\ell_{t-1} - \omega\xi_t^{(T)}(1+\varphi(\beta)) + \phi_{11}\omega\xi_t^{(T)}(1+\varphi(\beta)) \quad (4.26)$$

Summing and squaring equation (4.26)

$$\sum_{t=1}^{n} a_t^2 = \frac{\sum_{t=1}^{n}[\phi_{12}\,\varphi(\beta)\ell_{t-1} - \omega\xi_t^{T}(1+\varphi(\beta)) + \phi_{11}\omega\xi_t^{T}(1-\varphi(\beta))]^2}{[2\varphi(\beta) - 2\varphi(\beta)\phi_{11} - 1]^2} \quad (4.27)$$

Differentiating equation (4.27) with respect to ω and equating to 0 we have

$$\frac{\partial \sum a_t^2}{\partial\omega} = \frac{2(1+\varphi(\beta)) + \phi_{11}(1+\varphi(\beta))}{[2\varphi(\beta) - 2\varphi(\beta)\phi_{11} - 1]^2} = 0 \quad (4.28)$$

$$\sum_{t=1}^{n}[\phi_{12}\varphi(\beta)\ell_{t-1} - \omega(1+\varphi(\beta)) + \phi_{11}\,\omega(1+\varphi(\beta))] = 0 \quad (4.29)$$

$$\phi_{12}\,\varphi(\beta)\ell_{t-1} - \omega(1+\varphi(\beta)) + \phi_{11}\,\omega(1+\varphi(\beta)) = 0 \quad (4.30)$$

$$\omega(1+\varphi(\beta))-\phi_{11}\omega\omega(1+\varphi(\beta))=\phi_{12}\varphi(\beta)\ell_{t-1} \quad (4.31)$$

$$\omega[(1+\varphi(\beta))-\phi_{11}(1+\varphi(\beta))]=\phi_{12}\varphi(\beta)\ell_{t-1} \quad (4.32)$$

$$\hat{\omega}_C=\frac{\phi_{12}\varphi(\beta)\ell_{t-1}}{(1-\phi_{11})(1-\varphi(\beta))} \quad (4.33)$$

The corresponding variance is

$$V(\hat{\omega}_C)=\frac{\phi_{12}\varphi(\beta)^2\sigma_a{}^2}{(1-\phi_{11})^2(1-\varphi(\beta))^2} \quad (4.34)$$

The test statistic is

$$\lambda_i=\frac{\hat{\omega}_C}{S.e(\hat{\omega}_C)} \quad (4.35)$$

$$\lambda_i=\frac{\phi_{11}\varphi(\beta)\ell_{t-1}/(1-\phi_{11})(1+\varphi(\beta))}{\sqrt{\phi_{12}^2\varphi\left(\overset{2}{\beta}\right)\sigma_a^2\left|(1-\phi_{11})^2(1+\varphi(\beta))^2\right.}} \quad (4.36)$$

$$=\frac{\phi_{12}\ \varphi(\beta)\ell_{t-1}}{(1-\phi_{11})(1+\varphi(\beta))}*\frac{(1-\phi_{11})(1+\varphi(\beta))}{\phi_{12}\ \varphi(\beta)\sigma_a} \quad (4.37)$$

$$\lambda_i=\frac{\ell_{t-1}}{\sigma_a} \quad (4.38)$$

Table 3.1: Summary of Estimates and Test Statistic for the four models when X_{1t} contains outlier

MODELS	MAGNITUDE	VARIANCE	TEST STATISTIC
Innovative	$\left(\frac{\phi_{12}}{1-\phi_{11}}\ell_{t-1}\right)$	$\left(\frac{\phi_{12}}{1-\phi_{11}}\right)^2\sigma_a^2$	$\frac{\ell_{t-1}}{\sigma_a}$
Additive	$\frac{\varphi(\beta)(1+\phi_{12}\ell_t)}{1-\phi_{11}}$	$\left(\frac{\phi_{12}}{1-\phi_{11}}\right)^2\varphi\left(\overset{2}{\beta}\right)\sigma_a^2$	$\frac{(1+\phi_{12}\ell_t)}{\phi_{12}\sigma_a}$

Multiplicative	$Anti\log \frac{\ell_t\left[\phi_{11}+\phi_{12}\varphi(\beta)-1\right]}{1-\phi_{11}}$	$Anti\log\left[\frac{\phi_{11}+\phi_{12}\varphi(\beta)-1}{1-\phi_{11}}\right]^2 \sigma_a^2$	$Anti\log\left(\frac{\ell_t}{\sigma_a}\right)$
Convolution	$\frac{\phi_{12}\varphi(\beta)\ell_{t-1}}{(1-\phi_{11})(1-\varphi(\beta))}$	$\frac{\phi_{12}\varphi\left(\overset{2}{\beta}\right)\sigma_a^2}{(1-\phi_{11})^2(1-\varphi(\beta))^2}$	$\frac{\ell_{t-1}}{\sigma_a}$

3.8 Estimation of Magnitude of Outlier When X_{2t} Contains an Outlier

We now derive the estimate of magnitude of outliers for the four generating mechanisms when we consider X_{2t} as containing an outlier.

3.8.1 Innovative Outlier Model

The innovative outlier-generating model is defined as $X_t = Z_t + \varphi(\beta)\omega\xi_t^{(T)}$ with the unobservable free series $Z_t = \frac{\theta(\beta)}{\phi(\beta)}dt$ and

$\ell_t = a_t + \varphi(\beta)\omega\xi_t^{(T)}$ as defined in equation (3.24)

Given a vector model X_{1t} and X_{2t} such that X_{2t} contains outlier and X_{1t} is outlier free,

$$X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t}$$
$$X_{2t} = \phi_{21} X_{2t-1} + \phi_{22} X_{1t-1} + a_{2t}$$

as defined in equation (3.56)

If X_{2t} contains an outlier, then

$$X_{21t} = Z_{2t} + \varphi(\beta)\omega\xi_t^{(T)} \tag{4.39}$$

$$Z_{2t} + \varphi(\beta)\omega\xi_t^{(T)} = \phi_{21}\left(Z_{2t-1} + \varphi(\beta)\omega\xi_t^{(T)}\right) + \phi_{22}X_{2t-1} + a_{1t} \tag{4.40}$$

recall that $\ell_t = \pi\{\beta\}X_t$ *and* $Z_t = \dfrac{\theta(\beta)}{\phi[\beta]}a_t$

where $\pi(\beta) = \dfrac{\phi(\beta)}{\theta(\beta)}$

$$X_t = \frac{\ell_t}{\pi(\beta)} = \varphi(\beta)\ell_t$$

Then we have

$$\frac{\theta(\beta)}{\phi(\beta)}a_t + \varphi(\beta)\omega\xi_t^{(T)} = \phi_{21}\left(\frac{\theta(\beta)}{\phi(\beta)}a_t + \varphi(\beta)\omega\xi_t^{(T)}\right) + \phi_{22}\,\varphi(\beta)\ell_{t-1} + a_{1t} \tag{4.41}$$

$$\frac{\theta(\beta)}{\phi(\beta)}a_t - \phi_{21}\frac{\theta(\beta)}{\phi(\beta)}a_t - a_{1t} = \phi_{21}\varphi(\beta)\omega\xi_t^{(T)} + \phi_{22}\varphi(\beta)\ell_{t-1} - \varphi(\beta)\omega\xi_t^{(T)} \tag{4.42}$$

$$a_t\left[\frac{\theta(\beta)}{\phi(\beta)} - \phi_{21}\frac{\theta(\beta)}{\phi(\beta)} - 1\right] = \phi_{21}\varphi(\beta)\omega\xi_t^{T} - \varphi(\beta)\omega\xi_t^{(T)} + \phi_{22}\varphi(\beta)\ell_{t-1} \tag{4.43}$$

Making a_t the subject in equation (4.43) we have

$$a_t = \frac{\phi_{21}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{(T)} + \phi_{22}\varphi(\beta)\ell_{t-1}}{\frac{\theta(\beta)}{\phi(\beta)}(1-\phi_{21}) - 1} \tag{4.44}$$

Using the least squares theory to differentiate the sum of squares of equation (4.44) and equating it to zero, thus we have

$$\sum_{t=1}^{n} a_t = \sum_{t=1}^{n}\left[\frac{\phi_{21}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{(T)} + \phi_{22}\varphi(\beta)\ell_{t-1}}{\frac{\theta(\beta)}{\phi(\beta)}(1-\phi_{21})-1}\right]^2 \tag{4.45}$$

$$\frac{\partial \sum_{a_t^2}}{\partial \omega} = \frac{\left[2\phi_{21}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{(T)}\right]}{\left[\frac{\theta(\beta)}{\phi(\beta)}(1-\phi_{21})-1\right]^2}\sum_{t=1}^{n}\left[\phi_{21}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{T} + \phi_{22}\varphi(\beta)\ell_{t-1}\right] \tag{4.46}$$

$$\sum_{t=1}^{n}\left[\phi_{21}\varphi(\beta)\omega\xi_t^{(T)} - \varphi(\beta)\omega\xi_t^{T} + \phi_{22}\varphi(\beta)\ell_{t-1}\right] = t \tag{4.47}$$

Since $\xi_t^{(T)}$ is a time indictor where $\xi_t^{(T)} = 1$ for all $t = T$ *and* $\xi_t^{(T)} = 0$ otherwise, we have

$$\sum_{t=1}^{n}\omega(\varphi(\beta)-\varphi(\beta)) + \phi_{22}\varphi(\beta)\sum_{t=1}^{n}\ell_{t-1} = 0 \tag{4.48}$$

$$\sum_{t=1}^{n}\omega(\varphi(\beta)-\phi_{21}\varphi(\beta)) = \phi_{22}\varphi(\beta)\sum_{t=1}^{n}\ell_{t-1} \tag{4.49}$$

$$\omega\varphi(\beta)(1-\phi_{21}) = \phi_{22}\varphi(\beta)\ell_{t-1} \tag{4.50}$$

Therefore, the estimator of the magnitude of outlier is

$$\hat{\omega}_I = \frac{\phi_{22}}{1-\phi_{21}}\ell_{t-1} \tag{4.51}$$

The variance of the magnitude is

$$V(\hat{\omega}_I) = \left(\frac{\phi_{22}}{1-\phi_{21}}\right)^2 V(\ell_{t-1}) \tag{4.52}$$

$$V(\hat{\omega}_I) = \left(\frac{\phi_{22}}{1-\phi_{21}}\right)^2 \sigma_a^2 \tag{4.53}$$

With the estimate and its corresponding variance in, then the test statistic for innovative model is

$$\lambda_I = \frac{\ell_{t-1}}{\sigma_a} \tag{4.54}$$

3.8.2 Additive Outlier Model

The additive outlier model as defined in equation (3.57)

$$X_t = Z_t + \omega \xi_t^{(T)}$$

recall that $\ell_t = \pi(\beta) X_t$ and $ARMA(P_1 \in)$ is represented as $Z_t = \frac{\theta(\beta)}{\phi(\beta)} dt$

$$\ell_t = a_t + \pi(\beta)\omega\xi_t^{(T)}$$

With $X_{2t} = \phi_{11} X_{2t-1} + \phi_{12} X_{1t-1} + a_{1t}$ as defined in equation (3.56), then

$$Z_t + \omega\xi_t^{(T)} = \phi_{21}\left(Z_t + \omega\xi_t^{(T)}\right) + \phi_{22} X_{2t-1} + a_{1t} \tag{4.54}$$

$$\frac{\theta(\beta)}{\phi(\beta)} a_t + \omega\xi_t^{(T)} = \phi_{21}\left[\frac{\theta(\beta)}{\phi(\beta)} + \omega\xi_t^{(T)}\right] + \phi_{22}\,\varphi(\beta)\ell_t + a_{1t} \tag{4.55}$$

Hence

$$a_t\left[\frac{\theta(\beta)}{\phi(\beta)} - 1\right] = \phi_{21}\left[\frac{\theta(\beta)}{\phi(\beta)} + \omega\xi_t^{(T)}\right] + \phi_{22}\,\varphi(\beta)\ell_t - \omega\xi_t^{(T)} \tag{4.56}$$

$$a_t = \frac{\phi_{21}\left[\frac{\phi(\beta)}{\theta(\beta)} + \omega\xi_t^{(T)}\right] + \phi_{22}\,\varphi(\beta)\ell_t - \omega\xi_t^{(T)}}{\frac{\phi(\beta)}{\theta(\beta)} - 1} \quad (4.57)$$

Summing its square over n sample we have

$$\sum_{t=1}^{n} a_t^2 = \sum_{t=1}^{n}\left[\frac{\phi_{21}\left(\frac{\phi(\beta)}{\theta(\beta)} + \omega\xi_t^{(T)}\right) + \phi_{22}\,\varphi(\beta)\ell_t - \omega\xi_t^{(T)}}{\frac{\phi(\beta)}{\theta(\beta)} - 1}\right] \quad (4.58)$$

Differentiating (4.58) with respect to ω and setting to zero, we obtain the magnitude of outlier in the model as

$$\frac{\partial \sum a_t^2}{\partial \hat{\omega}_A} = \frac{2}{\left(\frac{\phi(\beta)}{\theta(\beta)} - 1\right)^2}\sum_{t=1}^{n}\phi_{21}\left(\frac{\phi(\beta)}{\theta(\beta)} + \omega\xi_t^{(T)}\right) + \phi_{22}\varphi(\beta)\ell_t - \omega\xi_t^{(T)} = 0 \quad (4.59)$$

$$\sum_{t=1}^{n}\phi_{21}\left(\frac{\phi(\beta)}{\theta(\beta)} + \omega\right) + \sum_{t=1}^{n}\phi_{22}\varphi(\beta)\ell_t - \sum_{t=1}^{n}\omega = 0 \quad (4.60)$$

$$\phi_{11}\left(\frac{\phi(\beta)}{\theta(\beta)}\right) + \phi_{11}\omega + \phi_{12}\varphi(\beta)\ell_t - \omega = 0$$

$$\omega(1 - \phi_{21}) = \phi_{21}\frac{\theta(\beta)}{\phi(\beta)} + \phi_{22}\varphi(\beta)\ell_t \quad (4.61)$$

$$\hat{\omega}_A = \frac{\phi_{21}\frac{\theta(\beta)}{\phi(\beta)} + \phi_{22}\varphi(\beta)\ell_t}{1 - \phi_{21}} \quad (4.62)$$

$$\hat{\omega}_A = \frac{\varphi(\beta)(1 + \phi_{22}\,\ell_t)}{1 - \phi_{21}} \quad (4.63)$$

The estimate of its variance is

$$V(\hat{\omega}_A) = \frac{\varphi\left(\overset{2}{\beta}\right)\phi_{22}^2\sigma_a^2}{(1-\phi_{21})^2} \quad (4.64)$$

$$= \left(\frac{\phi_{22}}{1-\phi_{21}}\right)^2 \varphi\left(\overset{2}{\beta}\right)\sigma_a^2 \quad (4.65)$$

With the derived estimates of the mean and its variance of the magnitude, the test statistic for testing the presence of outlier is constructed as follows

$$\lambda_A = \frac{(1+\phi_{22}\,\ell_t)}{\phi_{22}\,\sigma_a} \quad (4.66)$$

3.8.3 Multiplicative Outlier Model

Outlier may have multiplicative effect on a series as earlier stated; this is developed as follows for a multivariate situation.

$$X_t = Z_t \;\omega\xi_t^{(T)} \quad (4.67)$$

Now, specifying two variable-vector autoregressive equations as defined in equation (3.56)

$$X_{1t} = \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + a_{1t}$$
$$X_{2t} = \phi_{21} X_{2t-1} + \phi_{22} X_{1t-1} + a_{2t}$$

and considering when X_{2t} contains outlier

Assuming multiplicative model, the outlier free series is $Z_t = \frac{\theta(\beta)}{\phi(\beta)} a_t = \varphi(\beta)\mathrm{a}_t$

By taking the logarithm of both sides and further simplification, we have

$$\log\left(\pi(\beta)^{-1}\right) = \log\left(\pi(\beta)^{-1}\right)a_t + \log\omega\xi_t^{(T)} \quad (4.68)$$

and with $a_t = \log\left[\pi(\beta)^{-1}\right]\ell_t$, $\Omega_t^T = \log\omega\xi_t^{(T)}$ and $a_t = \ell_t - \Omega_t^T$

then

$$Z_t\,\omega\xi_t^{(T)} = \phi_{11}\,Z_{t-1}\,\omega\xi_t^{(T)} + \phi_{22}\,X_{1t-1} + \mathrm{a}_t \tag{4.69}$$

$$\log Z_t + \log \omega\xi_t^{(T)} = \phi_{21}\left[\log Z_{t-1} + long\ \omega\xi_t^{(T)}\right] + \phi_{22}\ X_{1t-1} + a_t \tag{4.70}$$

Since $Z_t = \dfrac{\theta(\beta)a_t}{\phi(\beta)} = \varphi(\beta)a_t$

$\log Z_t = \log \varphi(\beta)a_t$

Hence,

$$\log \varphi(\beta)a_t + \log \omega\xi_t^{(T)} = \phi_{21}\left[\log \varphi(\beta)\right]a_{t-1} + \log \omega\xi_t^{(T)} + \phi_{22}\,\varphi(\beta)\ell_t + a_t \tag{4.71}$$

If we let $\Omega_t^T = \log \omega\,\xi_t^T$, and $\ell_t = \log \varphi(\beta)a_t$,we have

$$\ell_t + \Omega_t^{(T)} = \phi_{21}\left(\ell_{t-1} + \Omega_t^{(T)}\right) + \phi_{22}\varphi(\beta)\ell_t + a_t \tag{4.72}$$

$$\ell_t + \Omega_t^{(T)} = \phi_{21}\ell_{t-1} + \phi_{21}\Omega_t^{(T)} + \phi_{22}\varphi(\beta)\ell_t + a_t \tag{4.73}$$

$$a_t = \ell_t + \Omega_t^{(T)} - \phi_{21}\ell_{t-1} - \phi_{21}\Omega_t^{(T)} - \phi_{22}\varphi(\beta)\ell_t \tag{4.74}$$

By summing the square of the equation (4.74) over sample size we have

$$\sum_{t=1}^{n} a_t^2 = \sum_{t=1}^{n}\left[\ell_t + \Omega_t^{(T)} - \phi_{21}\ell_{t-1} - \phi_{11}\Omega_t^{(T)} - \phi_{22}\varphi(\beta)\ell_t\right]^2 \tag{4.75}$$

Differentiating equation (4.75) with respect to Ω_t and equating it to 0 we have

$$\Omega_t^{(T)} \approx \frac{\ell_t\left[(\phi_{21} + \phi_{12}\,\varphi(\beta) - 1\right]}{1 - \phi_{21}} \tag{4.76}$$

recall that $\Omega_t^{(T)} = \log \omega\xi_t^{(T)}$ since $\xi_t^{(T)} = 1$ in the presence of outlier ,we have

$$\log \omega_m \cong \frac{\ell_t[\phi_{21}+\phi_{22}\,\varphi(\beta)-1]}{1-\phi_{21}} \tag{4.77}$$

$$\omega_m \cong Anti\log \frac{\ell_t[\phi_{21}+\phi_{22}\,\varphi(\beta)-1]}{1-\phi_{21}} \tag{4.78}$$

Its variance is

$$V(\hat{\omega}_m) = Anti\log\left[\frac{\phi_{21}+\phi_{22}\,\varphi(\beta)-1}{1-\phi_{21}}\right]^2 \sigma_a^2 \tag{4.79}$$

The test statistic is

$$\lambda_i = \frac{\hat{\omega}_m}{S.e(\hat{\omega}_m)} \tag{4.80}$$

$$\lambda_i = Anti\log\left(\frac{\ell_t}{\delta_a}\right) \tag{4.81}$$

3.8.4 Convolution Outlier Model

The outlier effect on a given series may neither be additive or innovative and the effect may be a combination of the two. By this we produce the convolution of the additive and innovative outliers as follows:

$$X_{tI} = Z_t + \varphi(\beta)\omega\xi_t^T \qquad \text{innovative}$$
$$X_{tA} = Z_t + \omega\xi_t^{(T)} \qquad \text{additive}$$

The convolution involved adding both innovative and additive models which gives

$$X_{tC} = 2\,Z_t + \omega\xi_t^T\,(1+\varphi(\beta)) \tag{4.82}$$

Now given

$$X_{2t} = \phi_{21} X_{2t-1} + \phi_{22} X_{1t-1} + a_{1t} \quad (4.83)$$

$$2Z_t + \omega\xi_t^{(T)}(1+\varphi(\beta)) = \phi_{21}\left(2Z_{t-1} + \omega\xi_t^{(T)}(1+\varphi(\beta))\right) + \phi_{22} X_{2t-1} + a_{1t} \quad (4.84)$$

$$a_t\left[2\varphi(\beta) - 2\varphi(\beta)\phi_{21} - 1\right] = \phi_{22}\,\varphi(\beta)\ell_{t-1} - \omega\xi_t^{(T)}(1+\varphi(\beta)) + \phi_{21}\omega\xi_t^{(T)}(1+\varphi(\beta)) \quad (4.85)$$

Summing the square of the equation (4.85) over sample size, we have

$$\sum_{t=1}^{n} a_t^2 = \frac{\sum_{t=1}^{n}\left[\phi_{22}\,\varphi(\beta)\ell_{t-1} - \omega\xi_t^{T}(1+\varphi(\beta)) + \phi_{21}\omega\xi_t^{T}(1-\varphi(\beta))\right]^2}{\left[2\varphi(\beta) - 2\varphi(\beta)\phi_{21} - 1\right]^2} \quad (4.86)$$

Differentiating equation (4.86) with respect to ω and equating to zero

$$\frac{\partial\sum a_t^2}{\partial\omega} = \frac{2(1+\varphi(\beta)) + \phi_{21}(1+\varphi(\beta))}{\left[2\varphi(\beta) - 2\varphi(\beta)\phi_{21} - 1\right]^2} = 0 \quad (4.87)$$

$$\sum_{t=1}^{n}\left[\phi_{12}\varphi(\beta)\ell_{t-1} - \omega(1+\varphi(\beta)) + \phi_{11}\,\omega(1+\varphi(\beta))\right] = 0 \quad (4.88)$$

$$\phi_{22}\,\varphi(\beta)\ell_{t-1} - \omega(1+\varphi(\beta)) + \phi_{21}\,\omega(1+\varphi(\beta)) = 0 \quad (4.89)$$

$$\omega\left[(1+\varphi(\beta)) - \phi_{21}(1+\varphi(\beta))\right] = \phi_{22}\,\varphi(\beta)\ell_{t-1} \quad (4.90)$$

$$\hat{\omega}_C = \frac{\phi_{22}\,\varphi(\beta)\ell_{t-1}}{(1-\phi_{21})(1-\varphi(\beta))} \quad (4.91)$$

The corresponding variance is

$$V(\hat{\omega}_C) = \frac{\phi_{22}\,\varphi\left(\overset{2}{\beta}\right)\sigma_a^2}{(1-\phi_{21})^2(1-\varphi(\beta))^2} \quad (4.92)$$

The test statistic is

$$\lambda_{C_1} = \frac{\hat{\omega}_C}{S.e(\hat{\omega}_C)} \quad (4.93)$$

$$\lambda_{C_1} = \frac{\phi_{21}\varphi(\beta)\ell_{t-1} / (1-\phi_{21})(1+\varphi(\beta))}{\sqrt{\phi_{22}^{2}\varphi\left(\overset{2}{\beta}\right)\sigma_a^2 \left|(1-\phi_{21})^2 (1+\varphi(\beta))^2\right.}} \tag{4.94}$$

$$\lambda_{C_1} = \frac{\ell_{t-1}}{\delta_a} \tag{4.95}$$

Table 3.2: Summary of Estimates and Test Statistic for the four models when X_{2t} contains outlier

MODEL	MAGNITUDE	VARIANCE	TEST STATISTIC (λ)
Innovative	$\frac{\phi_{22}}{1-\phi_{21}}\ell_{t-1}$	$\left(\frac{\phi_{22}}{1-\phi_{21}}\right)^{2}\sigma_a^2$	$\frac{\ell_{t-1}}{\sigma_a}$
Additive	$\frac{\varphi(\beta)(1+\phi_{22}\ell_t)}{1-\phi_{21}}$	$\left(\frac{\phi_{22}}{1-\phi_{21}}\right)^{2}\varphi\left(\overset{2}{\beta}\right)\sigma_a^2$	$\frac{(1+\phi_{22}\ell_t)}{\phi_{22}\sigma_a}$
Multiplicative	$Anti\log\frac{\ell_t[\phi_{21}+\phi_{22}\varphi(\beta)-1]}{1-\phi_{21}}$	$Anti\log\left[\frac{\phi_{21}+\phi_{22}\varphi(\beta)-1}{1-\phi_{21}}\right]^{2}\sigma_a^2$	$Anti\log\left(\frac{\ell_t}{\sigma_a}\right)$
Convolution	$\frac{\phi_{22}\varphi(\beta)\ell_{t-1}}{(1-\phi_{21})(1-\varphi(\beta))}$	$\frac{\phi_{22}\varphi\left(\overset{2}{\beta}\right)\sigma_a^2}{(1-\phi_{21})^2(1-\varphi(\beta))^2}$	$\frac{\ell_{t-1}}{\sigma_a}$

By comparing the test statistic value λ with some critical value C (Table 3.3), Tsay (1988) the existence of outlier can be determined. Comparing λ with the critical value C, if $\lambda < C$, no significant disturbance is found. On the other hand, if $\lambda \geq C$ an outlier is detected.

Table 3.3 Default Critical Values for Outlier Identification

Number of Observations	Outlier Critical Value (C)
1	1.96
2	2.24
3	2.44
4	2.62
5	2.74
6	2.84
7	2.92
8	2.99
9	3.04
10	3.09
11	3.13
12	3.16
24	3.42
36	3.55
48	3.63
72	3.73
96	3.80
120	3.85
144	3.89
168	3.92
192	3.95
216	3.97
240	3.99
264	4.01
288	4.03
312	4.04
336	4.05
360	4.07

Source: Tsay Critical Values (1988)

CHAPTER FOUR

ANALYSIS AND INTERPRETATION OF DATA

4.0 Introduction

In this section, analysis of both simulated and real data sets and their results are presented. From the derived outlier generating mechanisms in chapter three and with the estimation of the magnitudes of outliers and their variances, the test statistics constructed were used to detect the existence of outliers in both the generated series and real data.

Simulated data used was assumed to be normally distributed while contaminated observations assumed a uniform distribution with interval {0,1}. The simulated data were of varying sizes of 10, 50, and 100. Also in this section, the results obtained from the analysis of real data sets of Gross Domestic Product (GDP), Consumer Price Index (CPI), Deposits and Loans are presented.

Statistical software R3.0.1 is used in analysing the data.

The results and outcomes for the four models i.e. Innovative, Additive, Multiplicative, and Convolution models are summarised below.

4.1 Analysis of Simulated Data when X_{1t} Contains Outliers

The results of the four models in terms of their outlier detection performance are tabulated below. The results from the simulation experiment are firstly tabulated.

Table 4.1: Innovative Model Detection Performance

Sample size	No of outliers Injected	No of outliers Detected	% Detected
10	2	None	0
50	5	2	40
100	8	2	25

Source: Computer Output

Table 4.2: Additive Model Detection Performance

Sample size	No of outliers injected	No of outliers Detected	% Detected
10	2	1	50
50	5	5	100
100	8	8	100

Source: Computer Output

Table 4.3: Multiplicative Model Detection Performance

Sample size	No of outliers injected	No of outliers Detected	% Detected
10	2	2	100
50	5	4	80
100	8	5	62.5

Source: Computer Output

Table 4.4: Convolution Model Detection Performance

Sample size	No of outliers injected	No of outliers Detected	% Detected
10	2	2	100
50	5	5	100
100	8	8	100

Source: Computer Output

4.1.1 Analysis of Simulated Data when X_{2t} Contains Outliers

The results of the four models in terms of their outlier detection performance are tabulated below.

Table 4.5: Innovative Model Detection Performance case

Sample size	No of outliers Injected	No of outliers Detected	% Detected
10	2	None	0
50	5	1	20
100	8	3	37.5

Source: Computer Output

Table 4.6: Additive Model Detection Performance

Sample size	No of outliers Injected	No of Outliers Detected	% Detected
10	2	1	50
50	5	3	60
100	8	6	75

Source: Computer Output

Table 4.7: Multiplicative Model Detection Performance

Sample size	No of outliers Injected	No of Outliers Detected	% Detected
10	2	2	100
50	5	3	60
100	8	4	50

Source: Computer Output

Table 4.8: Convolution Model Detection Performance

Sample size	No of Outliers Injected	No of Outliers Detected	% Detected
10	2	2	100
50	5	4	80
100	8	6	75

Source: Computer Output

Table 4.9: Summary of Result on Detection Rate of the Models on Simulated Data when X_{1t} contains outliers

	n=10			n=50			n=100		
Model Type	No of outliers injected	No of outliers Detected	% of outliers detected	No of outliers injected	No of outliers detected	% of outliers detected	No of outliers injected	No of outliers detected	% of outliers detected
Innovative	2	0	0%	5	2	40%	8	2	25%
Additive	2	1	50%	5	5	100%	8	8	100%
Multiplicative	2	2	100%	5	4	80%	8	5	80%
Convolution	2	2	100%	5	5	100%	8	8	100%

Source: **Tables 4.1- 4.4**

Table 4.10: Summary of Result on Detection Rate of the Models on Simulated Data When X_{2t} Contains Outliers

	n=10			n=50			n=100		
Model Type	No of	No of	% of	No of	No of	% of	No of	No of	% of

	outliers injected	outliers Detected	outliers detected	outliers injected	outliers detected	outliers detected	outliers injected	outliers detected	outliers detected
Innovative	2	0	0%	5	1	20%	8	3	37.5%
Additive	2	1	50%	5	3	60%	8	6	75%
Multiplicative	2	2	100%	5	3	60%	8	4	50%
Convolution	2	2	100%	5	5	100%	8	8	100%

Source: Tables 4.5- 4.8

The convolution model from the summary tables of (4.9) and (4.10) is more sensitivity to outlying observations than the other three models. The pattern of detection is not different when we consider X_{1t} and X_{2t} separately as containing outliers.

4.2 Detection of Outlier in Real Data

Introduction

In order to investigate the performance of the proposed models, two different real data were used. The first pair is GDP and CPI and the second pair is Deposit and Loan. The data was obtained from the Annual statistical bulletin of the Central bank of Nigeria, 2011.

4.2.1 The Assumed Model

The vector autoregressive model is given as $X_{1t} = \emptyset_{11}X_{1t-1} + \emptyset_{12}X_{2t-1}$ (4.96)

Where X_{1t} is the current value of GDP

X_{1t-1} is the past value of GDP

X_{2t-1} is the past value of CPI

$\emptyset_{11}$ and $\emptyset_{12}$ are the coefficients of the model to be estimated

4.2.2 VAR Modelling of GDP and CPI

The estimated VAR model via the use of statistical package R is as follows

X_{1t} =	1.02865 X_{1t-1} +	0.62087 X_{2t-1}
Standard error	(0.01229)	(0.87628)
t-value	(83.813)	(0.709)
P - value	(2 X 10^{-16})	(0.480)

Detection performance of the four generating models on GDP and CPI data are shown on the tables 4.11 to 4.14.

Table 4.11: Detection Performance of Innovative Model on GDP and CPI Data

GDP	CPI	($\acute{\omega}_I$)	T	Remarks

11.24	63.1	22.6113650	0.321803170	ND
11.96	63.9	18.6573683	0.265530200	ND
11.74	64.6	20.6425816	0.293783600	ND
12.68	65.8	12.5790522	0.179024083	ND
11.42	67.9	15.0541890	0.214250035	ND
12.34	70.9	6.9987988	0.099606355	ND
12.18	70.5	10.6966899	0.152234450	ND
13.13	72.1	2.9081796	0.041388984	ND
12.23	73.1	6.8013567	0.096796374	ND
13.36	73.6	2.4526353	0.034905713	ND
13.23	72.6	7.3759771	0.104974326	ND
14.29	72.9	2.7455734	0.039074784	ND
13.86	74.2	4.9030520	0.069779851	ND
15.02	75.7	-0.9426245	-0.013415358	ND
14.99	76.3	0.1923962	0.002738168	ND
15.75	81.1	-7.5726541	-0.107773419	ND
16.65	83.5	-12.8521915	-0.182911380	ND
17.12	82.2	-12.0282753	-0.171185469	ND
17.10	83.9	-12.2586593	-0.174464277	ND
18.05	85.6	-25.8590499	-0.368023967	ND
15.07	87.1	-13.5844287	-0.193332522	ND
17.46	87.8	-21.5424258	-0.306590111	ND
17.42	85.2	-14.4191792	-0.205212625	ND
18.39	84.9	1.3131528	0.018688687	ND
25.00	88.0	-18.4227708	-0.262191427	ND
26.45	89.4	-24.9093581	-0.354508028	ND
26.42	89.6	-22.6121455	-0.321814279	ND
27.36	91.5	-15.2230304	-0.216652973	ND
32.23	92.0	-22.2500642	-0.316661165	ND
34.96	92.3	-29.3707479	-0.418002176	ND
35.33	97.0	-34.9835355	-0.497882929	ND
36.56	96.2	8.2414729	0.117291994	ND
53.26	95.8	-34.6542245	-0.493196201	ND
54.38	92.5	-33.8218265	-0.481349578	ND
53.79	95.6	-33.0019193	-0.469680723	ND
55.41	95.2	-8.3066725	-0.118219910	ND
65.93	97.3	-38.0302202	-0.541243107	ND
67.10	95.9	-41.2163021	-0.586587174	ND
66.26	94.8	-31.5925616	-0.449622855	ND
68.26	99.1	-22.1969518	-0.315905274	ND
76.45	100.0	-41.7140367	-0.593670894	ND
78.24	105.2	-57.9655457	-0.824961093	ND
77.32	109.5	-55.0174012	-0.783003332	ND
80.13	108.1	85.4678691	1.216371999	ND
133.93	113.4	-75.4292968	-1.073503827	ND
133.26	114.4	-82.1625527	-1.169330996	ND
130.71	115.9	-66.6926646	-0.949164764	ND
134.72	117.9	5.6500031	0.080410401	ND
166.75	119.1	-73.6093948	-1.047603124	ND
171.23	119.7	-88.7242019	-1.262715872	ND

170.64	116.1	-68.5054107	-0.974963624	ND
175.25	116.4	17.2629028	0.245684276	ND
211.79	119.8	-53.8165914	-0.765913502	ND
225.29	120.0	-85.2378150	-1.213097887	ND
227.72	121.2	-74.0992812	-1.054575150	ND
235.07	122.2	-96.2970540	-1.370492108	ND
235.07	123.8	552.3621033	7.861174064	ND
475.14	126.6	-106.1447848	-1.510644239	ND
481.12	127.6	-89.6274389	-1.275570670	ND
493.98	129.7	350.2009229	4.984032026	D
670.62	130.8	-132.3757374	-1.883961096	ND
675.14	132.8	-160.4063098	-2.282890001	ND
670.70	135.0	-109.5034824	-1.558444959	ND
686.26	137.3	-156.5603922	-2.228155203	ND
686.35	139.9	-122.7184249	-1.746518984	ND
700.53	142.0	-167.4861923	-2.383650332	ND
699.92	152.9	-142.7692094	-2.031880173	ND
715.17	156.7	-374.0851356	-5.323950264	D
647.96	153.9	-99.5000821	-1.416077353	ND
678.29	150.2	-159.6559806	-2.272211375	ND
685.02	146.9	-139.9146342	-1.991254083	ND
697.17	144.7	46.5296062	0.662205701	ND
777.02	144.9	-116.3477677	-1.655852290	ND
799.25	147.2	-176.4037879	-2.510564852	ND
801.41	151.3	-148.8580771	-2.118536460	ND
816.33	154.6	750.1930238	10.676688185	D
1165.09	154.7	-279.8235308	-3.982426509	ND
1144.27	154.4	-274.4982751	-3.906637888	ND
1124.63	157.5	-161.1198042	-2.293044396	ND
1148.14	162.5	-191.4746653	-2.725052395	ND
1164.24	163.5	-188.3331247	-2.680342237	ND
1182.58	159.4	-236.8987028	-3.371523729	ND
1181.00	158.3	-186.4925287	-2.654147019	ND
1197.27	157.1	932.1238016	13.265912723	D
1625.55	156.4	36.5958886	0.520829812	ND
1735.60	157.6	-118.5925082	-1.687799262	ND
1792.35	159.2	-370.4825358	-5.272678346	D
1758.88	161.1	481.1571226	6.847790372	D
2039.52	161.9	-64.2113899	-0.913851457	ND
2127.69	164.3	-195.2475829	-2.778748263	ND
2171.58	165.1	-382.4142356	-5.442489361	D
2148.24	169.2	986.3419219	14.037540751	D
2631.26	172.4	-472.9035779	-6.730326573	D
2592.27	175.5	697.7065541	9.929704871	D
2985.54	179.4	180.8497330	2.573839194	ND
3202.00	182.3	-516.0370280	-7.344198446	D
3169.61	183.5	195.6664848	2.784710042	ND
3399.35	184.6	978.1905297 \	13.921530778	D
3924.77	185.1	2369.9105068	33.728380168	D
4978.50	186.9	-3314.9069374	-47.177495135	D

3968.28	189.3	742.5847865	10.568408348	D
4426.08	191.7	981.3571638	13.966598066	D
4986.49	196.4	-104.2784884	-1.484083256	ND
5165.74	199.3	-1762.3491124	-25.081614129	D
4740.81	120.4	-136.8234783	-1.947260996	ND
4853.84	121.8	1364.6061225	19.420967141	D
5524.36	122.6	-470.3112890	-6.693433321	D
5538.29	127.7	-524.1016452	-7.458973444	D
5535.96	128.3	-18.5965938	-0.264665261	ND
5720.25	129.6	1477.0825951	21.021723464	D
6461.89	130.6	-278.8314232	-3.968306912	ND
6578.22	138.3	-3648.3458060	-51.922970921	D
5460.76	139.5	586.0113008	8.340066799	D
5872.69	140.4	1431.0597137	20.366729431	D
6608.44	142.4	36.0207076	0.512643881	ND
6852.34	144.7	909.2988607	12.941069957	D
7426.52	146.7	976.6092929	13.899026741	D
8043.20	149.3	1998.0382444	28.435923342	D
9055.63	151.2	264.7342383	3.767676884	ND
9459.40	154.6	-3983.0791809	-56.686870019	D
8311.23	157.5	1546.7543080	22.013285808	D
9170.10	159.7	1435.0210214	20.423106451	D
10013.76	160.3	-826.1272233	-11.757377747	D
10048.57	164.9			

Note: D=Declared, ND= Not Declared

From the above Table, 29 observations were identified as outliers by the innovative model.

Table 4.12: Detection Performance of Multiplicative Model on GDP and CPI Data

GDP	CPI	$(\acute{\omega}_m)$	T	Remarks
11.24	63.1	2.017069e-117	NA	ND
11.96	63.9	5.139871e-97	NA	ND
11.74	64.6	2.920382e-107	NA	ND
12.68	65.8	1.203955e-65	NA	ND
11.42	67.9	2.026029e-78	NA	ND
12.34	70.9	7.582473e-37	NA	ND
12.18	70.5	6.241918e-56	NA	ND
13.13	72.1	9.798034e-16	NA	ND
12.23	73.1	7.921225e-36	NA	ND
13.36	73.6	2.198687e-13	NA	ND
13.23	72.6	8.574854e-39	NA	ND
14.29	72.9	6.766058e-15	NA	ND
13.86	74.2	4.963375e-26	NA	ND
15.02	75.7	7.324915e+04	NA	ND
14.99	76.3	1.016389e-01	NA	ND
15.75	81.1	1.207276e+39	NA	ND
16.65	83.5	2.133229e+66	NA	ND
17.12	82.2	1.193653e+62	NA	ND
17.10	83.9	1.844463e+63	NA	ND
18.05	85.6	2.859549e+133	NA	ND
15.07	87.1	1.282478e+70	NA	ND
17.46	87.8	1.508611e+111	NA	ND
17.42	85.2	2.606898e+74	NA	ND
18.39	84.9	1.670833e-07	NA	ND
25.00	88.0	1.197598e+95	NA	ND
26.45	89.4	3.589393e+128	NA	ND
26.42	89.6	5.003890e+116	NA	ND
27.36	91.5	3.670541e+78	NA	ND
32.23	92.0	6.770760e+114	NA	ND
34.96	92.3	3.800703e+151	NA	ND
35.33	97.0	3.523468e+180	NA	ND
36.56	96.2	2.927377e-43	NA	ND
53.26	95.8	7.037612e+178	NA	ND
54.38	92.5	3.560344e+174	NA	ND
53.79	95.6	2.089404e+170	NA	ND
55.41	95.2	7.413290e+42	NA	ND
65.93	97.3	1.864851e+196	NA	ND
67.10	95.9	5.172904e+212	NA	ND
66.26	94.8	1.112886e+163	NA	ND
68.26	99.1	3.601877e+114	NA	ND
76.45	100.0	1.916458e+215	NA	ND
78.24	105.2	1.429280e+299	NA	ND
77.32	109.5	8.709662e+283	NA	ND
80.13	108.1	0.000000e+00	NA	ND
133.93	113.4	Inf	NA	ND
133.26	114.4	Inf	NA	ND
130.71	115.9	Inf	NA	ND

134.72	117.9	6.931504e-30	NA	ND
166.75	119.1	Inf	NA	ND
171.23	119.7	Inf	NA	ND
170.64	116.1	Inf	NA	ND
175.25	116.4	8.084754e-90	NA	ND
211.79	119.8	5.530023e+277	NA	ND
225.29	120.0	Inf	NA	ND
227.72	121.2	Inf	NA	ND
235.07	122.2	Inf	NA	ND
235.07	123.8	0.000000e+00	NA	ND
475.14	126.6	Inf	NA	ND
481.12	127.6	Inf	NA	ND
493.98	129.7	0.000000e+00	NA	ND
670.62	130.8	Inf	NA	ND
675.14	132.8	Inf	NA	ND
670.70	135.0	Inf	NA	ND
686.26	137.3	Inf	NA	ND
686.35	139.9	Inf	NA	ND
700.53	142.0	Inf	NA	ND
699.92	152.9	Inf	NA	ND
715.17	156.7	Inf	NA	ND
647.96	153.9	Inf	NA	ND
678.29	150.2	Inf	NA	ND
685.02	146.9	Inf	NA	ND
697.17	144.7	7.324314e-241	NA	ND
777.02	144.9	Inf	NA	ND
799.25	147.2	Inf	NA	ND
801.41	151.3	Inf	NA	ND
816.33	154.6	0.000000e+00	NA	ND
1165.09	154.7	Inf	NA	ND
1144.27	154.4	Inf	NA	ND
1124.63	157.5	Inf	NA	ND
1148.14	162.5	Inf	NA	ND
1164.24	163.5	Inf	NA	ND
1182.58	159.4	Inf	NA	ND
1181.00	158.3	Inf	NA	ND
1197.27	157.1	0.000000e+00	NA	ND
1625.55	156.4	1.354615e-189	NA	ND
1735.60	157.6	Inf	NA	ND
1792.35	159.2	Inf	NA	ND
1758.88	161.1	0.000000e+00	NA	ND
2039.52	161.9	Inf	NA	ND
2127.69	164.3	Inf	NA	ND
2171.58	165.1	Inf	NA	ND
2148.24	169.2	0.000000e+00	NA	ND
2631.26	172.4	Inf	NA	ND
2592.27	175.5	0.000000e+00	NA	ND
2985.54	179.4	0.000000e+00	NA	ND
3202.00	182.3	Inf	NA	ND
3169.61	183.5	0.000000e+00	NA	ND

3399.35	184.6	0.000000e+00	NA	ND
3924.77	185.1	0.000000e+00	NA	ND
4978.50	186.9	Inf	NA	ND
3968.28	189.3	0.000000e+00	NA	ND
4426.08	191.7	0.000000e+00	NA	ND
4986.49	196.4	Inf	NA	ND
5165.74	199.3	Inf	NA	ND
4740.81	120.4	Inf	NA	ND
4853.84	121.8	0.000000e+00	NA	ND
5524.36	122.6	Inf	NA	ND
5538.29	127.7	Inf	NA	ND
5535.96	128.3	9.449216e+95	NA	ND
5720.25	129.6	0.000000e+00	NA	ND
6461.89	130.6	Inf	NA	ND
6578.22	138.3	Inf	NA	ND
5460.76	139.5	0.000000e+00	NA	ND
5872.69	140.4	0.000000e+00	NA	ND
6608.44	142.4	1.259722e-186	NA	ND
6852.34	144.7	0.000000e+00	NA	ND
7426.52	146.7	0.000000e+00	NA	ND
8043.20	149.3	0.000000e+00	NA	ND
9055.63	151.2	0.000000e+00	NA	ND
9459.40	154.6	Inf	NA	ND
8311.23	157.5	0.000000e+00	NA	ND
9170.10	159.7	0.000000e+00	NA	ND
10013.76	160.3	*Inf	NA	ND
10048.57	164.9			

***Inf.** Indicates infinite value.

From the table above, no outlier was detected Multiplicative model as a result of non multiplicative nature of the data.

Table 4.13: Detection Performance of Convolution Model on GDP and CPI Data

GDP	CPI	($\acute{\omega}_c$)	T	Remarks
11.24	63.1	-43.3445592	-0.321803170	ND
11.96	63.9	-35.7649972	-0.265530200	ND
11.74	64.6	-39.5705258	-0.293783600	ND
12.68	65.8	-24.1132491	-0.179024083	ND
11.42	67.9	-28.8579300	-0.214250035	ND
12.34	70.9	-13.4162555	-0.099606355	ND
12.18	70.5	-20.5048792	-0.152234450	ND
13.13	72.1	-5.5747968	-0.041388984	ND
12.23	73.1	-13.0377714	-0.096796374	ND
13.36	73.6	-4.7015471	-0.034905713	ND
13.23	72.6	-14.1392824	-0.104974326	ND
14.29	72.9	-5.2630908	-0.039074784	ND
13.86	74.2	-9.3988411	-0.069779851	ND
15.02	75.7	1.8069516	0.013415358	ND
14.99	76.3	-0.3688114	-0.002738168	ND
15.75	81.1	14.5162999	0.107773419	ND
16.65	83.5	24.6368398	0.182911380	ND
17.12	82.2	23.0574444	0.171185469	ND
17.10	83.9	23.4990761	0.174464277	ND
18.05	85.6	49.5701663	0.368023967	ND
15.07	87.1	26.0404923	0.193332522	ND
17.46	87.8	41.2954703	0.306590111	ND
17.42	85.2	27.6406562	0.205212625	ND
18.39	84.9	-2.5172310	-0.018688687	ND
25.00	88.0	35.3152887	0.262191427	ND
26.45	89.4	47.7496670	0.354508028	ND
26.42	89.6	43.3460556	0.321814279	ND
27.36	91.5	29.1815883	0.216652973	ND
32.23	92.0	42.6519685	0.316661165	ND
34.96	92.3	56.3018696	0.418002176	ND
35.33	97.0	67.0612291	0.497882929	ND
36.56	96.2	-15.7983832	-0.117291994	ND
53.26	95.8	66.4299608	0.493196201	ND
54.38	92.5	64.8343064	0.481349578	ND
53.79	95.6	63.2625960	0.469680723	ND
55.41	95.2	15.9233668	0.118219910	ND
65.93	97.3	72.9015314	0.541243107	ND
67.10	95.9	79.0090493	0.586587174	ND
66.26	94.8	60.5609463	0.449622855	ND
68.26	99.1	42.5501553	0.315905274	ND
76.45	100.0	79.9631751	0.593670894	ND
78.24	105.2	111.1162919	0.824961093	ND
77.32	109.5	105.4648851	0.783003332	ND
80.13	108.1	-163.8365097	-1.216371999	ND
133.93	113.4	144.5932004	1.073503827	ND
133.26	114.4	157.5004269	1.169330996	ND
130.71	115.9	127.8456280	0.949164764	ND

134.72	117.9	-10.8306993	-0.080410401	ND
166.75	119.1	141.1045631	1.047603124	ND
171.23	119.7	170.0786942	1.262715872	ND
170.64	116.1	131.3205479	0.974963624	ND
175.25	116.4	-33.0918949	-0.245684276	ND
211.79	119.8	103.1630085	0.765913502	ND
225.29	120.0	163.3955105	1.213097887	ND
227.72	121.2	142.0436445	1.054575150	ND
235.07	122.2	184.5953736	1.370492108	ND
235.07	123.8	-1058.8432834	-7.861174064	D
475.14	126.6	203.4728519	1.510644239	ND
481.12	127.6	171.8101425	1.275570670	ND
493.98	129.7	-671.3130623	-4.984032026	D
670.62	130.8	253.7559322	1.883961096	ND
675.14	132.8	307.4887700	2.282890001	ND
670.70	135.0	209.9112632	1.558444959	ND
686.26	137.3	300.1163887	2.228155203	ND
686.35	139.9	235.2434738	1.746518984	ND
700.53	142.0	321.0604579	2.383650332	ND
699.92	152.9	273.6795619	2.031880173	ND
715.17	156.7	717.0975902	5.323950264	D
647.96	153.9	190.7353764	1.416077353	ND
678.29	150.2	306.0504364	2.272211375	ND
685.02	146.9	268.2075214	1.991254083	ND
697.17	144.7	-89.1943179	-0.662205701	ND
777.02	144.9	223.0313261	1.655852290	ND
799.25	147.2	338.1549256	2.510564852	ND
801.41	151.3	285.3515369	2.118536460	ND
816.33	154.6	-1438.0726696	-10.676688185	D
1165.09	154.7	536.4040443	3.982426509	ND
1144.27	154.4	526.1958653	3.906637888	ND
1124.63	157.5	308.8564936	2.293044396	ND
1148.14	162.5	367.0448462	2.725052395	ND
1164.24	163.5	361.0227113	2.680342237	ND
1182.58	159.4	454.1198587	3.371523729	ND
1181.00	158.3	357.4944049	2.654147019	ND
1197.27	157.1	-1786.8224860	-13.265912723	D
1625.55	156.4	-70.1520083	-0.520829812	ND
1735.60	157.6	227.3343520	1.687799262	ND
1792.35	159.2	710.1916338	5.272678346	D
1758.88	161.1	-922.3478302	-6.847790372	D
2039.52	161.9	123.0891810	0.913851457	ND
2127.69	164.3	374.2772911	2.778748263	ND
2171.58	165.1	733.0639492	5.442489361	D
2148.24	169.2	-1890.7551999	-14.037540751	D
2631.26	172.4	906.5263062	6.730326573	D
2592.27	175.5	-1337.4594205	-9.929704871	D
2985.54	179.4	-346.6775218	-2.573839194	ND
3202.00	182.3	989.2104071	7.344198446	D
3169.61	183.5	-375.0802997	-2.784710042	ND

3399.35	184.6	-1875.1294957	-13.921530778	D
3924.77	185.1	-4542.9688375	-33.728380168	D
4978.50	186.9	6354.4673406	47.177495135	D
3968.28	189.3	-1423.4881589	-10.568408348	D
4426.08	191.7	-1881.1997334	-13.966598066	D
4986.49	196.4	199.8952796	1.484083256	ND
5165.74	199.3	3378.3119976	25.081614129	D
4740.81	120.4	262.2819708	1.947260996	ND
4853.84	121.8	-2615.8637941	-19.420967141	D
5524.36	122.6	901.5570520	6.693433321	D
5538.29	127.7	1004.6697691	7.458973444	D
5535.96	128.3	35.6484963	0.264665261	ND
5720.25	129.6	-2831.4740918	-21.021723464	D
6461.89	130.6	534.5022367	3.968306912	ND
6578.22	138.3	6993.6486029	51.922970921	D
5460.76	139.5	-1123.3466708	-8.340066799	D
5872.69	140.4	-2743.2511335	-20.366729431	D
6608.44	142.4	-69.0494227	-0.512643881	ND
6852.34	144.7	-1743.0685151	-12.941069957	D
7426.52	146.7	-1872.0983646	-13.899026741	D
8043.20	149.3	-3830.1131854	-28.435923342	D
9055.63	151.2	-507.4788232	-3.767676884	ND
9459.40	154.6	7635.3113522	56.686870019	D
8311.23	157.5	-2965.0303673	-22.013285808	D
9170.10	159.7	-2750.8447103	-20.423106451	D
10013.76	160.3	1583.6337365	11.757377747	D
10048.57	164.9			ND

From the above table, 31 observations were identified as outliers by the convolution model.

Table 4.14: Detection Performance of Additive Model on GDP and CPI Data

GDP	CPI	($\acute{\omega}_A$)	T	Remarks
11.24	63.1	8.333881e+00	0.321803170	ND
11.96	63.9	6.876555e+00	0.265530200	ND
11.74	64.6	7.608246e+00	0.293783600	ND
12.68	65.8	4.636267e+00	0.179024083	ND
11.42	67.9	5.548529e+00	0.214250035	ND
12.34	70.9	2.579550e+00	0.099606355	ND
12.18	70.5	3.942484e+00	0.152234450	ND
13.13	72.1	1.071869e+00	0.041388984	ND
12.23	73.1	2.506779e+00	0.096796374	ND
13.36	73.6	9.039689e-01	0.034905713	ND
13.23	72.6	2.718567e+00	0.104974326	ND
14.29	72.9	1.011937e+00	0.039074784	ND
13.86	74.2	1.807120e+00	0.069779851	ND
15.02	75.7	-3.474235e-01	-0.013415358	ND
14.99	76.3	7.091156e-02	0.002738168	ND
15.75	81.1	-2.791057e+00	-0.107773419	ND
16.65	83.5	-4.736938e+00	-0.182911380	ND
17.12	82.2	-4.433267e+00	-0.171185469	ND
17.10	83.9	-4.518180e+00	-0.174464277	ND
18.05	85.6	-9.530882e+00	-0.368023967	ND
15.07	87.1	-5.006819e+00	-0.193332522	ND
17.46	87.8	-7.939902e+00	-0.306590111	ND
17.42	85.2	-5.314483e+00	-0.205212625	ND
18.39	84.9	4.839893e-01	0.018688687	ND
25.00	88.0	-6.790089e+00	-0.262191427	ND
26.45	89.4	-9.180853e+00	-0.354508028	ND
26.42	89.6	-8.334169e+00	-0.321814279	ND
27.36	91.5	-5.610759e+00	-0.216652973	ND
32.23	92.0	-8.200716e+00	-0.316661165	ND
34.96	92.3	-1.082519e+01	-0.418002176	ND
35.33	97.0	-1.289390e+01	-0.497882929	ND
36.56	96.2	3.037563e+00	0.117291994	ND
53.26	95.8	-1.277252e+01	-0.493196201	ND
54.38	92.5	-1.246573e+01	-0.481349578	ND
53.79	95.6	-1.216353e+01	-0.469680723	ND
55.41	95.2	-3.061594e+00	-0.118219910	ND
65.93	97.3	-1.401682e+01	-0.541243107	ND
67.10	95.9	-1.519111e+01	-0.586587174	ND
66.26	94.8	-1.164408e+01	-0.449622855	ND
68.26	99.1	-8.181141e+00	-0.315905274	ND
76.45	100.0	-1.537456e+01	-0.593670894	ND
78.24	105.2	-2.136439e+01	-0.824961093	ND
77.32	109.5	-2.027779e+01	-0.783003332	ND
80.13	108.1	3.150093e+01	1.216371999	ND
133.93	113.4	-2.780101e+01	-1.073503827	ND
133.26	114.4	-3.028269e+01	-1.169330996	ND
130.71	115.9	-2.458095e+01	-0.949164764	ND
134.72	117.9	2.082424e+00	0.080410401	ND

166.75	119.1	-2.713025e+01	-1.047603124	ND
171.23	119.7	-3.270112e+01	-1.262715872	ND
170.64	116.1	-2.524907e+01	-0.974963624	ND
175.25	116.4	6.362596e+00	0.245684276	ND
211.79	119.8	-1.983521e+01	-0.765913502	ND
225.29	120.0	-3.141614e+01	-1.213097887	ND
227.72	121.2	-2.731081e+01	-1.054575150	ND
235.07	122.2	-3.549225e+01	-1.370492108	ND
235.07	123.8	2.035844e+02	7.861174064	D
475.14	126.6	-3.912183e+01	-1.510644239	ND
481.12	127.6	-3.303403e+01	-1.275570670	ND
493.98	129.7	1.290737e+02	4.984032026	D
670.62	130.8	-4.878979e+01	-1.883961096	ND
675.14	132.8	-5.912103e+01	-2.282890001	ND
670.70	135.0	-4.035975e+01	-1.558444959	ND
686.26	137.3	-5.770354e+01	-2.228155203	ND
686.35	139.9	-4.523039e+01	-1.746518984	ND
700.53	142.0	-6.173046e+01	-2.383650332	ND
699.92	152.9	-5.262051e+01	-2.031880173	ND
715.17	156.7	-1.378767e+02	-5.323950264	D
647.96	153.9	-3.667279e+01	-1.416077353	ND
678.29	150.2	-5.884448e+01	-2.272211375	ND
685.02	146.9	-5.156840e+01	-1.991254083	ND
697.17	144.7	1.714944e+01	0.662205701	ND
777.02	144.9	-4.288235e+01	-1.655852290	ND
799.25	147.2	-6.501722e+01	-2.510564852	ND
801.41	151.3	-5.486469e+01	-2.118536460	ND
816.33	154.6	2.764990e+02	10.676688185	D
1165.09	154.7	-1.031347e+02	-3.982426509	ND
1144.27	154.4	-1.011720e+02	-3.906637888	ND
1124.63	157.5	-5.938400e+01	-2.293044396	ND
1148.14	162.5	-7.057190e+01	-2.725052395	ND
1164.24	163.5	-6.941403e+01	-2.680342237	ND
1182.58	159.4	-8.731386e+01	-3.371523729	ND
1181.00	158.3	-6.873564e+01	-2.654147019	ND
1197.27	157.1	3.435533e+02	13.265912723	D
1625.55	156.4	1.348816e+01	0.520829812	ND
1735.60	157.6	-4.370969e+01	-1.687799262	ND
1792.35	159.2	-1.365489e+02	-5.272678346	D
1758.88	161.1	1.773403e+02	6.847790372	D
2039.52	161.9	-2.366642e+01	-0.913851457	ND
2127.69	164.3	-7.196249e+01	-2.778748263	ND
2171.58	165.1	-1.409466e+02	-5.442489361	D
2148.24	169.2	3.635365e+02	14.037540751	D
2631.26	172.4	-1.742983e+02	-6.730326573	D
2592.27	175.5	2.571540e+02	9.929704871	D
2985.54	179.4	6.665587e+01	2.573839194	ND
3202.00	182.3	-1.901960e+02	-7.344198446	D
3169.61	183.5	7.211689e+01	2.784710042	ND
3399.35	184.6	3.605321e+02	13.921530778	D

3924.77	185.1	8.734790e+02	33.728380168	D
4978.50	186.9	-1.221777e+03	-47.177495135	D
3968.28	189.3	2.736948e+02	10.568408348	D
4426.08	191.7	3.616993e+02	13.966598066	D
4986.49	196.4	-3.843397e+01	-1.484083256	ND
5165.74	199.3	-6.495498e+02	-25.081614129	D
4740.81	120.4	-5.042909e+01	-1.947260996	ND
4853.84	121.8	5.029535e+02	19.420967141	D
5524.36	122.6	-1.733428e+02	-6.693433321	D
5538.29	127.7	-1.931684e+02	-7.458973444	D
5535.96	128.3	-6.854155e+00	-0.264665261	ND
5720.25	129.6	5.444090e+02	21.021723464	D
6461.89	130.6	-1.027690e+02	-3.968306912	ND
6578.22	138.3	-1.344672e+03	-51.922970921	D
5460.76	139.5	2.159865e+02	8.340066799	D
5872.69	140.4	5.274463e+02	20.366729431	D
6608.44	142.4	1.327617e+01	0.512643881	ND
6852.34	144.7	3.351407e+02	12.941069957	D
7426.52	146.7	3.599493e+02	13.899026741	D
8043.20	149.3	7.364179e+02	28.435923342	D
9055.63	151.2	9.757322e+01	3.767676884	ND
9459.40	154.6	-1.468045e+03	-56.686870019	D
8311.23	157.5	5.700879e+02	22.013285808	D
9170.10	159.7	5.289064e+02	20.423106451	D
10013.76	160.3	-3.044861e+02	-11.757377747	D
10048.57	164.9			

From the above table, 31 observations were detected as outliers were by the additive model.

4.2.3. Assumed Model of Deposits and Loans

Here two cases are considered. The first case is when loan is contaminated. The vector autoregressive model is given as

$$X_{1t} = \emptyset_{11} X_{1t-1} + \emptyset_{12} X_{2t-1} + \ell_t \quad (4.98)$$

where X_{1t} is the current value of deposit, X_{1t-1} is the immediate past value of deposit, and X_{2t-1} is the immediate past value of loan

4.2.. VAR Modelling of Deposits and Loans

The estimated VAR model via the use of statistical package R is as follows

$$X_{1t} = 0.4826\, X_{1t-1} — 0.1579\, X_{2t-1} \quad (4.99)$$

S.e	(0.1836)	(0.1561)
t	(2.628)	(–1.012)
P-value	(0.0142)	(0.3210)

When deposit is contaminated,

then, the vector autoregressive model is given as

$$X_{2t} = \emptyset_{21} X_{2t-1} + \emptyset_{22} X_{1t-1} + \ell_t \quad (5.00)$$

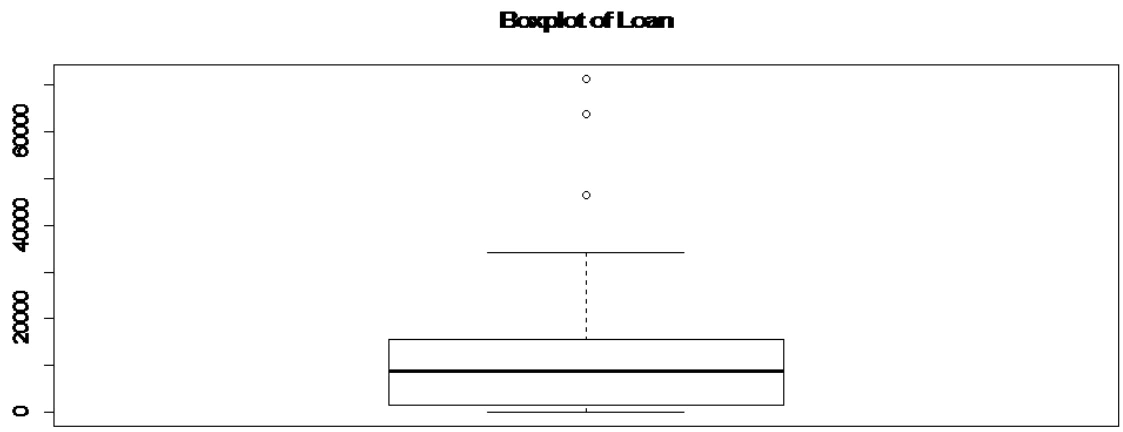
where X_{2t} is the current value of loan , X_{2t-1} is the immediate past value of loan and X_{1t-1} is the immediate past value of deposit.

The estimated VAR model via the use of statistical package R is as follows

$$X_{2t} = 0.9605\, X_{2t-1} — 0.3339\, X_{1t-1} \quad (5.10)$$

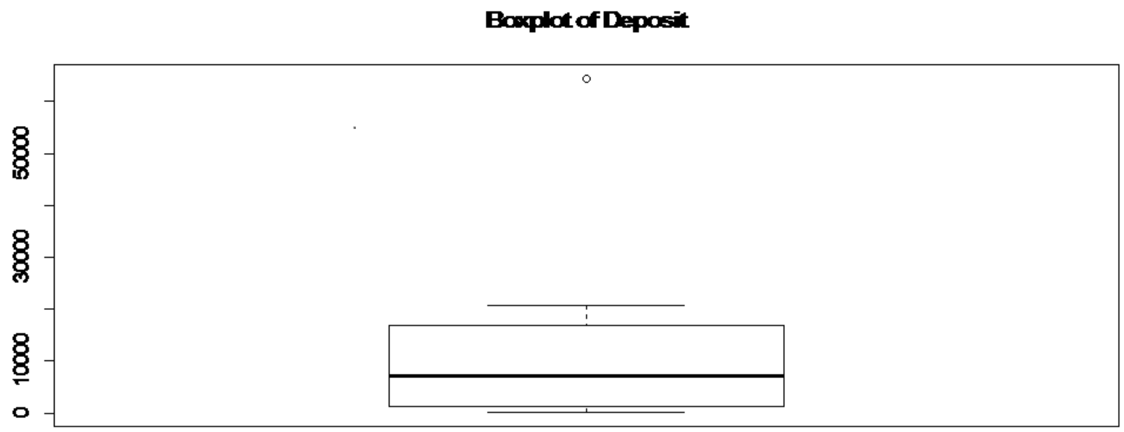
S.e	(0.1712)	(0.2015)
t	(5.610)	(–1.657)
P	(6.78e.06)	(0.1095)

Figure 4.1: Box Plot of Loan Data



From the box plot of loan data above, it shows that there is presence of outlier in the data.

Figure 4.2: Box Plots of Deposit Data



From the box plot of Deposit data on figure 4.2 above, it shows that there is presence of outlier in the data.

The detection performance of the four generating models on Deposits and Loans data for the two cases are shown on the tables 4.15 to 4.20.

Table 4.15: Detection Performance of Additive Model on Deposit and Loan Data for Case One

Deposit	Loan	(W_a)	T	Remarks
111.7	35.9	-7425.1230	-3.5562723	ND
131.2	44.2	-7287.8225	-3.4905121	ND
276.6	58.2	-7320.9776	-3.5063918	ND
311.4	114.9	-6766.7184	-3.2409286	ND
873.5	373.6	-6641.4233	-3.1809184	ND
1229.2	492.8	-6645.0523	-3.1826565	ND
1378.4	659.9	-2347.0674	-1.1241310	ND
5722.0	3721.1	-1321.7090	-0.6330342	ND
8360.1	4730.8	-214.7493	-0.1028544	ND
10580.7	5962.1	-7060.4273	-3.3816009	ND
4612.2	1895.3	10107.6624	4.8410781	D
19542.2	10910.4	-10360.6635	-4.9622533	D
4855.2	1602.2	-790.9809	-0.3788413	ND
8807.1	8659.3	2051.1317	0.9823922	ND
12442.0	4411.2	6231.8871	2.9847704	ND
19047.6	11158.6	3575.8049	1.7126364	ND
18513.8	11852.7	1289.6953	0.6177012	ND
15860.5	7498.1	6662.9285	3.1912183	ND
20640.9	11150.3	1167.7164	0.5592793	ND
16875.9	12341.0	1158.2959	0.5547673	ND
14861.6	8942.2	7283.8830	3.4886253	ND
20551.8	11251.9	48840.8554	23.3923914	D
64490.0	34118.5	-14779.9643	-7.0788832	D
18461.9	16105.5	-10755.6646	-5.1514396	D
3118.6	24274.6	-2097.8978	-1.0047909	ND
3082.3	27263.5	8721.0511	4.1769588	D
13411.8	46521.5	-3338.5137	-1.5989855	ND
3296.2	15590.5	-2683.9759	-1.2854938	ND
3953.1	63769.4	747.8194	0.3581691	ND
94.7	71294.2			

Table 4.16: Detection Performance of Additive Model on Deposit and Loan Data For Case Two

Deposit	Loan	(W_m)	t	Remarks
111.7	35.9	-6512.6824	-2.8431714	ND
131.2	44.2	-6500.1432	-2.8376973	ND
276.6	58.2	-6408.3408	-2.7976201	ND
311.4	114.9	-6192.4792	-2.7033837	ND
873.5	373.6	-6134.0659	-2.6778829	ND
1229.2	492.8	-5962.6850	-2.6030650	ND
1378.4	659.9	-3012.1602	-1.3149862	ND
5722.0	3721.1	-3492.2985	-1.5245950	ND
8360.1	4730.8	-2349.9174	-1.0258780	ND
10580.7	5962.1	-6857.8767	-2.9938691	ND
4612.2	1895.3	4070.3484	1.7769480	ND
19542.2	10910.4	-8911.3893	-3.8903489	ND
4855.2	1602.2	2181.8965	0.9525270	ND
8807.1	8659.3	-7524.7308	-3.2849904	ND
12442.0	4411.2	4516.4853	1.9717132	ND
19047.6	11158.6	935.5785	0.4084354	ND
18513.8	11852.7	-4263.9115	-1.8614498	ND
15860.5	7498.1	2684.7738	1.1720627	ND
20640.9	11150.3	1963.8506	0.8573370	ND
16875.9	12341.0	-3835.6962	-1.6745084	ND
14861.6	8942.2	1065.8428	0.4653035	ND
20551.8	11251.9	23614.0175	10.3089165	D
64490.0	34118.5	-1690.5372	-0.7380196	ND
18461.9	16105.5	8410.6161	3.6717318	ND
3118.6	24274.6	-1569.6781	-0.6852574	ND
3082.3	27263.5	14805.4833	6.4634699	D
13411.8	46521.5	-31173.0466	-13.6088801	D
3296.2	15590.5	43336.2504	18.9188386	D
3953.1	63769.4	4806.4960	2.0983200	ND
94.7	71294.2			

Table 4.17: Detection Performance of Convolution Model on Deposit and Loan Data For Case One

Deposit	Loan	(W_c)	t	Remarks
111.7	35.9	543.83931	3.4944194	ND
131.2	44.2	533.78300	3.4298029	ND
276.6	58.2	536.21138	3.4454064	ND
311.4	114.9	495.61569	3.1845603	ND
873.5	373.6	486.43869	3.1255938	ND
1229.2	492.8	486.70449	3.1273017	ND
1378.4	659.9	171.90659	1.1045794	ND
5722.0	3721.1	96.80611	0.6220241	ND
8360.1	4730.8	15.72891	0.1010655	ND
10580.7	5962.1	517.12786	3.3227859	ND
4612.2	1895.3	-740.31692	-4.7568790	D
19542.2	10910.4	758.84751	4.8759466	D
4855.2	1602.2	57.93392	0.3722523	ND
8807.1	8659.3	-150.23132	-0.9653058	ND
12442.0	4411.2	-456.44297	-2.9328574	ND
19047.6	11158.6	-261.90317	-1.6828492	ND
18513.8	11852.7	-94.46133	-0.6069578	ND
15860.5	7498.1	-488.01380	-3.1357146	ND
20640.9	11150.3	-85.52721	-0.5495519	ND
16875.9	12341.0	-84.83722	-0.5451185	ND
14861.6	8942.2	-533.49445	-3.4279489	ND
20551.8	11251.9	-3577.25755	-22.9855362	D
64490.0	34118.5	1082.53098	6.9557628	D
18461.9	16105.5	787.77864	5.0618425	D
3118.6	24274.6	153.65662	0.9873149	ND
3082.3	27263.5	-638.75715	-4.1043105	D
13411.8	46521.5	244.52322	1.5711749	ND
3296.2	15590.5	196.58282	1.2631357	ND
3953.1	63769.4	-54.77264	-0.3519396	ND
94.7	71294.2			

Table 4.18: Detection Performance of Convolution Model on Deposit and Loan Data For Case Two

Deposit	Loan	(W_c)	T	Remarks
111.7	35.9	13212.666	2.7937212	ND
131.2	44.2	13187.227	2.7883423	ND
276.6	58.2	13000.982	2.7489621	ND
311.4	114.9	12563.051	2.6563648	ND
873.5	373.6	12444.544	2.6313074	ND
1229.2	492.8	12096.854	2.5577908	ND
1378.4	659.9	6110.949	1.2921152	ND
5722.0	3721.1	7085.034	1.4980783	ND
8360.1	4730.8	4767.417	1.0080353	ND
10580.7	5962.1	13912.983	2.9417979	ND
4612.2	1895.3	-8257.758	-1.7460422	ND
19542.2	10910.4	18079.066	3.8226855	ND
4855.2	1602.2	-4426.543	-0.9359600	ND
8807.1	8659.3	15265.869	3.2278558	ND
12442.0	4411.2	-9162.862	-1.9374199	ND
19047.6	11158.6	-1898.064	-0.4013316	ND
18513.8	11852.7	8650.451	1.8290743	ND
15860.5	7498.1	-5446.760	-1.1516774	ND
20640.9	11150.3	-3984.180	-0.8424257	ND
16875.9	12341.0	7781.705	1.6453843	ND
14861.6	8942.2	-2162.339	-0.4572106	ND
20551.8	11251.9	-47907.162	-10.1296173	D
64490.0	34118.5	3429.693	0.7251835	ND
18461.9	16105.5	-17063.117	-3.6078707	ND
3118.6	24274.6	3184.499	0.6733390	ND
3082.3	27263.5	-30036.765	-6.3510532	D
13411.8	46521.5	63242.614	13.3721859	D
3296.2	15590.5	-87918.829	-18.5897902	D
3953.1	63769.4	-9751.224	-2.0618247	ND
94.7	71294.2			

Table 4.19: Detection Performance of Innovative Model on Deposit and Loan Data For Case One

Deposit	Loan	(W_I)	t	Remarks
111.7	35.9	3094.70968	-3.4944194	ND
131.2	44.2	-3037.48437	-3.4298029	ND
276.6	58.2	-3051.30304	-3.4454064	ND
311.4	114.9	-2820.29389	-3.1845603	ND
873.5	373.6	-2768.07229	-3.1255938	ND
1229.2	492.8	-2769.58479	-3.1273017	ND
1378.4	659.9	-978.23192	-1.1045794	ND
5722.0	3721.1	-550.87378	-0.6220241	ND
8360.1	4730.8	-89.50514	-0.1010655	ND
10580.7	5962.1	-2942.70855	-3.3227859	ND
4612.2	1895.3	4212.76267	4.7568790	D
19542.2	10910.4	-4318.21071	-4.8759466	D
4855.2	1602.2	-329.67215	-0.3722523	ND
8807.1	8659.3	854.88915	0.9653058	ND
12442.0	4411.2	2597.38207	2.9328574	ND
19047.6	11158.6	1490.35620	1.6828492	ND
18513.8	11852.7	537.53081	0.6069578	ND
15860.5	7498.1	2777.03540	3.1357146	ND
20640.9	11150.3	486.69136	0.5495519	ND
16875.9	12341.0	482.76499	0.5451185	ND
14861.6	8942.2	3035.84241	3.4279489	ND
20551.8	11251.9	20356.33200	22.9855362	D
64490.0	34118.5	-6160.12677	-6.9557628	D
18461.9	16105.5	-4482.84285	-5.0618425	D
3118.6	24274.6	-874.38077	-0.9873149	ND
3082.3	27263.5	3634.83828	4.1043105	D
13411.8	46521.5	-1391.45583	-1.5711749	ND
3296.2	15590.5	-1118.65169	-1.2631357	ND
3953.1	63769.4	311.68289	0.3519396	ND
94.7	71294.2			

Table 4.20: Detection Performance of Convolution Model on Deposit and Loan Data For Case Two

Deposit	Loan	(W_I)	t	Remarks
111.7	35.9	-4689.5805	-2.7937212	ND
131.2	44.2	-4680.5514	-2.7883423	ND
276.6	58.2	-4614.4474	-2.7489621	ND
311.4	114.9	-4459.0121	-2.6563648	ND
873.5	373.6	-4416.9505	-2.6313074	ND
1229.2	492.8	-4293.5445	-2.5577908	ND
1378.4	659.9	-2168.9631	-1.2921152	ND
5722.0	3721.1	-2514.6958	-1.4980783	ND
8360.1	4730.8	-1692.1026	-1.0080353	ND
10580.7	5962.1	-4938.1442	-2.9417979	ND
4612.2	1895.3	2930.9316	1.7460422	ND
19542.2	10910.4	-6416.8149	-3.8226855	ND
4855.2	1602.2	1571.1159	0.9359600	ND
8807.1	8659.3	-5418.3252	-3.2278558	ND
12442.0	4411.2	3252.1809	1.9374199	ND
19047.6	11158.6	673.6810	0.4013316	ND
18513.8	11852.7	-3070.3104	-1.8290743	ND
15860.5	7498.1	1933.2223	1.1516774	ND
20640.9	11150.3	1414.1079	0.8424257	ND
16875.9	12341.0	-2761.9658	-1.6453843	ND
14861.6	8942.2	767.4803	0.4572106	ND
20551.8	11251.9	17003.7213	10.1296173	D
64490.0	34118.5	-1217.3034	-0.7251835	ND
18461.9	16105.5	6056.2237	3.6078707	ND
3118.6	24274.6	-1130.2765	-0.6733390	ND
3082.3	27263.5	10660.9691	6.3510532	D
13411.8	46521.5	-22446.7435	-13.3721859	ND
3296.2	15590.5	31205.0892	18.5897902	D
3953.1	63769.4	3461.0086	2.0618247	ND
94.7	71294.2			

Table 4.21: Detection Performance of Multiplicative Model on Deposit and Loan Data For Case One

Deposit	Loan	(W_m)	t	Remarks
111.7	35.9	Inf	NA	ND
131.2	44.2	Inf	NA	ND
276.6	58.2	Inf	NA	ND
311.4	114.9	Inf	NA	ND
873.5	373.6	Inf	NA	ND
1229.2	492.8	Inf	NA	ND
1378.4	659.9	Inf	NA	ND
5722.0	3721.1	Inf	NA	ND
8360.1	4730.8	Inf	NA	ND
10580.7	5962.1	Inf	NA	ND
4612.2	1895.3	0	NA	ND
19542.2	10910.4	Inf	NA	ND
4855.2	1602.2	0	NA	ND
8807.1	8659.3	Inf	NA	ND
12442.0	4411.2	0	NA	ND
19047.6	11158.6	0	NA	ND
18513.8	11852.7	Inf	NA	ND
15860.5	7498.1	0	NA	ND
20640.9	11150.3	0	NA	ND
16875.9	12341.0	Inf	NA	ND
14861.6	8942.2	0	NA	ND
20551.8	11251.9	0	NA	ND
64490.0	34118.5	Inf	NA	ND
18461.9	16105.5	0	NA	ND
3118.6	24274.6	Inf	NA	ND
3082.3	27263.5	0	NA	ND
13411.8	46521.5	Inf	NA	ND
3296.2	15590.5	0	NA	ND
3953.1	63769.4	0	NA	ND
94.7	71294.2			

From Table 4.21 above the multiplicative model could not detect any outlier as a result of non multiplicative nature of the data analysed.

Table 4.22: Summary of Outlier Detection of the Four Models on Deposits and Loan Data

Model	No of outliers for Case 1	No of outliers for Case 2
Convolution	6	4
Innovative	6	4
Multiplicative	—	—
Additive	6	4

Source : From Tables 4.15 -4.21

4.3 Discussion of Results

Results obtained from the analysed simulated data with varying sample sizes of 10, 50, and100 gave an average percentage rates of outlier detection for IO, AO, MO, and CO as

21%, 71%, 86% and 100% respectively of the injected outliers. However, with real data on GDP and CPI, CO detected 30 observations as outliers, IO and AO identified the same number of 29 observations as outliers while MO detected nO outlier, as the data did not exhibit any multiplicative effect of outliers on the observations .

For the second real data set of Deposit and Loan, 6 outliers were equally detected by all the models when we consider the case of deposit depending on loan .For the second case of Deposit depending on Loan, 4 outliers were equally detected by all models except for MO as a result of non multiplicative nature of data.

All the four derived outlier-generating mechanisms were able to detect potential outlier independently in multivariate time series. However as the sample size increases, CO was found to be most sensitive to outliers for the simulated data sets.

Of the four-outlier detection models, CO has been found to be most efficient with minimum standard error of the estimate and is therefore recommended for outlier detection in multivariate time series data.

CHAPTER FIVE

SUMMARY, CONCLUSION, CONTRIBUTION TO KNOWLEDGE AND AREA OF FURTHER RESEARCH

5.0 Introduction

This chapter is divided into four sections. The first section gives the summary of the findings, while the second gives the conclusion based on the analysed data. The third section highlights contribution to knowledge and the last section highlights the suggested the area for further studies.

5.1 Summary

This project work is undertaken to develop test statistic for detecting outliers assuming different outliers generating mechanism in multivariate time series models. In line with the main objectives of the study, the test statistics were derived for each generating mechanism namely; the Additive, Innovative, Multiplicative, and Convolution models. Attempts were also made to determine the model with greatest detective power in terms of their sensitivity to the number of outliers detected by applying the models to both simulated and two different pairs of real data. All these were achieved using theoretical and analytical methods. The convolution model was found to be most sensitive to outlier detection.

5.2 Conclusion

Having considered necessary statistical techniques in accordance with main purpose of the work, which is outlier generating mechanisms, identification and estimation of the magnitude of outliers and deriving models for discovering outlier's occurrences in multivariate time series data using both simulated and real data, there is no doubt that the research has been successful as the proposed models were able to detect outliers in multivariate time series and enjoy ability to identify outliers independently. This is indeed a great promise over the conventional approaches of outlier detection in multivariate time series.

5.3 Contribution to Knowledge

The research work made use of the theory of maximum likelihood and Vector Autoregressive (VAR) model in multivariate time series from which these contributions were made.

(a) New outlier generating mechanisms; Convolution, Multiplicative, Innovative and additive were developed by extending the outlier detection in univariate time series to multivariate time series. For each model, estimates of the magnitude of outliers and their residual variances were obtained.

(b) For each generating mechanism, appropriate test statistic was developed, both simulated data and empirical data were used to validate the performance of the models.

(c) All the derived outlier generating mechanisms were able to detect varying number of potential outliers independently in multivariate time series with the swamping effect ameliorated. The pair wise relative efficiency of the variances indicated that convolution model has the minimum residual variance and was the most sensitive to outliers.

5.4 Area of Further Research

The outlier detection in multivariate time series can be extended to frequency domain since this work is limited to time domain. In addition the present work only considered integration of order one, this can be extended to higher orders of integration.

We hope that this will be a subject for further research.

REFERENCES

Arnold, Z. (1962) An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, Vol. 57, No. 298 (Jun., 1962), pp. 348 – 368

Azami, Z., Ibrahim, A. and Mohd, S. (2007): "Detection Procedure for a Single Additive Outlier in Bilinear Model*" Journal of Pak. Stat. Oper. Res.* Vol No 1 PP. 1-5

Baragona, R. & Battaglia, F. (2007). Outlier Detection in Multivariate Time Series by Independent Component Analysis. Neutral Computation, 19:1962-1984.

Baragona, R., Battaglia, F. and Calzini (2001).''Genetic Algorithms for the Identification of Addiditive and Innovational Outliers in Time Series.'' Computational Statistics and Data Analysis.30,147.

Barnett, V. (1978). The study of outliers: purpose and model. Applied Statistics, 27(3), 242–250

Barnett, V. and Lewis, T. (1994): "Outlier in Statistical Data" John Wiley & Sons U. K.

Box, G.E.P., Jenkins, G.M. and Reinsel, G. (1994) *Time Series Analysis: Forecasting and Control*, 3rd Ed., New *Jersey:'* Prentice-Hall.

Chaloner K. and Brant, R. (1988): "A Bayesian Approach to Outlier Detection and Residual Analysis," Biometrika, 25, 651 – 660.

Chang, I. (1982): "Outlier in time series". Technical Report, Department of statistics, University of Wisconsin.

Chang, I., et. al. (1988). Estimation of time series parameters in the presence of *outliers. Technometrics*, 3, 193.204.

Chattfield, S. (1980) "The Analysis of Time series an introduction, New York, Charian and Hall.

Chen, C and Liu, L. M. (1993). "Joint Estimation of Model Parameters and Outlier effects in Time Series. *Journal of the American Statistical Association*, 88, 284 – 297

Cucina, D., Di Salvatore, A. and Protopapas, M. (2008) ''Meta-heuristic Methods for Outliers Detection in Multivariate Time Series.'' Comisef working paper series, 003,270.

Enders, W. (2003) Applied Econometric Time Series, 2nd Edition, John Wiley & Sons, ISBN 0-471-23065-0

Forni, M., & Reichlin, I. (1998). Let's get real: a dynamic factor analytical approach to disaggregated business cycle. *Review of Economic Studies*, 65, 453/474.

Fox, A. J. (1972). Outerliers in Time Series. *Journal of the Royal Statistical Society*. B34: 350 – 363

Galeano, P., Pena, D., & Tsay, R. S. (2004). Outlier detection in multivariate time series via projection pursuit. Working paper 0-42. *Statistics and Econometrics Series II*, Dept. De Estadistica, Universidad Carlos III de Madrid.

Georgiev, I. (2005). A factor model for innovational outliers in multivariate time series, ICEE, First Italian Congress of Econometrics and Empirical Economics ,Venice,24-25.

Hamilton, J. D. (1995). Time Series Analysis. Princeton University Press.

Hawkins, D. (1980), Identification of Outliers, Chapman and Hall.

Helbling, J and Cleroux, R. (2009).On Outlier Detection in Multivariate Time Series Mathematical Volume 34, Number 1, pp. 19-26

Kaya, A. (1999). An Investigation: The Analysis of Outliers in Time Series. PhD Thesis. Dokuz Fylit Unversity, Izmir, Turkey.

Kaya, A. (2010). Modelling Outlier Factors in Data analysis, (advances in Information Systems), LNCS 3261, 88 – 95

Khattree and Naik, D.N. (1987). Detection of Outliers in Bivariate Time Series Data. Communications in Statistics – Theory and Methods, 16(12): 3701 – 3714

Ljung, G. M. (1993). On Outlier Detection in Time Series. J. R. Statist. Soc. B. 55 No. 2, 559 -567

Lutkepohl, H. (2005). ´´New *Introduction to Multiple Time Series Analysis´´, Springer, Berlin.*

Nelson, C.R. and Plosser, C.I. (1982). Trends and random walks in macroeconomic Time Series. *Journal of Monetary Economics*, 10, 139 – 162.

Olivier, D. and Amelie, C. (2008) "The Impact of Outliers on Transitory and Permanent Components in macroeconomic Series". Economic Bulleting, Vol. 3, No 60 PP 1 – 9

Pankratz, A. (1983). Forcasting with Univariate Box-Jenkins Models: Concepts and Cases, New York: John Wiley and Sins

Pankratz, A. (1993). Detecting and treating outliers in dynamic regression models. *Biometrika*, 80, 84'7-54.

Pena, D. and Box, G.E.P. (1987). Identifying a Simplifying Structure in Time Series. *Journal of the American Statistical Association*, 82, 836-843.

Rao, C. R. (1968). A note on a previous lemma in the theory of least squares and some further results, Sankhya, Series A30: 259–266.

Rao, C. R. (1973).Linear Statistical Inference and Its Applications, 2nd Edn, Wiley, New York

Rao, C. R. (1974). Characterization of prior distributions and solution to a compound decision problem, Discussion Paper, Indian Statistical Institute, New Delhi.

Rolin, B. (1990): Comparing classical and Resistant Outlier rules. *Journal of American Stat. Ass.* Vol. 412 pp. 1083 – 1090.

Rosner, B. (1975): On the Detection of Many Outliers, "Technometrics. 17, 221 – 227

Ruey, S. Tsay (1988): Outliers, Level Shifts, and Variance Changes in Time Series. *Journal of Forecasting*, Vol. 7, I-20 Department of Statistics, Carnegie Mellon University, U. S.A.

Shangodoyin, D.K. (1994): "On the Specification of time series Models in the Presence of aberrant Observations". Ph.D Thesis in the Dept. of Statistics, Univ. of Ibadan.

Shittu, I. O. and Shangodoyin, D.K. (2008) ''Detection of Outliers in Time Series Data: A Frequency Domain Approach'' *Assian Journal of Scientific Research* 1, (2) 130-137

Shittu, I.O. (2000) '' On Performance of Some generating Models in Detection of Outliers Under Classical Rule'' Mphil in the Thesis Dept. of Statistics, Univ. of Ibadan.

Sims, C. (1980) "Macroeconomics and reality" *Econometricsa* 48 (1), 11-46, JSTOR 112017

Sridevi, S., Abirami, S. and Rajaram, S. (2012) Detecting and Revamping of X-Outliers in Time Series Database. *International Journal of Computer Applications* 60(19):28-33.

Tsay, R. S. (1986): "Time series model specification in the presence of outlier". *Jour. Amer. Stat. Asso.* 81, 132 – 141.

Tsay, R. S. (1988). Outliers, level shifts and variance changes in time series. *Journal of Forecasting*, 7, 1-20.

Tsay, R. S., Pena, D. & Pankratz, A. E. (2000). Outliers in multivariate time series. Biometrika, 87, 789-804.

Yanjie, Ji et al. (2013). Detection of Outliers in a Time Series of Available Parking Spaces. Mathematical Problems in Engineering Volume 2013: 1-12.

APPENDIX A

OUTPUT FOR INNOVATIVE

[1] "INNOVATIVE"

> WI=(P11/(1-P12))*e

> WI

1 2 3 4 5 6 7 8

22.6113650 18.6573683 20.6425816 12.5790522 15.0541890 6.9987988 10.6966899 2.9081796

9 10 11 12 13 14 15 16

6.8013567 2.4526353 7.3759771 2.7455734 4.9030520 -0.9426245 0.1923962 -7.5726541

17 18 19 20 21 22 23 24

-12.8521915 -12.0282753 -12.2586593 -25.8590499 -13.5844287 -21.5424258 -14.4191792 1.3131528

25 26 27 28 29 30 31 32

-18.4227708 -24.9093581 -22.6121455 -15.2230304 -22.2500642 -29.3707479 -34.9835355 8.2414729

33 34 35 36 37 38 39 40

-34.6542245 -33.8218265 -33.0019193 -8.3066725 -38.0302202 -41.2163021 -31.5925616 -22.1969518

41 42 43 44 45 46 47 48

-41.7140367 -57.9655457 -55.0174012 85.4678691 -75.4292968 -82.1625527 -66.6926646 5.6500031

49 50 51 52 53 54 55 56

-73.6093948 -88.7242019 -68.5054107 17.2629028 -53.8165914 -85.2378150 -74.0992812 -96.2970540

57 58 59 60 61 62 63 64

552.3621033 -106.1447848 -89.6274389 350.2009229 -132.3757374 -160.4063098 -109.5034824 -156.5603922

65 66 67 68 69 70 71 72

-122.7184249 -167.4861923 -142.7692094 -374.0851356 -99.5000821 -159.6559806 -139.9146342 46.5296062

73 74 75 76 77 78 79 80

-116.3477677 -176.4037879 -148.8580771 750.1930238 -279.8235308 -274.4982751 -161.1198042 -191.4746653

81 82 83 84 85 86 87 88

-188.3331247 -236.8987028 -186.4925287 932.1238016 36.5958886 -118.5925082 -370.4825358 481.1571226

89 90 91 92 93 94 95 96

-64.2113899 -195.2475829 -382.4142356 986.3419219 -472.9035779 697.7065541 180.8497330 -516.0370280

97 98 99 100 101 102 103 104

195.6664848 978.1905297 2369.9105068 -3314.9069374 742.5847865 981.3571638 -104.2784884 -1762.3491124

105 106 107 108 109 110 111 112

-136.8234783 1364.6061225 -470.3112890 -524.1016452 -18.5965938 1477.0825951 -278.8314232 -3648.3458060

113 114 115 116 117 118 119 120

586.0113008 1431.0597137 36.0207076 909.2988607 976.6092929 1998.0382444 264.7342383 -3983.0791809

121 122 123

1546.7543080 1435.0210214 -826.1272233

> ABS(W1)

> var(e)

[1] 83164.05

> W=(P11/(1-P12))

> var(W1)=W^2*var

> varW1

[1] 612201.8

> t=e/sd(e)

> t

1 2 3 4 5 6 7 8

0.0288987778 0.0238453159 0.0263825461 0.0160768373 0.0192402211 0.0089449147 0.0136710572 0.0037168405

9 10 11 12 13 14 15 16

0.0086925710 0.0031346256 0.0094269728 0.0035090192 0.0062664156 -0.0012047347 0.0002458947 -0.0096783387

17 18 19 20 21 22 23 24

-0.0164259268 -0.0153729089 -0.0156673545 -0.0330495279 -0.0173617730 - 0.0275326048 -0.0184286378 0.0016782937

25 26 27 28 29 30 31 32

-0.0235454853 -0.0318357607 -0.0288997754 -0.0194560113 -0.0284370121 - 0.0375377036 -0.0447112069 0.0105331320

33 34 35 36 37 38 39 40

-0.0442903263 -0.0432264682 -0.0421785741 -0.0106164613 -0.0486050658 - 0.0526770832 -0.0403773244 -0.0283691311

41 42 43 44 45 46 47 48

-0.0533132201 -0.0740836931 -0.0703157750 0.1092334301 -0.0964034895 - 0.1050090233 -0.0852375120 0.0072210671

49 50 51 52 53 54 55 56

-0.0940775377 -0.1133952328 -0.0875543181 0.0220630993 -0.0687810629 - 0.1089394062 -0.0947036441 -0.1230738245

57 58 59 60 61 62 63 64

0.7059542710 -0.1356598574 -0.1145496277 0.4475792887 -0.1691847008 -0.2050095740 -0.1399524888 -0.2000942441

65 66 67 68 69 70 71 72

-0.1568420348 -0.2140581189 -0.1824682261 -0.4781048476 -0.1271675003 - 0.2040506051 -0.1788198955 0.0594678275

73 74 75 76 77 78 79 80

-0.1486999254 -0.2254553792 -0.1902501904 0.9587949030 -0.3576324580 - 0.3508264389 -0.2059214657 -0.2447169292

81 82 83 84 85 86 87 88

-0.2407018385 -0.3027717688 -0.2383494385 1.1913141306 0.0467718979 -0.1515688481 -0.4735004935 0.6149497290

89 90 91 92 93 94 95 96

-0.0820662835 -0.2495389605 -0.4887499728 1.2606083732 -0.6044011684 0.8917138212
0.2311375829 -0.6595284900

97 98 99 100 101 102 103 104

0.2500743441 1.2501903701 3.0288979535 -4.2366641313 0.9490710867 1.2542375320
-0.1332746107 -2.2523954405

105 106 107 108 109 110 111 112

-0.1748691996 1.7440543345 -0.6010880566 -0.6698355891 -0.0237676421 1.8878064959
-0.3563644807 -4.6628204372

113 114 115 116 117 118 119 120

0.7489601082 1.8289862954 0.0460367796 1.1621423891 1.2481694478 2.5536213001
0.3383473724 -5.0906312053

121 122 123

1.9768514231 1.8340491012 -1.0558436906

```
f=sd(e)/sqrt(124)
> t=e/f
> t
          1           2           3           4           5           6           7           8
 0.321803170   0.265530200   0.293783600   0.179024083   0.214250035   0.099606355
0.152234450  0.041388984
          9          10          11          12          13          14          15          16
 0.096796374   0.034905713   0.104974326   0.039074784   0.069779851  -0.013415358
0.002738168  -0.107773419
         17          18          19          20          21          22          23          24
 -0.182911380  -0.171185469  -0.174464277  -0.368023967  -0.193332522  -0.306590111  -
0.205212625  0.018688687
         25          26          27          28          29          30          31          32
 -0.262191427  -0.354508028  -0.321814279  -0.216652973  -0.316661165  -0.418002176  -
0.497882929  0.117291994
         33          34          35          36          37          38          39          40
 -0.493196201  -0.481349578  -0.469680723  -0.118219910  -0.541243107  -0.586587174  -
0.449622855  -0.315905274
         41          42          43          44          45          46          47          48
 -0.593670894  -0.824961093  -0.783003332   1.216371999  -1.073503827  -1.169330996  -
0.949164764  0.080410401
         49          50          51          52          53          54          55          56
 -1.047603124  -1.262715872  -0.974963624   0.245684276  -0.765913502  -1.213097887  -
1.054575150  -1.370492108
         57          58          59          60          61          62          63          64
  7.861174064  -1.510644239  -1.275570670   4.984032026  -1.883961096  -2.282890001  -
1.558444959  -2.228155203
         65          66          67          68          69          70          71          72
 -1.746518984  -2.383650332  -2.031880173  -5.323950264  -1.416077353  -2.272211375  -
1.991254083  0.662205701
         73          74          75          76          77          78          79          80
 -1.655852290  -2.510564852  -2.118536460  10.676688185  -3.982426509  -3.906637888  -
2.293044396  -2.725052395
         81          82          83          84          85          86          87          88
 -2.680342237  -3.371523729  -2.654147019  13.265912723   0.520829812  -1.687799262  -
5.272678346  6.847790372
         89          90          91          92          93          94          95          96
 -0.913851457  -2.778748263  -5.442489361  14.037540751  -6.730326573   9.929704871
2.573839194  -7.344198446
```

97 98 99 100 101 102 103 104

2.784710042 13.921530778 33.728380168 -47.177495135 10.568408348 13.966598066 -1.484083256 -25.081614129

105 106 107 108 109 110 111 112

-1.947260996 19.420967141 -6.693433321 -7.458973444 -0.264665261 21.021723464 -3.968306912 -51.922970921

113 114 115 116 117 118 119 120

8.340066799 20.366729431 0.512643881 12.941069957 13.899026741 28.435923342 3.767676884 -56.686870019

121 122 123

22.013285808 20.423106451 -11.757377747

APPENDIX B

OUTPUT FOR ADDITIVE MODEL

```
gdp=x1t
cpi=x2t
gdpt=matrix(gdp)
cpit=matrix(cpi)
gdpt
       [,1]
 [1,]  11.24
 [2,]  11.96
 [3,]  11.74
 [4,]  12.68
 [5,]  11.42
 [6,]  12.34
 [7,]  12.18
 [8,]  13.13
 [9,]  12.23
[10,]  13.36
[11,]  13.23
[12,]  14.29
[13,]  13.86
[14,]  15.02
[15,]  14.99
[16,]  15.75
[17,]  16.65
[18,]  17.12
[19,]  17.10
[20,]  18.05
[21,]  15.07
[22,]  17.46
[23,]  17.42
[24,]  18.39
[25,]  25.00
[26,]  26.45
```

[27,]	26.42
[28,]	27.36
[29,]	32.23
[30,]	34.96
[31,]	35.33
[32,]	36.56
[33,]	53.26
[34,]	54.38
[35,]	53.79
[36,]	55.41
[37,]	65.93
[38,]	67.10
[39,]	66.26
[40,]	68.26
[41,]	76.45
[42,]	78.24
[43,]	77.32
[44,]	80.13
[45,]	133.93
[46,]	133.26
[47,]	130.71
[48,]	134.72
[49,]	166.75
[50,]	171.23
[51,]	170.64
[52,]	175.25
[53,]	211.79
[54,]	225.29
[55,]	227.72
[56,]	235.07
[57,]	235.07
[58,]	475.14
[59,]	481.12
[60,]	493.98
[61,]	670.62

[62,] 675.14
[63,] 670.70
[64,] 686.26
[65,] 686.35
[66,] 700.53
[67,] 699.92
[68,] 715.17
[69,] 647.96
[70,] 678.29
[71,] 685.02
[72,] 697.17
[73,] 777.02
[74,] 799.25
[75,] 801.41
[76,] 816.33
[77,] 1165.09
[78,] 1144.27
[79,] 1124.63
[80,] 1148.14
[81,] 1164.24
[82,] 1182.58
[83,] 1181.00
[84,] 1197.27
[85,] 1625.55
[86,] 1735.60
[87,] 1792.35
[88,] 1758.88
[89,] 2039.52
[90,] 2127.69
[91,] 2171.58
[92,] 2148.24
[93,] 2631.26
[94,] 2592.27
[95,] 2985.54
[96,] 3202.00

[97,] 3169.61
[98,] 3399.35
[99,] 3924.77
[100,] 4978.50
[101,] 3968.28
[102,] 4426.08
[103,] 4986.49
[104,] 5165.74
[105,] 4740.81
[106,] 4853.84
[107,] 5524.36
[108,] 5538.29
[109,] 5535.96
[110,] 5720.25
[111,] 6461.89
[112,] 6578.22
[113,] 5460.76
[114,] 5872.69
[115,] 6608.44
[116,] 6852.34
[117,] 7426.52
[118,] 8043.20
[119,] 9055.63
[120,] 9459.40
[121,] 8311.23
[122,] 9170.10
[123,] 10013.76
[124,] 10048.57

```
> cpit
      [,1]
 [1,]  63.1
 [2,]  63.9
 [3,]  64.6
 [4,]  65.8
 [5,]  67.9
 [6,]  70.9
 [7,]  70.5
 [8,]  72.1
 [9,]  73.1
[10,]  73.6
[11,]  72.6
[12,]  72.9
[13,]  74.2
[14,]  75.7
[15,]  76.3
[16,]  81.1
[17,]  83.5
[18,]  82.2
[19,]  83.9
[20,]  85.6
[21,]  87.1
[22,]  87.8
[23,]  85.2
[24,]  84.9
[25,]  88.0
[26,]  89.4
[27,]  89.6
[28,]  91.5
[29,]  92.0
[30,]  92.3
[31,]  97.0
[32,]  96.2
[33,]  95.8
```

[34,] 92.5
[35,] 95.6
[36,] 95.2
[37,] 97.3
[38,] 95.9
[39,] 94.8
[40,] 99.1
[41,] 100.0
[42,] 105.2
[43,] 109.5
[44,] 108.1
[45,] 113.4
[46,] 114.4
[47,] 115.9
[48,] 117.9
[49,] 119.1
[50,] 119.7
[51,] 116.1
[52,] 116.4
[53,] 119.8
[54,] 120.0
[55,] 121.2
[56,] 122.2
[57,] 123.8
[58,] 126.6
[59,] 127.6
[60,] 129.7
[61,] 130.8
[62,] 132.8
[63,] 135.0
[64,] 137.3
[65,] 139.9
[66,] 142.0
[67,] 152.9
[68,] 156.7

[69,] 153.9

[70,] 150.2

[71,] 146.9

[72,] 144.7

[73,] 144.9

[74,] 147.2

[75,] 151.3

[76,] 154.6

[77,] 154.7

[78,] 154.4

[79,] 157.5

[80,] 162.5

[81,] 163.5

[82,] 159.4

[83,] 158.3

[84,] 157.1

[85,] 156.4

[86,] 157.6

[87,] 159.2

[88,] 161.1

[89,] 161.9

[90,] 164.3

[91,] 165.1

[92,] 169.2

[93,] 172.4

[94,] 175.5

[95,] 179.4

[96,] 182.3

[97,] 183.5

[98,] 184.6

[99,] 185.1

[100,] 186.9

[101,] 189.3

[102,] 191.7

[103,] 196.4

[104,] 199.3

[105,] 120.4

[106,] 121.8

[107,] 122.6

[108,] 127.7

[109,] 128.3

[110,] 129.6

[111,] 130.6

[112,] 138.3

[113,] 139.5

[114,] 140.4

[115,] 142.4

[116,] 144.7

[117,] 146.7

[118,] 149.3

[119,] 151.2

[120,] 154.6

[121,] 157.5

[122,] 159.7

[123,] 160.3

[124,] 164.9

> summary(fastVARX(gdpt,cpit,1,1,getdiag=FALSE))

```
Call:
lm(formula = varxz$y.p ~ varxz$Z)

Residuals:
    Min       1Q   Median       3Q      Max
-1468.05   -47.01    -9.53     5.96   873.48

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -47.11304  103.18480  -0.457    0.649
varxz$Z.l1    1.02865    0.01227  83.813   <2e-16 ***
varxz$Z.l1    0.62087    0.87628   0.709    0.480
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 290.8 on 120 degrees of freedom
Multiple R-squared: 0.9883,     Adjusted R-squared: 0.9881
F-statistic:  5057 on 2 and 120 DF,  p-value: < 2.2e-16

>
> resid(VARXZ(gdpt, cpit, 1, 1))
NULL
> resid(fastVARX(gdpt,cpit,1,1,getdiag=FALSE))
            1             2             3             4             5             6             7             8
 8.333881e+00  6.876555e+00  7.608246e+00  4.636267e+00  5.548529e+00  2.579550e+00
3.942484e+00  1.071869e+00
            9            10            11            12            13            14            15            16
 2.506779e+00  9.039689e-01  2.718567e+00  1.011937e+00  1.807120e+00 -3.474235e-01
7.091156e-02 -2.791057e+00
```

17 18 19 20 21 22 23 24

-4.736938e+00 -4.433267e+00 -4.518180e+00 -9.530882e+00 -5.006819e+00 -
7.939902e+00 -5.314483e+00 4.839893e-01

25 26 27 28 29 30 31 32

-6.790089e+00 -9.180853e+00 -8.334169e+00 -5.610759e+00 -8.200716e+00 -
1.082519e+01 -1.289390e+01 3.037563e+00

33 34 35 36 37 38 39 40

-1.277252e+01 -1.246573e+01 -1.216353e+01 -3.061594e+00 -1.401682e+01 -
1.519111e+01 -1.164408e+01 -8.181141e+00

41 42 43 44 45 46 47 48

-1.537456e+01 -2.136439e+01 -2.027779e+01 3.150093e+01 -2.780101e+01 -
3.028269e+01 -2.458095e+01 2.082424e+00

49 50 51 52 53 54 55 56

-2.713025e+01 -3.270112e+01 -2.524907e+01 6.362596e+00 -1.983521e+01 -
3.141614e+01 -2.731081e+01 -3.549225e+01

57 58 59 60 61 62 63 64

2.035844e+02 -3.912183e+01 -3.303403e+01 1.290737e+02 -4.878979e+01 -
5.912103e+01 -4.035975e+01 -5.770354e+01

65 66 67 68 69 70 71 72

-4.523039e+01 -6.173046e+01 -5.262051e+01 -1.378767e+02 -3.667279e+01 -
5.884448e+01 -5.156840e+01 1.714944e+01

73 74 75 76 77 78 79 80

-4.288235e+01 -6.501722e+01 -5.486469e+01 2.764990e+02 -1.031347e+02 -
1.011720e+02 -5.938400e+01 -7.057190e+01

81 82 83 84 85 86 87 88

-6.941403e+01 -8.731386e+01 -6.873564e+01 3.435533e+02 1.348816e+01 -
4.370969e+01 -1.365489e+02 1.773403e+02

89 90 91 92 93 94 95 96

-2.366642e+01 -7.196249e+01 -1.409466e+02 3.635365e+02 -1.742983e+02
2.571540e+02 6.665587e+01 -1.901960e+02

97 98 99 100 101 102 103 104

7.211689e+01 3.605321e+02 8.734790e+02 -1.221777e+03 2.736948e+02 3.616993e+02
-3.843397e+01 -6.495498e+02

105 106 107 108 109 110 111 112
-5.042909e+01 5.029535e+02 -1.733428e+02 -1.931684e+02 -6.854155e+00 5.444090e+02 -1.027690e+02 -1.344672e+03

113 114 115 116 117 118 119 120

2.159865e+02 5.274463e+02 1.327617e+01 3.351407e+02 3.599493e+02 7.364179e+02 9.757322e+01 -1.468045e+03

121 122 123

5.700879e+02 5.289064e+02 -3.044861e+02

>

> e=resid(fastVARX(gdpt,cpit,1,1,getdiag=FALSE))

> e

1 2 3 4 5 6 7 8

8.333881e+00 6.876555e+00 7.608246e+00 4.636267e+00 5.548529e+00 2.579550e+00 3.942484e+00 1.071869e+00

9 10 11 12 13 14 15 16

2.506779e+00 9.039689e-01 2.718567e+00 1.011937e+00 1.807120e+00 -3.474235e-01 7.091156e-02 -2.791057e+00

17 18 19 20 21 22 23 24

-4.736938e+00 -4.433267e+00 -4.518180e+00 -9.530882e+00 -5.006819e+00 -7.939902e+00 -5.314483e+00 4.839893e-01

25 26 27 28 29 30 31 32

-6.790089e+00 -9.180853e+00 -8.334169e+00 -5.610759e+00 -8.200716e+00 -1.082519e+01 -1.289390e+01 3.037563e+00

33 34 35 36 37 38 39 40

-1.277252e+01 -1.246573e+01 -1.216353e+01 -3.061594e+00 -1.401682e+01 -1.519111e+01 -1.164408e+01 -8.181141e+00

41 42 43 44 45 46 47 48

-1.537456e+01 -2.136439e+01 -2.027779e+01 3.150093e+01 -2.780101e+01 -3.028269e+01 -2.458095e+01 2.082424e+00

49 50 51 52 53 54 55 56

-2.713025e+01 -3.270112e+01 -2.524907e+01 6.362596e+00 -1.983521e+01 -3.141614e+01 -2.731081e+01 -3.549225e+01

57 58 59 60 61 62 63 64

2.035844e+02 -3.912183e+01 -3.303403e+01 1.290737e+02 -4.878979e+01 -5.912103e+01 -4.035975e+01 -5.770354e+01

65 66 67 68 69 70 71 72

-4.523039e+01 -6.173046e+01 -5.262051e+01 -1.378767e+02 -3.667279e+01 -5.884448e+01 -5.156840e+01 1.714944e+01

73 74 75 76 77 78 79 80

-4.288235e+01 -6.501722e+01 -5.486469e+01 2.764990e+02 -1.031347e+02 -1.011720e+02 -5.938400e+01 -7.057190e+01

81 82 83 84 85 86 87 88

-6.941403e+01 -8.731386e+01 -6.873564e+01 3.435533e+02 1.348816e+01 -4.370969e+01 -1.365489e+02 1.773403e+02

89 90 91 92 93 94 95 96

-2.366642e+01 -7.196249e+01 -1.409466e+02 3.635365e+02 -1.742983e+02 2.571540e+02 6.665587e+01 -1.901960e+02

97 98 99 100 101 102 103 104

7.211689e+01 3.605321e+02 8.734790e+02 -1.221777e+03 2.736948e+02 3.616993e+02 -3.843397e+01 -6.495498e+02

105 106 107 108 109 110 111 112

-5.042909e+01 5.029535e+02 -1.733428e+02 -1.931684e+02 -6.854155e+00 5.444090e+02 -1.027690e+02 -1.344672e+03

113 114 115 116 117 118 119 120

2.159865e+02 5.274463e+02 1.327617e+01 3.351407e+02 3.599493e+02 7.364179e+02 9.757322e+01 -1.468045e+03

121 122 123

5.700879e+02 5.289064e+02 -3.044861e+02

```
> sd(e)
[1] 288.3818
> sd(e)/sqrt(124)
[1] 25.89745
> f=sd(e)/sqrt(124)
> t=e/f
> t
          1            2            3            4            5            6            7            8
 0.321803170  0.265530200  0.293783600  0.179024083  0.214250035  0.099606355
0.152234450  0.041388984
          9           10           11           12           13           14           15           16
 0.096796374  0.034905713  0.104974326  0.039074784  0.069779851 -0.013415358
0.002738168 -0.107773419
         17           18           19           20           21           22           23           24
-0.182911380 -0.171185469 -0.174464277 -0.368023967 -0.193332522 -0.306590111 -
0.205212625  0.018688687
         25           26           27           28           29           30           31           32
-0.262191427 -0.354508028 -0.321814279 -0.216652973 -0.316661165 -0.418002176 -
0.497882929  0.117291994
         33           34           35           36           37           38           39           40
-0.493196201 -0.481349578 -0.469680723 -0.118219910 -0.541243107 -0.586587174 -
0.449622855 -0.315905274
         41           42           43           44           45           46           47           48
-0.593670894 -0.824961093 -0.783003332  1.216371999 -1.073503827 -1.169330996 -
0.949164764  0.080410401
         49           50           51           52           53           54           55           56
-1.047603124 -1.262715872 -0.974963624  0.245684276 -0.765913502 -1.213097887 -
1.054575150 -1.370492108
         57           58           59           60           61           62           63           64
 7.861174064 -1.510644239 -1.275570670  4.984032026 -1.883961096 -2.282890001 -
1.558444959 -2.228155203
```

65 66 67 68 69 70 71 72

-1.746518984 -2.383650332 -2.031880173 -5.323950264 -1.416077353 -2.272211375 -
1.991254083 0.662205701

73 74 75 76 77 78 79 80

-1.655852290 -2.510564852 -2.118536460 10.676688185 -3.982426509 -3.906637888 -
2.293044396 -2.725052395

81 82 83 84 85 86 87 88

-2.680342237 -3.371523729 -2.654147019 13.265912723 0.520829812 -1.687799262 -
5.272678346 6.847790372

89 90 91 92 93 94 95 96

-0.913851457 -2.778748263 -5.442489361 14.037540751 -6.730326573 9.929704871
2.573839194 -7.344198446

97 98 99 100 101 102 103 104

2.784710042 13.921530778 33.728380168 -47.177495135 10.568408348 13.966598066
-1.484083256 -25.081614129

105 106 107 108 109 110 111 112

-1.947260996 19.420967141 -6.693433321 -7.458973444 -0.264665261 21.021723464 -
3.968306912 -51.922970921

113 114 115 116 117 118 119 120

8.340066799 20.366729431 0.512643881 12.941069957 13.899026741 28.435923342
3.767676884 -56.686870019

121 122 123

22.013285808 20.423106451 -11.757377747

APPENDIX C

OUTPUT FOR MULTIPLICATIVE MODEL

[1] "Multiplicative"

> w=e*(P11+P12*fi-1)/(1-P11)

> Wm=10^w

> Wm

```
        1          2          3          4          5          6          7          8
2.017069e-117  5.139871e-97 2.920382e-107  1.203955e-65  2.026029e-78  7.582473e-37
6.241918e-56  9.798034e-16
        9         10         11         12         13         14         15         16
 7.921225e-36  2.198687e-13  8.574854e-39  6.766058e-15  4.963375e-26  7.324915e+04
1.016389e-01  1.207276e+39
       17         18         19         20         21         22         23         24
 2.133229e+66       1.193653e+62       1.844463e+63    2.859549e+133       1.282478e+70
1.508611e+111  2.606898e+74  1.670833e-07
       25         26         27         28         29         30         31         32
 1.197598e+95    3.589393e+128     5.003890e+116          3.670541e+78     6.770760e+114
3.800703e+151 3.523468e+180  2.927377e-43
       33         34         35         36         37         38         39         40
7.037612e+178     3.560344e+174     2.089404e+170          7.413290e+42     1.864851e+196
5.172904e+212 1.112886e+163 3.601877e+114
       41         42         43         44         45         46         47         48
1.916458e+215 1.429280e+299 8.709662e+283   0.000000e+00              Inf              Inf
Inf  6.931504e-30
       49         50         51         52         53         54         55         56
      Inf        Inf        Inf  8.084754e-90 5.530023e+277        Inf        Inf        Inf
       57         58         59         60         61         62         63         64
 0.000000e+00          Inf          Inf  0.000000e+00          Inf        Inf        Inf        Inf
       65         66         67         68         69         70         71         72
      Inf        Inf        Inf        Inf        Inf        Inf        Inf 7.324314e-241
       73         74         75         76         77         78         79         80
      Inf        Inf        Inf  0.000000e+00        Inf        Inf        Inf        Inf
       81         82         83         84         85         86         87         88
      Inf           Inf           Inf   0.000000e+00 1.354615e-189           Inf              Inf
0.000000e+00
       89         90         91         92         93         94         95         96
```

```
        Inf           Inf           Inf  0.000000e+00           Inf  0.000000e+00  0.000000e+00
Inf
         97          98          99         100         101         102         103         104
 0.000000e+00  0.000000e+00  0.000000e+00           Inf  0.000000e+00  0.000000e+00
Inf         Inf
        105         106         107         108         109         110         111         112
        Inf  0.000000e+00           Inf           Inf  9.449216e+95  0.000000e+00           Inf
Inf
        113         114         115         116         117         118         119         120
 0.000000e+00  0.000000e+00 1.259722e-186  0.000000e+00  0.000000e+00  0.000000e+00
0.000000e+00         Inf
        121         122         123
 0.000000e+00  0.000000e+00          Inf
>
>
> sd(e)
[1] 288.3818
> se=sd(e)/sqrt(124)
> t=10^(e/se)
> t
         1          2          3          4          5          6          7          8          9          2
```

APPENDIX D

OUTPUT FOR CONVOLUTION MODEL

```
[1] "CONVOLUTION"
> Wc=P12*fi*e/(1-P11)*(1-fi)
> Wc
         1           2           3           4           5           6           7           8
 -43.3445592  -35.7649972  -39.5705258  -24.1132491  -28.8579300  -13.4162555  -
20.5048792   -5.5747968
         9          10          11          12          13          14          15          16
 -13.0377714   -4.7015471  -14.1392824   -5.2630908   -9.3988411    1.8069516   -
0.3688114   14.5162999
        17          18          19          20          21          22          23          24
  24.6368398   23.0574444   23.4990761   49.5701663   26.0404923   41.2954703
27.6406562   -2.5172310
        25          26          27          28          29          30          31          32
  35.3152887   47.7496670   43.3460556   29.1815883   42.6519685   56.3018696
67.0612291  -15.7983832
        33          34          35          36          37          38          39          40
  66.4299608   64.8343064   63.2625960   15.9233668   72.9015314   79.0090493
60.5609463   42.5501553
        41          42          43          44          45          46          47          48
  79.9631751  111.1162919  105.4648851 -163.8365097  144.5932004  157.5004269
127.8456280  -10.8306993
        49          50          51          52          53          54          55          56
 141.1045631  170.0786942  131.3205479  -33.0918949  103.1630085  163.3955105
142.0436445  184.5953736
        57          58          59          60          61          62          63          64
-1058.8432834  203.4728519  171.8101425 -671.3130623  253.7559322  307.4887700
209.9112632  300.1163887
```

```
        65          66          67          68          69          70          71          72 
 235.2434738   321.0604579   273.6795619   717.0975902   190.7353764   306.0504364 
268.2075214   -89.1943179 
        73          74          75          76          77          78          79          80 
 223.0313261   338.1549256   285.3515369 -1438.0726696   536.4040443   526.1958653 
308.8564936   367.0448462 
        81          82          83          84          85          86          87          88 
 361.0227113   454.1198587   357.4944049 -1786.8224860   -70.1520083   227.3343520 
710.1916338  -922.3478302 
        89          90          91          92          93          94          95          96 
 123.0891810   374.2772911   733.0639492 -1890.7551999   906.5263062 -1337.4594205  -
346.6775218   989.2104071 
        97          98          99         100         101         102         103         104 
 -375.0802997 -1875.1294957 -4542.9688375  6354.4673406 -1423.4881589 -1881.1997334 
199.8952796  3378.3119976 
       105         106         107         108         109         110         111         112 
 262.2819708 -2615.8637941   901.5570520  1004.6697691    35.6484963 -2831.4740918 
534.5022367  6993.6486029 
       113         114         115         116         117         118         119         120 
-1123.3466708 -2743.2511335   -69.0494227 -1743.0685151 -1872.0983646 -3830.1131854 
-507.4788232  7635.3113522 
       121         122         123 
-2965.0303673 -2750.8447103  1583.6337365 
> 
> wc=Wc/e
> varWc=wc^2*var(e)
> varWc
      1       2       3       4       5       6       7       8       9      10      11      12      13      14      15 
2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 
2249626 2249626 2249626 2249626 2249626 
     16      17      18      19      20      21      22      23      24      25      26      27      28      29      30 
2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 
2249626 2249626 2249626 2249626 2249626 
     31      32      33      34      35      36      37      38      39      40      41      42      43      44      45 
2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 
2249626 2249626 2249626 2249626 2249626 
     46      47      48      49      50      51      52      53      54      55      56      57      58      59      60 
```

2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626

61 62 63 64 65 66 67 68 69 70 71 72 73 74 75

2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626

76 77 78 79 80 81 82 83 84 85 86 87 88 89 90

2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626

91 92 93 94 95 96 97 98 99 100 101 102 103 104 105

2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626

106 107 108 109 110 111 112 113 114 115 116 117 118 119 120

2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626 2249626

121 122 123

2249626 2249626 2249626

> sdWc=sqrt(varWc)

> seWc=sdWc/sqrt(124)

> t=Wc/seWc

> t

1 2 3 4 5 6 7 8

-0.321803170 -0.265530200 -0.293783600 -0.179024083 -0.214250035 -0.099606355 -
0.152234450 -0.041388984

9 10 11 12 13 14 15 16

-0.096796374 -0.034905713 -0.104974326 -0.039074784 -0.069779851 0.013415358 -
0.002738168 0.107773419

17 18 19 20 21 22 23 24

0.182911380 0.171185469 0.174464277 0.368023967 0.193332522 0.306590111
0.205212625 -0.018688687

25 26 27 28 29 30 31 32

0.262191427 0.354508028 0.321814279 0.216652973 0.316661165 0.418002176
0.497882929 -0.117291994

33 34 35 36 37 38 39 40

0.493196201 0.481349578 0.469680723 0.118219910 0.541243107 0.586587174
0.449622855 0.315905274

41 42 43 44 45 46 47 48

0.593670894 0.824961093 0.783003332 -1.216371999 1.073503827 1.169330996
0.949164764 -0.080410401

49 50 51 52 53 54 55 56

1.047603124 1.262715872 0.974963624 -0.245684276 0.765913502 1.213097887
1.054575150 1.370492108

57 58 59 60 61 62 63 64

-7.861174064 1.510644239 1.275570670 -4.984032026 1.883961096 2.282890001
1.558444959 2.228155203

65 66 67 68 69 70 71 72

1.746518984 2.383650332 2.031880173 5.323950264 1.416077353 2.272211375
1.991254083 -0.662205701

73 74 75 76 77 78 79 80

1.655852290 2.510564852 2.118536460 -10.676688185 3.982426509 3.906637888
2.293044396 2.725052395

81 82 83 84 85 86 87 88

2.680342237 3.371523729 2.654147019 -13.265912723 -0.520829812 1.687799262
5.272678346 -6.847790372

89 90 91 92 93 94 95 96

0.913851457 2.778748263 5.442489361 -14.037540751 6.730326573 -9.929704871 -
2.573839194 7.344198446

97 98 99 100 101 102 103 104

-2.784710042 -13.921530778 -33.728380168 47.177495135 -10.568408348 -13.966598066 1.484083256 25.081614129

105 106 107 108 109 110 111 112

1.947260996 -19.420967141 6.693433321 7.458973444 0.264665261 -21.021723464 3.968306912 51.922970921

113 114 115 116 117 118 119 120

-8.340066799 -20.366729431 -0.512643881 -12.941069957 -13.899026741 -28.435923342 -3.767676884 56.686870019

121 122 123

-22.013285808 -20.423106451 11.757377747