ESTIMATION AND OUTLIER DETECTION OF RANDOM COEFFICIENT AUTOREGRESSIVE MODELS

NORLI ANIDA ABDULLAH

FACULTY OF SCIENCE

UNIVERSITY OF MALAYA

KUALA LUMPUR

OCTOBER 2009

ESTIMATION AND OUTLIER DETECTION OF RANDOM COEFFICIENT AUTOREGRESSIVE MODELS

NORLI ANIDA ABDULLAH

DISSERTATION SUBMITTED IN FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

FACULTY OF SCIENCE

UNIVERSITY OF MALAYA

KUALA LUMPUR

OCTOBER 2009

Abstract

The class of random coefficient autoregressive (RCA) models has been considered in many areas of science due to its rich applications. We review two methods of RCA parameter estimation, namely least squares and estimating functions. An iterative method based on the estimating functions is proposed to improve the existing RCA parameter estimation. This study is then followed by investigating the robustness of the three estimates when outliers exist in the RCA process. Simulation studies are carried out to investigate the performance of parameter estimation and robustness of the estimates.

Further, the outlier detection procedure for the RCA process is proposed. In this study, a procedure by Chang et al. (1988) has been extended to detect additive and innovational outliers in the RCA process. A simulation study is carried out to investigate the performance of the procedures. It is found that, in general, these procedures work well in detecting outliers. Finally, we apply the suggested procedures to a real data set to show the importance of the study in practice.

To my dear parents

Acknowledgments

Alhamdulillah to the most gracious and mighty Allah S.W.T who has made this project possible.

My greatest debt is to my respected supervisors Assoc. Prof. Dr. Ibrahim Mohamed and Visiting Prof. Dr. Shelton Peiris for their guidance, patience, advice and support throughout my research work.

Special thanks to the staff of Institute of Mathematical Sciences and Center for Foundation Studies in Sciences, notably Assoc. Prof. Dr. Abdul Ghapor Hussin, Assoc. Prof. Dr. Noor Hasnah Moin, Dr Noor Azlinna Azizan, Prof Sim Chiaw Hock, Prof Chia Gek Ling, Assoc. Prof. Dr. Omar Rijal, Mr Malik, Ms Ng and my friends; Ali, Akmal, Ivan, Isfahani, Sima, Siew Ching, Iqbal, Shahrina, Li Yin, Arina, Siti, Aimi and Muhamad Hafiz for their assistance and motivation right up to the completion of this thesis.

My heartfelt appreciation goes to my parents, and to Along, Ajol, Kak Long Su, for their continuous love and prayers which have given me the strength to complete this thesis.

To my late sister Maria Liza, may Allah S.W.T bless you.

Contents

List of Figures

List of Tables

List of Symbols

y_t	Observation at time t
y_t^*	Contaminated observation with outlier at time t
$\{y_t\}$	A sequence of y_t, for $t = 1, 2, ...$
n	Sample size
u_t	Residual at time t
u_t^*	Contaminated residual with outlier at time t
$\{u_t\}$	A sequence of u_t, for $t = 1, 2, ...$
ACF	Autocorrelation function
AIC	Akaike's information criteria
AO	Additive outlier
AR	Autoregressive model
AR(1)	Autoregressive model of order one
IO	Innovational outlier
EF	Estimating functions
IT	Iterative method based on estimating functions
LC	Level change
LS	Least squares
ML	Maximum likelihood

RC	Random coefficient
RCA	Random coefficient autoregressive model
RCA(1)	Random coefficient autoregressive model of order one
TC	Temporary change
log	Natural logarithm

Chapter 1

Introduction

1.1 An Overview

In most time series data, the structure correlation can be approximated by linear time series models. Examples of the models include autoregressive (AR) model, moving average (MA) model, autoregressive moving average (ARMA) model and the autoregressive integrated moving average (ARIMA) model. These linear models have been widely used in various areas including engineering, economics, finance and natural sciences. However, in many applications, linear models are not a reasonable choice. Note that any random burst and cyclicity pattern in time series data cannot be explained well with linear models. See, for example, Tong (1977) and references therein for details.

It is known that the famous Wolfer's sunspot data has systematic periodic cycles with faster downturn trend than upturn. This pattern will never be well explained by any linear model. Granger and Andersen (1978) are among others who have shown that nonlinear models fit the Wolfer's sunspot data better than linear models. The Canadian lynx data is another good example. Moran (1953) carried out a rigorous statistical analysis and an AR(2) model was fitted to the data. Noting that the one-step-ahead predictors of the fitted model were not particularly good, he further suggested that the process may be better represented by a nonlinear model. On the other hand, Tong (1977) fitted the lynx data using an AR(12) model. He pointed out that an AR model does provide a good approximation but can be further improved by a nonlinear model. Later, Nicholls and Quinn (1982) showed that the Canadian lynx data is better fitted using a nonlinear random coefficient autoregressive model of order two, RCA(2). Their conclusion was based on one-step-ahead predictors of the last 14 time points, which have smaller sum squares of error compared to models given by Moran (1953) and Tong (1977).

At present, there are many types of nonlinear time series models available in the literature. The most mentioned models are:

1. Bilinear models given by Granger and Anderson (1978), with the applications in modeling seismological data such as earthquakes and sunspot data.

2. Threshold autoregressive models proposed by Tong (1983) and later Tiao and Tsay (1994) for modeling short-term interest rate yield.

3. The autoregressive conditional heteroscedastic (ARCH) models proposed by Engle (1982) and Bollerslev (1986) for modeling volatility in financial time series.

There are many data sets in engineering, economics, hydrology and meteorology which exhibit occasional random spikes. This has led to the consideration of nonlinear models with random coefficients. The entire issue of Annals of Economics and Social Measurement (Volume 2, Number 4, 1973) reported important issues and usefulness of models with random coefficients. Consequently, Nicholls and Quinn (1982) introduced a special form of random coefficient model, known as random coefficient autoregressive (RCA) models. Since then, the RCA models have been used extensively in modeling data with occasional random spikes. This model will be considered in this study.

Several methods of RCA parameter estimation have been proposed. Nicholls and Quinn (1982) considered the least squares (LS) and maximum likelihood (ML) methods. In their monograph, extensive theoretical properties of these methods and the resulting estimates had been explored. Since then, many approaches based on LS and ML have been extended to improve the RCA parameter estimation (see Tjøstheim, 1986; Hwang & Basawa, 1993; Schick, 1996; and later Aue et al., 2006). Whilst Thavaneswaran and Abraham (1988) applied Godambe's (1985) estimating functions (EF) theory to estimate the parameter of several linear and nonlinear models including RCA models. Consequently, Thavaneswaran and Peiris (1996) and Chandra and Taniguchi (2001) have further improved the RCA parameter estimation using several other approaches based on EF. Besides, the Bayesian method has been put forward by Wang and Gosh (2002) to estimate the parameter of non-stationary case of RCA models.

Outliers may also occur in the RCA process. They are usually displayed as 'bumps' or shocks in a time series plot and may disturb more than one observation. As a result, outliers might cause some changes in mean and variance, and hence, affects the parameter estimation and model forecasts. However, in practice, the presence of outliers cannot be identified easily at the beginning of the analysis. Barnett and Lewis (1978) pointed out that outliers in time series can be cloaked to some extent by the general structure of the process. Thus, the process of identifying outliers will be more difficult. One of the ways of overcoming this problem is to use robust methods. It is done by down-weighting the effects of outliers so that the results obtained would not be adversely affected.

Tukey (1976) defined an estimate with robustness property such the estimator should not change much in the presence of outliers. In other words, an estimator which is not sensitive to outliers is said to be robust. Later, Papantoni-Kazakos and Gray (1979) and Boente et al. (1982) gave a formal definition of qualitative robustness estimates. That is, "a sequence of estimates is robust, if a small change in the distribution of the stochastic process produces a small change in the distribution of the estimates". The above definition is actually the generalization of qualitative robustness from Hampel's (1971) independent and identically distributed observations case. An excellent reference on statistical robustness is available in Huber (2004).

The detection and handling of outliers in RCA data is another important and challenging problem in practice. There are four common types of outliers being considered in the literature. They are additive outlier (AO), innovational outlier (IO), level change (LC) and temporary change (TC). A bulk of work is available on detecting these outliers in linear time series data. For example, Abraham and Box (1979) proposed a procedure for detecting multiple outliers using the Bayesian approach while McCulloch and Tsay (1994) applied the Gibbs sampling to handle AO. Vogelsang (1999) proposed a method to detect AO in the context of unit-root testing, and this was further studied by Perron and Rodriguez (2000) and Harvey et al. (2001) on the occurrence of IO.

Another method adopted by many authors is by examining the maximum value of the standardized-outlier-effects statistics. Chang et al. (1988) had initially considered this approach for detecting AO and IO in ARIMA models. Then, Chen and Liu (1993) extended the work of Chang et al. (1988) to two other types of outliers, namely LS and TC. Franses and Ghijsels (1999) and Charles and Darne (2005) used the same approach to detect AO and IO for the generalized autoregressive conditional heteroskedasticity (GARCH) models. Zaharim et al. (2006) and Ismail et al. (2008) had also applied the same procedure to bilinear models for detection of AO, IO, LS and TC. They have performed a large scale simulation study and showed that the detection procedure works well. This suggests that the approach has been successfully applied to several nonlinear time series models described above. However, no work has been carried out in detecting outliers in RCA models. It is our interest to explore this problem in the study.

1.2 Problem Statement

To the best of our knowledge, no study has been carried out to detect outliers and to investigate the robustness property of the RCA parameter estimates when AO or IO exists in the data set. Therefore, the following problems will be addressed in this project to close the gap in the time series literature:

1. A new iterative method in estimating RCA parameters.
2. The robustness property of the RCA parameter estimates when AO or IO occurs in the data sets.
3. The detection of outliers in RCA models.

1.3 Objectives

This study has several objectives as given below:

1. Propose an iterative (IT) method based on the estimating functions approach in estimating the RCA(1) parameters.
2. Investigate the robustness of least squares (LS), estimating functions (EF) and IT estimates when AO or IO exists in RCA(1) process.

3. Measure the effects of AO and IO on observations and residuals for the RCA(1) model.

4. Derive the test statistics to verify the existence of AO and IO for the RCA(1) model.

5. Propose the AO and IO detection procedures for the RCA(1) process.

6. Apply the suggested estimation method and outlier detection procedures to a real data set from the RCA(1) model.

1.4 Outline of Thesis

This thesis is organized as follows:

Chapter 2 consists of literature review on three important topics covered in this thesis; the RCA models, theory of estimating functions and outliers.

Chapter 3 proposes a new iterative (IT) method based on estimating functions. A simulation study is carried out to justify the IT method. We then apply the theory to a real data set to illustrate the usefulness of this IT method in practice.

Chapter 4 discusses the effects of AO and IO in RCA(1) process. A simulation study is performed to compare the robustness properties of LS, EF and IT estimates for data with and without outliers.

Chapter 5 proposes separate outlier detection procedures for AO and IO cases. We first formulate the effects of AO and IO on observations generated from RCA(1) process and the resulting residuals. Then, we derive two statistics using least squares method to measure the effects of AO and IO. Consequently, test statistics are defined to verify the existence of outliers. A simulation study is carried out to investigate the performance of the proposed outlier detection procedures.

Then, a real data set is considered in chapter 6 to illustrate the application of outlier detection procedures in practice. Finally, the summary, the significance of this study and possible future research are presented in Chapter 7.

Chapter 2

Literature Review

2.1 Random Coefficient Autoregressive Model (RCA)

Quality control engineers and economists had observed regularly recurring cycles in the production lines and prices of particular commodities since more than one century ago. They had noticed an inconsistency between the observed continuation of these regularly recurring cycles and the economic theory assuming the tendency towards equilibrium. Classical economic theory relies on the assumption that an equilibrium of the disturbed price and production will tend to gravitate back towards a normal trend. However in reality, prices and production might tend to fluctuate continually, or even diverge further and further away from equilibrium. As a displacement from equilibrium is uncertain, many authors allow random disturbances. Conlisk (1974, 1976) has contributed to the development of random coefficient (RC) models by proposing a general n-variable model with random coefficients, that is

$$y_t = (b_t + u_t) + (A + U_t)y_{t-1}, \tag{2.1}$$

where y_t is an (nx1) random vector of dependent variables at time t, b_t is an (nx1) constant intercept vector; u_t is an (nx1) random vector of intercept shocks; A is an (nxn) constant matrix of coefficients; and U_t is an (nxn) random matrix of coefficient shocks. Further, shocks u_t and U_t are assumed to have zero means and are serially uncorrelated properties $E[u_t|y_{t-1}] = 0$ and $E[U_t|y_{t-1}] = 0$ for all $t = 1, 2, .., n$.

The model in (2.1) is a RC model which allows random disturbances in errors. The development of RC models was then extended by Andel (1976) with the derivation of second order stationarity conditions. See also Turnovsky (1968) for motivation of RC models.

In 1982, Nicholls and Quinn (1982) proposed a special form of RC model, known as the random coefficient autoregressive model. The univariate random coefficient autoregressive models of order p, denoted by RCA(p), is given by

$$y_t = \sum_{i=1}^{p}[\theta_i + b_i(t)]y_{t-i} + e_t, \tag{2.2}$$

where

1. θ_i are the parameters to be estimated,

2. $\{e_t\}$ is a sequence of i.i.d random variables with mean 0 and variance σ_e^2,

3. $\{b_i(t)\}$ is a sequence of i.i.d random variables with mean 0 and variance σ_b^2,

4. $b_i(t)$ is independent of e_t for all i and t.

5. θ_i and σ_b^2 satisfy $\sum(\theta_i^2 + \sigma_b^2) < 1$ to ensure stationarity.

In this study, we consider a special case of (2.2) when $p = 1$. The random coefficient autoregressive model of order 1, RCA(1), is given by

$$y_t = (\theta + b_t)y_{t-1} + e_t, \tag{2.3}$$

where $\theta^2 + \sigma_b^2 < 1$.

For the applications of RCA models, Nicholls and Quinn (1982) fitted the popular Canadian lynx data (see Campbell & Walker, 1977) to the RCA(2) model using the first 100 observations and forecasted the remaining 14 observations. Based on error sum of squares, the one-step-ahead forecasts of the transformed and untransformed lynx data using the fitted RCA(2) are better than the models proposed by Moran (1953) and Tong (1977). Longitudinal data on percentage of protein in cow's milk (Rahiala, 1999), NASDAQ and IBM index stock data (Wang & Gosh, 2002) have also been modeled using the RCA models. Another application of RCA is to estimate the time varying hedge ratios for corn and soybeans data (Bera et al., 1997). Recently, the RCA model has been used by Ghahramani and Thavaneswaran (2008) to estimate the volatility of the Japan index market, IBM stock return and dollar exchange rate data. They have combined the estimating functions of RCA and GARCH models to obtain the volatility estimates.

2.1.1 Properties of RCA

A number of authors have studied the properties of RCA models. Conlisk (1974, 1976) looked into the stability of RCA models, while Andel (1976) and Nicholls and Quinn (1981) discussed the problem of its second order stationarity. Furthermore, Feigin and Tweedie (1985) studied the stationarity, ergodicity and finiteness of moments for the RCA model.

There is a similarity between the RCA and the autoregressive (AR) models such that the RCA is obtained by adding a random additive perturbation to the ordinary AR coefficients. For clearer illustration of those similarities, we have plotted two simulated data from AR(1) ($\theta = 0.3$ and $\sigma_e^2 = 1.0$) and the RCA(1) ($\theta = 0.3$, $\sigma_e^2 = 1.0$, and $\sigma_b^2 = 0.3$) models and these are given in Figures 2.1 and 2.2 respectively.

From Figure 2.1, it can be seen that there is no significant fluctuation throughout the series of AR(1) process. The process appear to have no systematic change in mean and variance. On the other hand, the peak behavior of the RCA(1) in Figure 2.2 with the random coefficient b_t is significantly different from AR(1). It is clear that the additive random perturbation has caused

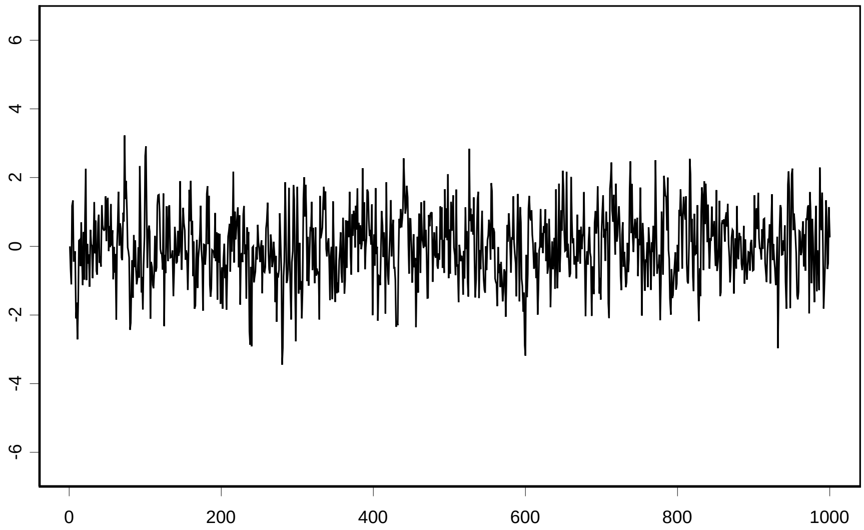

Figure 2.1: Time series plot of AR(1)

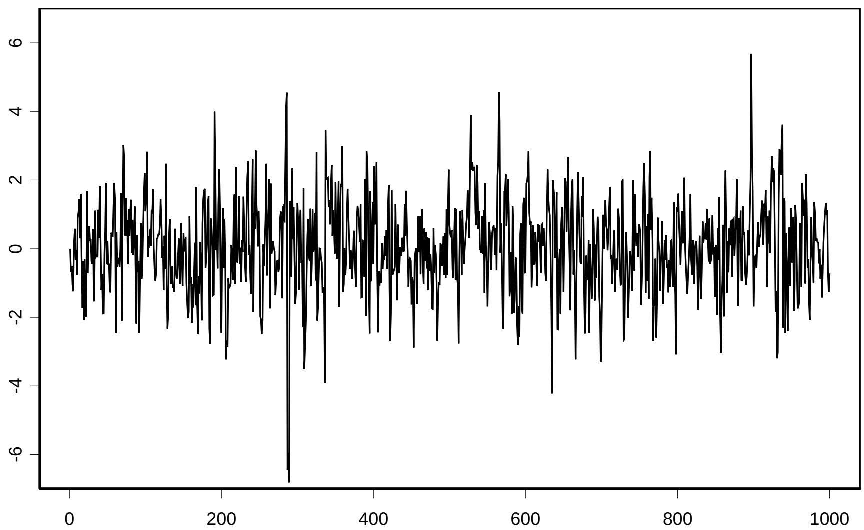

Figure 2.2: Time series plot of RCA(1)

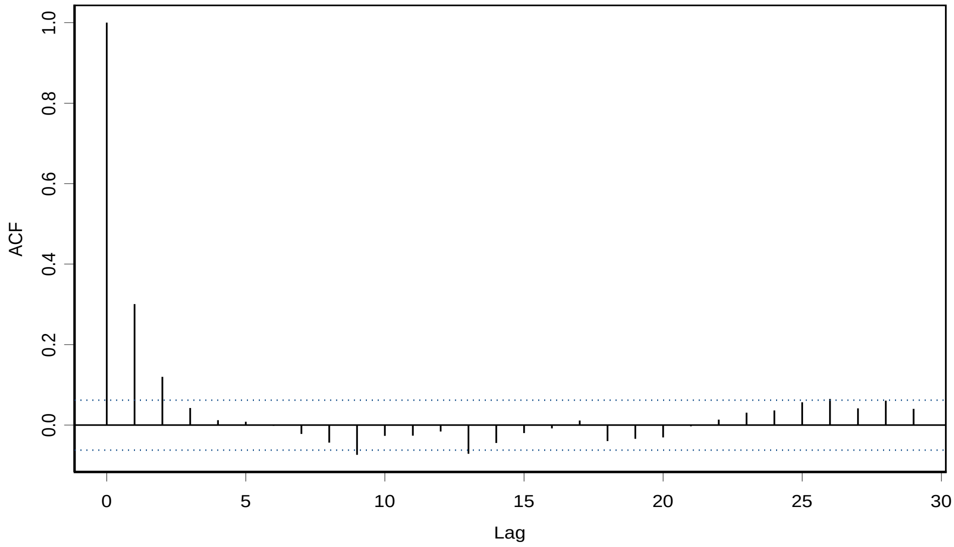

Figure 2.3: ACF plot of AR(1)

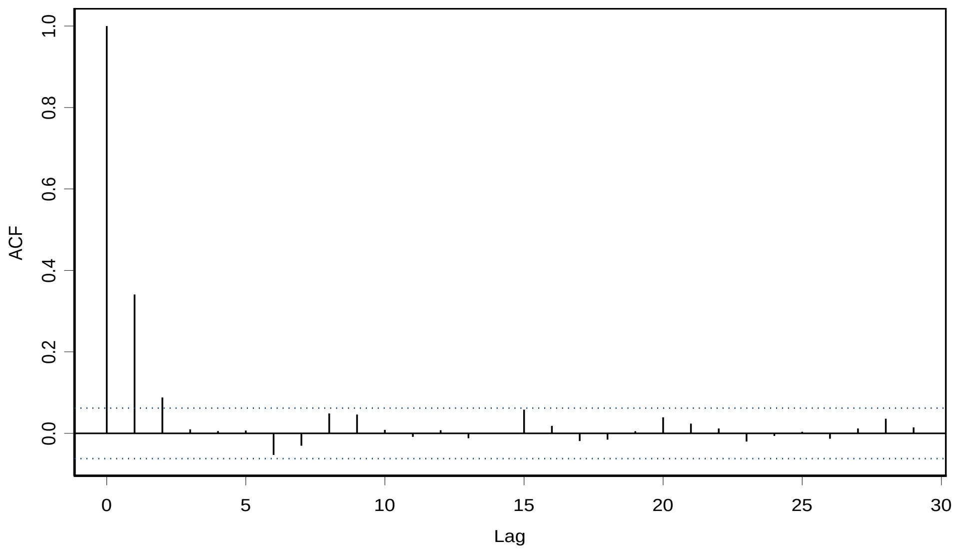

Figure 2.4: ACF plot of RCA(1)

some jumps in amplitude in the series. Figures 2.3 and 2.4 give corresponding autocorrelation function (ACF) plots of series in Figure 2.1 and 2.2. By looking at the spikes that quickly die out, both series appear to be stationary.

For further comparison of AR(1) and RCA(1) models, Figure 2.5 gives a time series plot for a larger value of σ_b^2=0.5. It is clear that as the variance of b_t gets larger, larger jumps in amplitude can be observed in the data set.

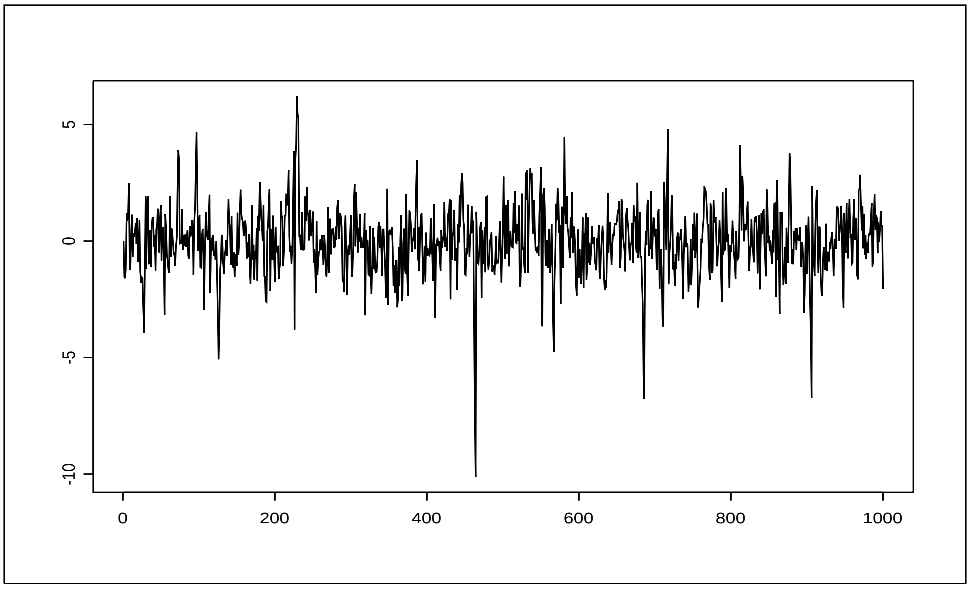

Figure 2.5: Time series plot of RCA(1) with larger value of σ_b^2

For the stationarity of RCA(1), it is noted earlier that the stationarity condition $\theta^2 + \sigma_b^2 < 1$ must be satisfied. To further study this condition, Figures 2.6 and 2.7 show the effects of θ and σ_b^2 on the stationarity of series.

Figure 2.6 is the plot for the RCA(1) process with $\theta = 0.8$ and $\sigma_b^2 = 0.3$. If θ is relatively large compared to σ_b^2, $\theta^2 + \sigma_b^2$ will be closer to unity (i.e., close to the boundary of stationarity). Thus, large values of $\{y_t\}$ are expected to arise and are generally associated with non-stationarity. If θ is not relatively large compared to σ_b^2, values of $\{y_t\}$ are expected to resemble realizations of constant coefficient autoregressive process. Figure 2.7 gives the ACF plot of the series in Figure 2.6. It can be seen that the spikes die out slowly. This is true since the stationarity condition is close to unity.

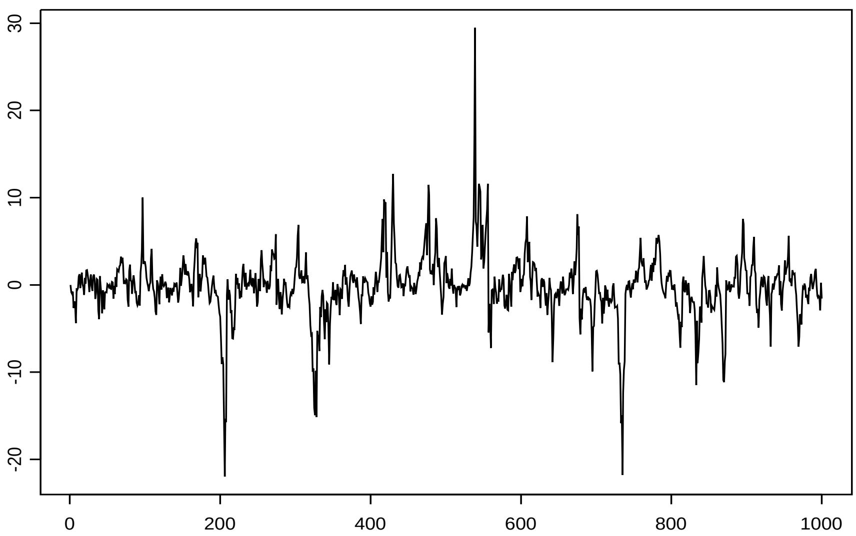

Figure 2.6: Time series plot of RCA(1) which closer to stationarity bound

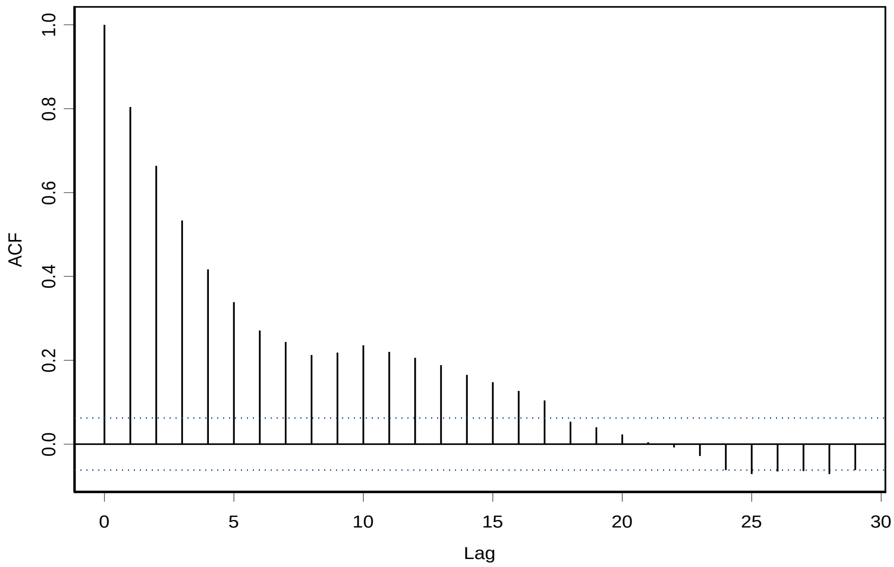

Figure 2.7: ACF plot of RCA(1) which closer to stationarity bound

2.1.2 Estimation Methods of RCA(1) Parameters

Several methods of estimating the RCA(1) parameters are available in the literature. The method of least squares (LS) proposed by Nicholls and Quinn (1980) is reviewed. From (2.3), let

$$u_t = e_t + b_t y_{t-1} = y_t - \theta y_{t-1}. \tag{2.4}$$

The LS estimate of $\hat{\theta}$ is obtained by minimizing

$$\sum_{t=2}^{n} u_t^2 = \sum_{t=2}^{n} (y_t - \theta y_{t-1})^2 \tag{2.5}$$

with respect to θ. Hence we have

$$\hat{\theta}_{LS} \quad = \quad \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}. \tag{2.6}$$

Note that $\hat{\theta}_{LS}$ depends only on the observations. Following Nicholls and Quinn (1980), by regressing u_t^2 on 1 and y_{t-1}^2, we can obtain the estimates of σ_b^2 and

σ_e^2. This is equivalent to minimizing $\sum_{t=2}^n [\hat{u}_t^2 - (\sigma_e^2 + \sigma_b^2 y_{t-1}^2)]^2$ with respect to σ_b^2 and σ_e^2. Thus the LS estimates of σ_b^2 and σ_e^2 are given by

$$\sigma^2_{b,LS} = \frac{\sum_{t=2}^n \hat{u}^2_{t,LS}(y^2_{t-1} - \bar{z})}{\sum_{t=2}^n (y^2_{t-1} - \bar{z})^2} \tag{2.7}$$

$$\sigma^2_{e,LS} = \frac{\sum_{t=2}^n \hat{u}^2_{t,LS}}{n-1} - \sigma^2_{b,LS}\bar{z} \tag{2.8}$$

respectively, where $\bar{z} = \sum_{t=2}^n \frac{y^2_{t-1}}{n-1}$ and $\hat{u}_{t,LS} = y_t - \hat{\theta}_{LS} y_{t-1}$.

Nicholls and Quinn (1982) showed that if the second moments of b_t and e_t are finite, then $\hat{\theta}_{LS}$ is a consistent estimator for θ. Further, under the finite fourth moments of $\{y_t\}$ (i.e., $\theta^4 + 6\theta\sigma_b^2 + 3\sigma_b^2 < 1$), $\sqrt{n}(\hat{\theta}_{LS} - \theta)$ converges in distribution to a normal random variable.

Another method available to estimate the RCA(1) parameters is based on the maximum likelihood criterion (see Nicholls and Quinn, 1981). By assuming the normality of b_t and e_t, one can write the likelihood function explicitly as follows

$$\begin{aligned} f_n(y_1, .., y_n | y_0) &= \prod_{t=1}^n f(y_t | y_{t-1}) \\ &= (2\pi)^{-n/2} \prod_{t=1}^n \{(\sigma_e^2 + \sigma_b^2 y_{t-1}^2)^{-1/2} exp[-\frac{(y_t - \theta y_{t-1})^2}{2(\sigma_e^2 + \sigma_b^2 y_{t-1}^2)}]\} \\ &= L_n(\theta, \sigma_e^2, \sigma_b^2). \end{aligned} \tag{2.9}$$

Taking the log-likelihood function, we have

$$\begin{aligned} logL_n(\theta,\sigma_e^2,\sigma_b^2) &= -\frac{n}{2}log(2\pi)-\frac{1}{2}\sum_{t=1}^{n}log(\sigma_e^2+\sigma_b^2y_{t-1}^2) \\ &\quad -\sum_{t=1}^{n}[\frac{(y_t-\theta y_{t-1})^2}{2(\sigma_e^2+\sigma_b^2y_{t-1}^2)}]. \end{aligned} \tag{2.10}$$

Let

$$\begin{aligned} \tilde{l}(\theta,\sigma_e^2,\sigma_b^2) &= -\frac{2}{n}log[L_n(\theta,\sigma_e^2,\sigma_b^2)]-log(2\pi) \\ &= n^{-1}\sum_{t=1}^{n}log(\sigma_e^2+\sigma_b^2y_{t-1}^2) \\ &\quad +n^{-1}\sum_{t=1}^{n}[\frac{(y_t-\theta y_{t-1})^2}{(\sigma_e^2+\sigma_b^2y_{t-1}^2)}]. \end{aligned} \tag{2.11}$$

We now minimize the monotone function $\tilde{l}(\theta,\sigma_e^2,\sigma_b^2)$ instead of maximization of likelihood function $L_n(\theta,\sigma_e^2,\sigma_b^2)$. By re-parameterizing $\tau=\frac{\sigma_b^2}{\sigma_e^2}$, we are able to minimize the monotone function above in terms of τ alone. Thus $\tilde{l}(\theta,\sigma_e^2,\sigma_b^2)$ is reduced to $\tilde{l}(\theta,\sigma_e^2,\tau)$, giving:

$$\begin{aligned} \tilde{l}(\theta,\sigma_e^2,\tau) &= log(\sigma_e^2)+n^{-1}\sum_{t=1}^{n}log(1+\tau y_{t-1}^2) \\ &\quad +(\sigma_e^2n)^{-1}\sum_{t=1}^{n}[\frac{(y_t-\theta y_{t-1})^2}{(1+\tau y_{t-1}^2)}]. \end{aligned} \tag{2.12}$$

Minimizing $\tilde{l}(\theta,\sigma_e^2,\tau)$ with respect to θ and σ_e^2, we have

$$\hat{\theta}(\tau)=\sum_{t=1}^{n}[\frac{y_ty_{t-1}}{1+\tau y_{t-1}^2}]\{\sum_{t=1}^{n}[\frac{y_t^2}{1+\tau y_{t-1}^2}]\}^{-1} \tag{2.13}$$

and

$$\hat{\sigma}_e^2(\tau)=n^{-1}\sum_{t=1}^{n}\{\frac{[y_t-\hat{\theta}(\tau)y_{t-1}]^2}{1+\tau y_{t-1}^2}\}. \tag{2.14}$$

Substituting equation (2.14) in (2.12), we now have the following function of τ:

$$\tilde{l}(\tau) = log(\hat{\sigma}_e^2(\tau))+n^{-1}\sum_{t=1}^{n}log(1+\tau y_{t-1}^2). \tag{2.15}$$

An estimate of τ can be obtained by minimizing the above function in (2.15). Thus the maximum likelihood estimates of θ, σ_e^2 and σ_b^2 are given by $\hat{\theta}_{MLE} = \theta(\hat{\tau})$, $\hat{\sigma}^2_{e,MLE} = \sigma_e^2(\hat{\tau})$, and $\hat{\sigma}^2_{b,MLE} = \hat{\sigma}_e^2(\hat{\tau})$ respectively. These estimates can be calculated using Newton Raphson algorithm given good choice of initial values. The common choice are the least squares estimates. Nicholls and Quinn (1981) had shown that under normality assumptions of b_t, e_t and second order stationary condition of $\theta^2 + \sigma_b^2$, $\hat{\theta}_{MLE}$ is a consistent estimator for θ. Moreover, if the fourth moments of b_t and e_t are finite, then $\sqrt{n}(\hat{\theta}_{MLE} - \theta)$ follows the central limit theorem.

On the other hand, Schick (1996) introduced a class of asymptotically normal estimators of θ indexed by a family of bounded measurable functions to estimate the RCA(1) parameters. Let Φ be the set of all bounded measurable functions ϕ such that $x\phi(x) > 0$ for $x \neq 0$ with mean and variance given by

$$E(\phi) = \frac{\sum_{j=1}^{n} \phi(x_{j-1})x_j}{\sum_{j=1}^{n} \phi(x_{j-1})x_{j-1}} \tag{2.16}$$

and

$$V(\phi) = \frac{E[\phi^2(x_0)w(x_0)]}{E[\phi(x_0)x_0]^2} \tag{2.17}$$

respectively, where $w(x) = \sigma_e^2 + \sigma_b^2 x^2$ and x_0 is the initial observation. He has shown that for every $\phi \in \Phi$, $\sqrt{n}(E(\phi) - \theta)$ converges in distribution to a normal random variable with mean 0 and variance $V(\phi)$. Furthermore, an asymptotically optimal estimator which possesses the smallest variance within this class of estimators is defined by taking

$$\phi(x) = \phi^*(x) = \frac{x}{1 + \rho x^2}, \quad \text{where } \rho = \frac{\sigma_b^2}{\sigma_e^2}. \tag{2.18}$$

Hence, Schicks estimator has the form

$$\hat{\theta}(\phi^*(x)) = (\sum_{t=1}^{n} \frac{x_{t-1}x_t}{\sigma_e^2 + \sigma_b^2 x_{t-1}^2})(\sum_{t=1}^{n} \frac{x_{t-1}^2}{\sigma_e^2 + \sigma_b^2 x_{t-1}^2})^{-1}. \tag{2.19}$$

Note that the estimate of θ using the maximum likelihood is asymptotically equivalent to that in (2.19). However, (2.19) does not require the finite fourth moments of e_t and b_t. Further, consistent estimators of ρ from a covariance assumption and the ergodic theorem have also been constructed in Schick (1996). It is shown that the asymptotic normality still holds.

Wang and Gosh (2002) obtained a Bayesian estimation of the RCA(1) parameter. The prior density which reflects prior beliefs about the unknown parameter is chosen to be N(μ_0, V_0) as non-informative prior for θ and inverse gamma (IG(a,b)) for σ_b^2 and σ_e^2. Then the joint posterior density of the parameters is given by

$$f(\psi|x_0, .., x_n) \propto L(\psi)p(\psi), \tag{2.20}$$

where ψ is the set of parameters to be estimated; i.e., $\psi = (\theta, \sigma_b^2, \sigma_e^2)^T$, $L(\psi)$ is the likelihood function and $p(\psi)$ is the prior density of ψ. Since it is not an easy task to find the normality constant of joint posterior densities and marginal density with respect to parameters, Markov Chain Monte Carlo (MCMC) method is employed. Gibbs sampler is used to obtain the dependent samples from the posterior distribution. Wang and Gosh (2002) had also derived the conditional density of one parameter given the others with observed data. Using arbitrary starting values $\theta^{(0)}$, $\sigma_b^{(0)}$ and $\sigma_e^{(0)}$ for θ, σ_b^2 and σ_e^2 respectively, the Gibbs sampling algorithm for RCA(1) that can be obtained from (2.20) is given by

1. $\theta^{(k)}$ from $f(\theta|\sigma_b^{2(k-1)}, \sigma_e^{2(k-1)})$

2. $\sigma_b^{2(k)}$ from $f(\sigma_b^2|\theta^k, \sigma_e^{2(k-1)})$

3. $\sigma_e^{2(k)}$ from $f(\sigma_e^2|\theta^k, \sigma_b^{2(k)})$,

where the superscripts (k) represent the respective value at the k^{th} iteration. Repeating the above sampling steps, the discrete-time Markov Chain can be obtained, whose stationary distribution is the joint posterior density of the parameter. An extensive simulation study was carried out in Wang and Gosh (2002) for weakly stationary and non stationary cases.

Another method that has been considered in estimating the model is the theory of estimating functions for stochastic processes. This method was originally proposed by Durbin (1960) and had been extended by Godambe (1985) with the optimality theorem for a certain class of estimating functions. Thavaneswaran and Abraham (1988) had applied Godambe's estimating functions theory to nonlinear time series estimation problems, and this was further extended by Thavaneswaran and Peiris (1996) to nonparametric estimation problems. A review of estimating functions is given in the next section for later reference.

2.2 Theory of Estimating Functions

The theory of estimating functions was first proposed by Durbin (1960). He considered the class of functions of the form

$$g(y, \theta) = T_1(y) + \theta T_2(y), \tag{2.21}$$

where $T_1(y)$ and $T_2(y)$ are functions of the data and $E[g(y, \theta)] = 0$.

Among all such unbiased estimating functions in the form of (2.21), $g_1(y, \theta)$ is the best unbiased linear estimating functions if

$$Var[(g_1(y, \theta))] \leq Var[(g(y, \theta))]. \tag{2.22}$$

Godambe (1960) extended the work of Durbin (1960) with the optimality criterion for a general class of estimating functions. Following Godambe (1985), any real function g of random variates $y_1, y_2, ..., y_i$ and the parameter θ is called a regular unbiased estimating function if

$$E_{i-1}[g\{y_1, ..., y_i; \theta(F)\}] = 0 \quad (F \in \mathfrak{F}), \tag{2.23}$$

where $\mathfrak{F}$ is a class of probability distributions F on R^n and $\theta = \theta(F)$ be a real parameter. E_{i-1} denotes the expectation holding the first $i-1$ values namely $y_1, ..., y_{i-1}$ fixed. Among all regular unbiased estimating function g, g^* is said to be optimum if

$$\frac{E_{i-1}[g*]}{[E_{i-1}(\frac{\partial g*}{\partial \theta})]^2} \leq \frac{E_{i-1}[g]}{[E_{i-1}(\frac{\partial g}{\partial \theta})]^2} \tag{2.24}$$

for all $F \in \mathfrak{F}$ at $g = g^*$. An estimate of θ can be obtained by solving the optimum estimating equations, $g^*(y_1, ..., y_i; \theta) = 0$.

Godambe (1985) extended his work on estimating functions to the stochastic processes. He restricted the estimating functions $g(.)$ in the linear form such that

$$g = \sum_{i=1}^{n} h_i a_{i-1} \quad g \in L, \tag{2.25}$$

where L is the class of estimating functions, h_i is a real function of $y_1, ..., y_i$ $(i = 1, ..., n)$ and parameter θ satisfying

$$E_{i-1}[h_i\{y_1, ..., y_i; \theta(F)\}] = 0 \;\; (F \in \mathfrak{F}), \tag{2.26}$$

where a_{i-1} is a function of the random variate $y_1, ..., y_{(i-1)}$ $(i = 1, ..., n)$ and parameter θ. Because of (2.26),

$$E(g) = E(\sum_{i=1}^{n} h_i a_{i-1}) = 0 \tag{2.27}$$

and hence (2.25) is an unbiased estimating functions.

Optimality Theorem

In the class L of unbiased estimating functions $g(.) = \sum_{i=1}^{n} h_i a_{i-1}$, the optimum function g^* is the one which minimizes $\frac{Ei-1[g]}{[Ei-1(\partial g/\partial \theta)]^2}$ and is given by

$$g^* = \sum_{i=1}^{n} h_i a_{i-1}^*, \tag{2.28}$$

where $a_{i-1}^* = \frac{E_{i-1}(\frac{\partial h_i}{\partial \theta})}{E_{i-1}(h_i^2)}$ •

Note that the optimality theorem only depends on the first two conditional moments of the distribution of h_i, i.e., $E(h_i)$ and $E(h_i^2)$. Proof of this result can be found in Godambe (1985).

Thavaneswaran and Abraham (1988) extended Godambe's (1985) optimality estimating functions theory to nonlinear time series estimation problems. Later Thavaneswaran and Peiris (1996) applied this theory to nonparametric estimation problems. Further extensions of the estimating functions are based on the least absolute deviation (LAD), generalized kernel smoothers and the smoothed least absolute deviation (SLAD) estimating function (see Thavaneswaran & Peiris, 2001, 2003, 2004).

To further comprehend the use of estimating functions, we illustrate the optimal estimating function theorem for two models below.

Example 1: AR(1) model

The AR(1) model is given by

$$y_t = \alpha y_{t-1} + e_t, \tag{2.29}$$

where $|\alpha| < 1$ and e_t is an independent and identically distributed random variable with mean zero and variance σ^2.

1. Let h_t be a function of random variate $y_1, y_2, ..., y_n$ and parameter α such that $h_t = y_t - \alpha y_{t-1}$ so that $E[h_t] = 0$.

2. Let $g(.)$ be the class of linear estimating functions in the form of

$$g(.) = \sum_{t=2}^{n} a_{i-1} h_i, \tag{2.30}$$

where a_i is a suitably chosen function. Now $E[g] = 0$, and consequently the condition of unbiased estimating functions is satisfied.

3. Using the optimality theorem from Godambe (1985), the corresponding optimal estimating function (EF) is

$$g^*_{opt}(\alpha) = \sum_{t=2}^{n} a^*_{t-1} h_t, \quad \text{where} \quad a^*_{t-1} = \frac{E_{i-1}[\frac{\partial h_t}{\partial \alpha}]}{E_{i-1}[h_t^2]}. \tag{2.31}$$

4. The EF estimate of $\hat{\alpha}$ can be easily obtained by solving $g^*_{opt}(\alpha) = 0$ and is given by

$$\hat{\alpha} = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}. \tag{2.32}$$

Note that the estimate in (2.32) is the standard conditional LS estimate of α.

Example 2: GARCH(1,1) model

The GARCH(1,1) model is given by

$$y_t | F_{t-1} \sim iid(0, k_t), \tag{2.33}$$

where F_{t-1} is the information set available up to time $t-1$ and

$$k_t = \omega + \varepsilon_{t-1}^2 + \beta k_{t-1} \tag{2.34}$$

where ω is a non-negative constant and ε is the error term.

1. In this case, define h_t to be $h_t = y_t^2 - k_t$ so that $E[h_t] = 0$

2. Let $g(.)$ be the class of linear estimating functions in the form of

$$g(.) = \sum_{t=2}^{n} a_{i-1} h_i, \tag{2.35}$$

so that $g(.)$ satisfies the unbiased estimating functions condition.

3. Using the optimality theorem, the corresponding optimal estimating functions (EF) for α and β are

$$g_{opt}^{\alpha*}(\alpha,\beta)=\sum_{t=2}^{n}a_{t-1}^{\alpha*}h_t \quad \text{where} \quad a_{t-1}^{\alpha*}=\frac{E_{i-1}[\frac{\partial h_t}{\partial\alpha}]}{E_{i-1}[h_t^2]} \tag{2.36}$$

and

$$g_{opt}^{\beta*}(\alpha,\beta)=\sum_{t=2}^{n}a_{t-1}^{\beta*}h_t \quad \text{where} \quad a_{t-1}^{\beta*}=\frac{E_{i-1}[\frac{\partial h_t}{\partial\beta}]}{E_{i-1}[h_t^2]} \tag{2.37}$$

4. The EF estimates $\hat{\alpha}$ and $\hat{\beta}$ are easily obtained by solving $g_{opt}^{\alpha*}(\alpha,\beta)=0$ and $g_{opt}^{\beta*}(\alpha,\beta)=0$ respectively.

In the next section, we review the problem of outliers in time series models.

2.3 Outliers

The term "outliers" refers to observations that is obviously depart from the rest of the data. Outliers are also known as 'contaminants', 'discordant observations' or 'extreme values'. Chatfield (1975) discussed the idea of outliers in time series with an example of product sale data. Further, Barnett and Lewis (1978) considered another two examples; the moisture content of Malaysian tobacco in eight hours and the percentages of monthly road accidents in the British Isles.

Outliers in data may occur due to various reasons such as data management errors and unexpected or unusual events (eg. disaster, sudden political or economic crisis). The presence of such outliers in time series may cause distortion in model specification, affects the parameter estimation and forecasting. These have been investigated by Abraham and Chuang (1989). They found out that a number of suspected outliers in time series data may result in large residuals, consequently affecting model specification and parameter estimation. Hogg (1979) pointed out that outliers may influence the least squares estimator by pulling the least squares "fit" towards them. Therefore, proper actions on possible occurrence of outliers are necessary. Next, we discuss the types of outliers that may arise in time series.

2.3.1 Types of Outliers

Four types of outliers are frequently found in time series literature. They are additive outlier (AO), innovational outlier (IO), level change (LC) and temporary change (TC). The most common type is AO which is deterministic in nature and most likely caused by an isolated incident such as recording error or a sudden disturbance. Among authors who had looked at the occurrence of AO in their study are Chang et al. (1988), Vogelsang (1999), and Battaglia and Orfei (2005).

Suppose that an AO occurs in time series $\{y_t\}$ at time $t = d$ and let y_d be the affected observation. Following Tsay (1986), the contaminated observation will differ from the original observations according to the following rule:

$$y_t^* = \begin{cases} y_t & \text{for } t \neq d \\ y_t + \omega & \text{for } t = d. \end{cases}$$

That is, the shock caused by an AO affect the observation at time $t = d$ only with magnitude of ω, while the rest remains unaffected.

IO is the other type of outliers often found in time series data. It represents an extraordinary shock at time $t = d$ and influences not only at the observation y_d, but also the subsequent observations $y_{d+1}, y_{d+2}, ...$ through the memory of dynamic system associated with the IO. If the process is stationary, then the outlier effects will die out exponentially. Battaglia and Orfei (2005) described the effect of an IO at time $t = d$ by

$$y_t^* = y_t + \eta_t, \tag{2.38}$$

where $\eta_t = e_t + \omega\delta_t$ and

$$\delta_t = \begin{cases} 0 & \text{for } t \neq d \\ 1 & \text{for } t = d. \end{cases}$$

Another type of outliers is the LC (see Box & Tiao, 1965). This outlier will cause a permanent change in the series after its impact. On the other hand, Tsay (1986) noticed, in some cases, the changes are not permanent but decay exponentially with a rate δ. Such outlier is defined as TC. The parameter δ is used to model the pace of the dynamically dampening effect of the TC. In this study, we consider the occurrence of AO and IO in RCA(1) processes.

Pena (1990) had investigated the effects of outliers in time series analysis. He noted that outliers may not necessarily influence the parameter estimates, but in general, affect the variance of the estimates. Martin (1980) discussed the effect on innovation variance estimates whilst Chen and Liu (1993) studied the impact of outliers on forecasts. Later, Chick (1994) showed that the order selection criteria such as Akaike information criteria or AIC is indirectly affected by outliers. This is because, the calculation of AIC depends on the estimate of innovation variance. Consequently, these will lead to incorrect model selection and forecasts.

2.3.2 Treatment of Outliers

Barnett and Lewis (1978) suggested four different approaches in handling outliers.

- The first approach is to accommodate outliers using a robust method. That is, the outlying observations are down-weighted through appropriate weight functions to reduce their influence.
- The second approach is by placing outliers within a homogenous probability model setting so that no observation will appear as outlier.
- Thirdly, enhancing the importance of outliers by setting up a mixture model to explain their existence.
- The last one is the rejection of outliers by using suitable method. This rejection approach was originally studied by Peirce (1852) and later extended by Wright (1884). Wright (1884) suggested that the best rule to

reject an outlier is when the residual has exceeded 3.37 times the standard deviation of the observations.

In this study, we use the idea of the last approach. A rejection rule is derived and the effects of outliers are measured. In the next chapter, we look at different estimation methods of the RCA(1) model. A new iterative method based on estimating functions theory is proposed.

Chapter 3

An Iterative Estimation Method Based on the Estimating Functions

A number of authors have proposed several approaches to further improve the RCA parameter estimation. For example, Chandra and Taniguchi (2001) used the estimating functions to estimate the coefficient parameter θ by first estimating the nuisance parameters using least squares and moment methods. Whilst Aue et al. (2006) proposed the quasi-maximum likelihood method by assuming the error term e_t and random coefficient b_t follow a joint normal distribution. In this chapter, we present a new iterative estimation method for RCA(1) model based on the estimating function theory. This method is expected to improve not only the estimation of parameter coefficient θ, but also the estimation of σ_b^2 and σ_e^2. We also compare the performance of this new iterative method with the least squares and the estimating function method through an extensive simulation study.

3.1 Optimal Estimation for RCA(1) Parameters using Estimating Functions

Let

$$h_t = e_t + b_t y_{t-1} = y_t - \theta y_{t-1}. \tag{3.1}$$

Clearly, h_t satisfies the unbiased estimating function (EF) condition in (2.26). The corresponding optimal (linear) EF for estimating θ is

$$g^*(\theta) = \sum_{t=2}^{n} a_{t-1}^* h_t,$$

where

$$a_{t-1}^* = \frac{E(\frac{\partial h_t}{\partial \theta}|F_{t-1})}{E(h_t^2|F_{t-1})}.$$

Now, it is easy to see that for the RCA(1) model in (2.3),

$$a_{t-1}^* = \frac{-y_{t-1}}{\{\sigma_e^2 + y_{t-1}^2\sigma_b^2\}}. \tag{3.2}$$

To estimate the parameter θ, we solve $g^*(\theta) = 0$ and the corresponding optimal estimate is given by

$$\theta_{EF}^* = \frac{\sum_{t=2}^{n} a_{t-1}^* y_t}{\sum_{t=2}^{n} a_{t-1}^* y_{t-1}}. \tag{3.3}$$

Thus, the EF estimate in (3.3) reduces to

$$\theta_{EF}^* = \sum_{t=2}^{n} \frac{y_t y_{t-1}}{\sigma_e^2 + \sigma_b^2 y_{t-1}^2} \Big/ \sum_{t=2}^{n} \frac{y_{t-1}^2}{\sigma_e^2 + \sigma_b^2 y_{t-1}^2}. \tag{3.4}$$

Chandra and Taniguchi (2001) had shown that θ^*_{EF} is asymptotically normal with the second moments of $\{b_t\}$ and $\{e_t\}$ are assumed to exist. Note that the estimation of θ using the EF method in (3.4) is weighted with σ_e^2 and σ_b^2 for each observation. Following Thavaneswaran and Peiris (1996), $\hat{\theta}_{EF}$ of equation (3.4) can be obtained by using the LS estimates $\hat{\sigma}^2_{e,LS}$ and $\hat{\sigma}^2_{b,LS}$, given by equations (2.7) and (2.8) respectively. Now, we suggest a new estimation method based on EF.

3.2 A New Iterative Method based on Estimating Functions

It is clear that the evaluation of (3.4) depends on (2.7) and (2.8). We therefore suggest the following efficient iterative algorithm to estimate θ. Let $\theta^{(k)}_{IT}$ be an estimate of θ based on n observations at the step k, $k = 1, 2, ...$

1. Take the initial values $\theta^{(0)}_{IT} = \hat{\theta}_{LS}$ from (2.6), $u^{(0)}_{t,IT}=\hat{u}_{t,LS}$ from (2.4) and $\sigma^{2(0)}_{b,IT}=\sigma^2_{b,LS}$ and $\sigma^{2(0)}_{e,IT}=\sigma^2_{e,LS}$ from (2.7) and (2.8) respectively.

2. Now obtain the following values for k=1,2,...:

$$\theta^{(k)}_{IT} = \sum_{t=2}^{n} \frac{y_t y_{t-1}}{\sigma^{2(k-1)}_{e,IT} + \sigma^{2(k-1)}_{b,IT} y^2_{t-1}} \Big/ \sum_{t=2}^{n} \frac{y^2_{t-1}}{\sigma^{2(k-1)}_{e,IT} + \sigma^{2(k-1)}_{b,IT} y^2_{t-1}} \tag{3.5}$$

$$u^{(k)}_{t,IT} = y_t - \theta^{(k)}_{IT} y_{t-1} \tag{3.6}$$

$$\sigma^{2(k)}_{b,IT} = \frac{\sum_{t=2}^{n} \hat{u}^{2(k)}_{t,IT}(y^2_{t-1} - \bar{z})}{\sum_{t=2}^{n}(y^2_{t-1} - \bar{z})^2} \tag{3.7}$$

$$\sigma^{2(k)}_{e,IT} = \frac{\sum_{t=2}^{n} \hat{u}^{2(k)}_{t,IT}}{n-1} - \sigma^{2(k)}_{b,IT}\bar{z} \tag{3.8}$$

where $\bar{z} = \frac{\sum_{t=2}^{n} y^2_{t-1}}{n-1}$.

3. The process in step two will continue until $\theta^{(k)}_{IT}$, $\sigma^{2(k)}_{b,IT}$ and $\sigma^{2(k)}_{e,IT}$ converge with a certain pre-specified tolerance.

The subsequent sections consider the model selection and diagnostic checks as a tool to verify the fitted model.

3.3 A Model Selection of RCA(1) Estimates

Suppose that we fit the RCA(1) model using LS, EF and IT methods. We use Akaike's information criterion (AIC) to select the best fitted model when different methods are used. The criterion is given by

$$AIC = -2lnL(\theta) + 2p, \tag{3.9}$$

where $lnL(\theta)$ is the log likelihood function and p is the number of parameters to be estimated.

$$\begin{aligned} lnL(\cdot) &= -\frac{n}{2}log(2\pi) - \frac{1}{2}\sum_{t=1}^{n} log(\sigma_e^2 + \sigma_b^2 y_{t-1}^2) \\ &\quad - \sum_{t=1}^{n}[\frac{(y_t - \theta y_{t-1})^2}{2(\sigma_e^2 + \sigma_b^2 y_{t-1}^2)}]. \end{aligned} \tag{3.10}$$

Thus the AIC is

$$\begin{aligned} AIC^{(A)} &= n\ ln(2\pi) + \sum_{t=1}^{n}\{ln(\sigma_e^{2(A)} + \sigma_b^{2(A)} y_{t-1}^2)\} \\ &\quad + \sum_{t=1}^{n}\{\frac{(y_t - \theta^{(A)} y_{t-1})^2}{(\sigma_e^{2(A)} + \sigma_b^{2(A)} y_{t-1}^2)}\} + 2p, \end{aligned} \tag{3.11}$$

where A is either LS, EF or IT.

3.4 Diagnostic Checks

To check the adequacy of the fitted RCA(1) model, a diagnostic check based on the autocorrelation can be used. Since $E[u_t u_{t-1}|F_{t-1}] = 0$ and $E[u_t^2|F_{t-1}] = \sigma_e^2 + \sigma_b^2 y_{t-1}^2 = \gamma_t$, then $E[\varepsilon_t \varepsilon_{t-1}|F_{t-1}] = 0$ where $\varepsilon_t = \gamma_t^{-1/2} u_t$. This implies that diagnostic checks may be based on the autocorrelation of the $\hat{\varepsilon}_t = \hat{\gamma}_t^{-1/2}\hat{u}_t$ where $\hat{u}_t = y_t - \hat{\theta} y_{t-1}$ and $\hat{\gamma}_t = \hat{\sigma}_e^2 + \hat{\sigma}_b^2 y_{t-1}^2$. The residual autocorrelation coefficients for the RCA(1) model at k^{th} lag are given by

$$r(k) = \frac{\sum_{t=1}^{n} \varepsilon_t \varepsilon_{t-k}}{\sum_{t=1}^{n} \varepsilon_t^2}. \tag{3.12}$$

See Nicholls (1986) for details.

From Heyde and Hannan (1972), $\hat{r}(k) \to r(k)$ and $n^{1/2}[\hat{r}(k) - r(k)]$ will converge to a normal distribution. Consequently, the $\hat{r}(k)$ can be used to form an analogue of the Box-Pierce test

$$Q = n \sum_{k=1}^{j} \hat{r}^2(k), \tag{3.13}$$

where n is the sample size and j is a fixed maximum number of lags. Typically, j should be between 10 to 20.

The above Box-Pierce test is used to confirm the validity of the fitted model, i.e., $\mathrm{H}_0 : y_t \sim \mathrm{RCA}(1)$. If residuals of the model are white noise, then the Box-Pierce statistics is distributed approximately to Chi-square distribution with $j-m$ degree of freedom, where m is the number of fitted parameters. This Box-Pierce test is appropriate for large samples. However, for small samples, Nicholls (1986) noted that the Box-Pierce test may lead to a test statistic for which the true significance level may be much lower than that given by the Chi-square distribution.

The next section compares the performance of the fitted RCA(1) model by using LS, EF and IT methods via simulation study.

3.5 A Simulation Study

This section reports a simulation study for the RCA(1) parameter estimation given in (2.3). Below are the steps taken to look at the performance of three methods considered in fitting the RCA(1) model:

1. We generate the RCA(1) series of length $n = 50$ and $n = 500$ with known parameter values of θ, σ_b^2 and σ_e^2. We assume that e_t and b_t follow a normal distribution with mean 0 and variance 1 and σ_b^2 respectively. The initial value y_0 is chosen to be 0 and the first 200 values from the series are ignored to remove the initial value effect.

2. We then obtain the estimates of θ, σ_b^2 and σ_e^2 using LS, EF and IT. This estimation is repeated s times.

3. Let $\boldsymbol{\xi}$ $=(\theta, \sigma_b^2, \sigma_e^2)$, $\hat{\xi}_i^{(j)}$ be an estimate of ξ_i at step $j = 1, ..., s$ and the mean of ξ_i, $\bar{\xi}_i = \frac{1}{s}\sum_{j=1}^{s} \hat{\xi}_i^{(j)}$. The following calculations for each parameter are obtained:

   - Bias $=$ $\bar{\xi}_i - \xi_i$
   - Root mean square error (RMSE):
   $$\sqrt{\frac{1}{s}\sum_{j=1}^{s}[\hat{\xi}_i^{(j)} - \xi_i]^2} \tag{3.14}$$
   - Standard error (SE):
   $$\sqrt{\frac{1}{s-1}\sum_{j=1}^{s}[\hat{\xi}_i^{(j)} - \bar{\xi}_i]^2}. \tag{3.15}$$

4. Throughout our simulation, we fixed the tolerance of IT method to be 10^{-6} and number of simulation to be $s = 1,000$.

Table 3.1 gives the simulation results for the true parameter values $\theta = -0.3$ and $\sigma_b^2 = 0.25$. In Table 3.2, we use the true parameter values $\theta = 0.5$ and $\sigma_b^2 = 0.25$; Table 3.3 uses the true parameter values $\theta = 0.7$ and $\sigma_b^2 = 0.16$. In

Table 3.1: Parameter estimation for true parameter values θ=-0.3 and σ_b^2=0.25

	$n=50$			$n=500$		
Statistics	LS	EF	IT	LS	EF	IT
Bias θ	0.02497	0.01187	0.01011	0.00137	-0.00144	-0.00141
Bias σ_b^2	-0.07264	-0.07264	-0.05010	-0.05839	-0.05839	-0.05019
Bias σ_e^2	0.08917	0.08917	0.05817	0.07297	0.07297	0.06140
RMSE θ	0.16178	0.15937	0.16063	0.08692	0.07951	0.07946
RMSE σ_b^2	0.15198	0.15198	0.16276	0.12332	0.12332	0.12586
RMSE σ_e^2	0.32189	0.32189	0.32097	0.18580	0.18580	0.18490
SE θ	0.15992	0.15900	0.16039	0.08695	0.07954	0.07949
SE σ_b^2	0.13356	0.13356	0.15493	0.10867	0.10867	0.11548
SE σ_e^2	0.30945	0.30945	0.31582	0.17096	0.17096	0.17450

all cases, we use $\sigma_e^2 = 1.0$. Each table gives the bias, RMSE and SE for LS, EF and IT estimators for two sample sizes $n = 50$ and $n = 500$. The first three rows of Table 3.1 give the bias for each parameter θ, σ_b^2 and σ_e^2 using LS, EF and IT methods. It is clear that the bias of each parameter using IT is smaller than that of LS and EF. As for the RMSE and SE, the performance of LS, EF and IT are close to each other, with their differences range from 0.03 to 0.0001. When we increase the sample size to $n = 500$, the bias, RMSE and SE for each parameter (using LS, EF and IT) have improved. Similar results are observed in Tables 3.2 and 3.3

Table 3.2: Parameter estimation for true parameter values θ=0.5 and σ_b^2=0.25

	$n = 50$			$n = 500$		
Statistics	LS	EF	IT	LS	EF	IT
Bias θ	-0.04315	-0.01956	-0.01580	-0.01060	-0.00035	0.00009
Bias σ_b^2	-0.08633	-0.08633	-0.06025	-0.06222	-0.06222	-0.05201
Bias σ_e^2	0.14511	0.14511	0.10286	0.09100	0.09100	0.07404
RMSE θ	0.15717	0.15016	0.15079	0.08278	0.07504	0.07522
RMSE σ_b^2	0.15686	0.15686	0.17362	0.11924	0.11924	0.12368
RMSE σ_e^2	0.39137	0.39137	0.39319	0.22345	0.22345	0.22633
SE θ	0.15120	0.14895	0.15004	0.08214	0.07508	0.07525
SE σ_b^2	0.13103	0.13103	0.16291	0.10177	0.10177	0.11227
SE σ_e^2	0.36366	0.36366	0.37968	0.20418	0.20418	0.21399

Table 3.3: Parameter estimation for true parameter values θ=0.7 and σ_b^2=0.16

	$n = 50$			$n = 500$		
Statistics	LS	EF	IT	LS	EF	IT
Bias θ	-0.04055	-0.01455	-0.01149	-0.01621	-0.00501	-0.00447
Bias σ_b^2	-0.04168	-0.04168	-0.02342	-0.03343	-0.03343	-0.02616
Bias σ_e^2	0.10416	0.10416	0.06231	0.07889	0.07889	0.06215
RMSE θ	0.12892	0.12309	0.12363	0.06995	0.06412	0.06431
RMSE σ_b^2	0.10385	0.10385	0.11925	0.08580	0.08580	0.09120
RMSE σ_e^2	0.35141	0.35141	0.35461	0.22290	0.22290	0.22983
SE θ	0.12243	0.12229	0.12316	0.06808	0.06395	0.06419
SE σ_b^2	0.09516	0.09516	0.11698	0.07906	0.07906	0.08741
SE σ_e^2	0.33579	0.33579	0.3492	0.20857	0.20857	0.22138

In addition, it is our interest to find the best method for fitting the RCA(1) model (whether using LS, EF or IT). For that, we consider a similar set of parameter values as in Table 3.2 with $n = 500, \theta = 0.7$, and $\sigma_b^2 = 0.16$. Figures

3.1, 3.2 and 3.3 give the plot of simulated time series, autocorrelation function (ACF) and partial autocorrelation function (PACF) respectively. From the three figures, the series seems to be stationary and may be suitable to be fitted using the AR(1) or RCA(1) model. We then calculate the AIC values for the fitted AR(1) using maximum likelihood (ML), and RCA(1) using LS, EF and IT. Their AIC values are tabulated in Table 3.4. It can be seen that the AIC values of the RCA(1) model using LS, EF and IT are smaller than that of the AR(1) model. Furthermore, the RCA(1) model using IT has smaller AIC value compared to LS and EF.

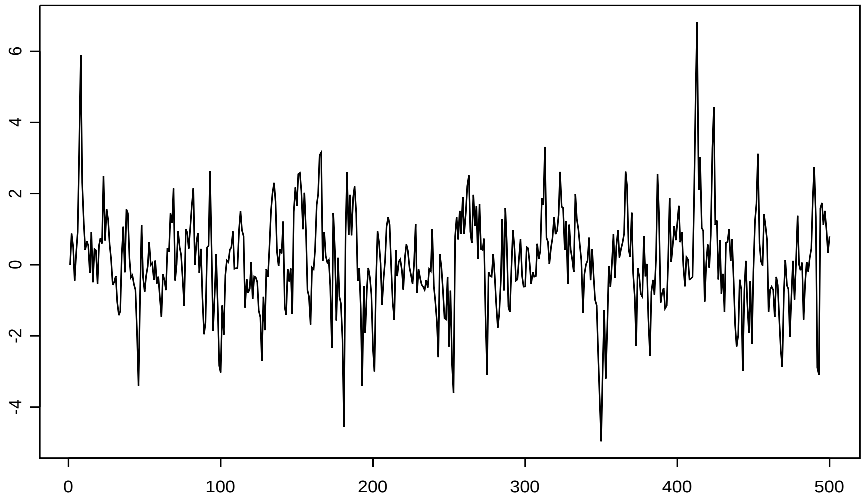

Figure 3.1: Time series plot of true parameter values θ=0.7 and σ_b^2=0.16

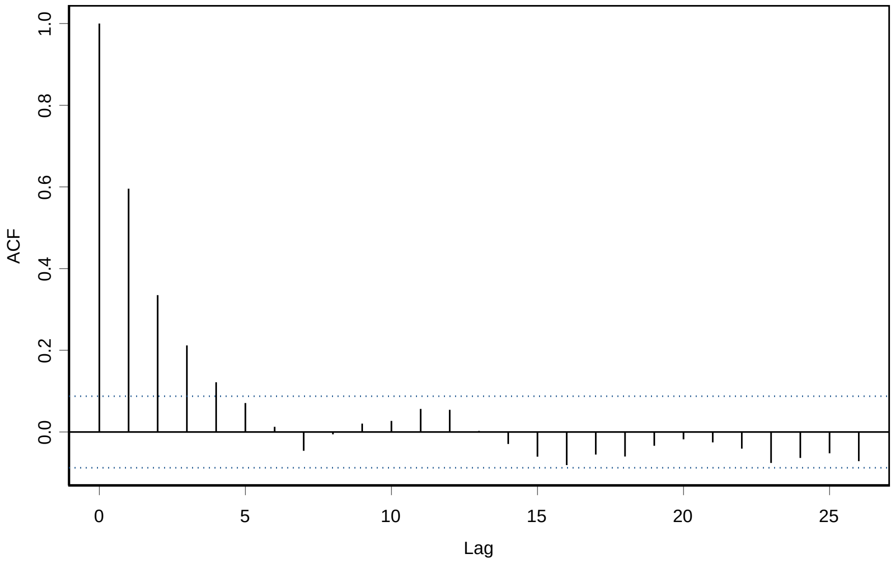

Figure 3.2: ACF plot of true parameter values θ=0.7 and σ_b^2=0.16

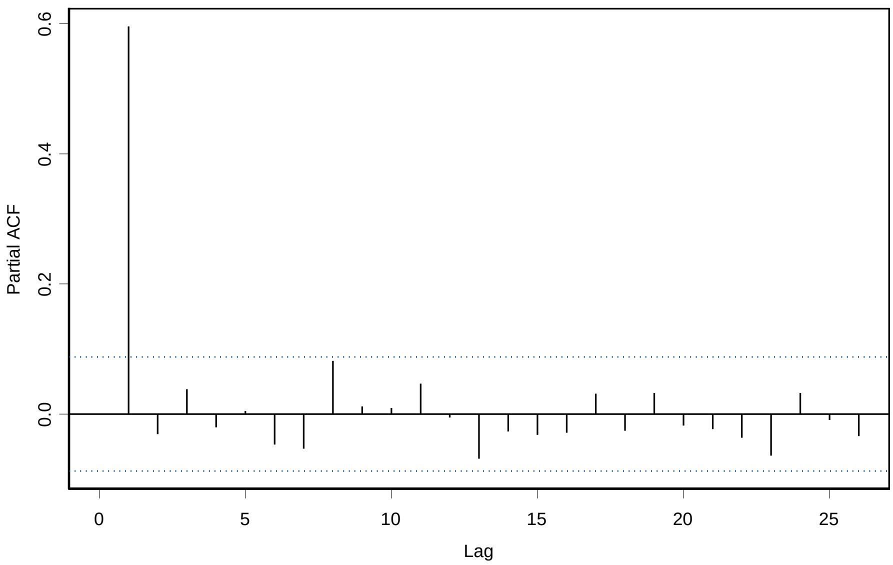

Figure 3.3: PACF plot of true parameter values θ=0.7 and σ_b^2=0.16

Table 3.4: AIC values using different fitted models and methods for simulated data

Model	AIC
AR(1) using ML	1514.72
RCA(1) using LS	1470.20
RCA(1) using EF	1468.50
RCA(1) using IT	1466.98

3.6 Application to a Real Data Set

In this section, we consider a set of differenced Indian consumer price index (CPI) data obtained from DataStream (http://www.library.uitm.edu.my/index.php?option=com_content&task=view&id=444&Itemid=557). The data consists of quarterly interest rate between the years of 1990 to 2006. There are n=67 observations altogether.

Figures 3.4, 3.5 and 3.6 give the time series, ACF and PACF plots of the quarterly Indian CPI data respectively. It can be seen that the series exhibits occasional random spikes and stable around the mean value, with possible outliers at $t = 3$ to $t = 10$. Further, the ACF plot shows that the series is stationary. Due to high spikes at the second and third lags though not significant enough, the PACF plot indicates that the series might be suitable for an AR(2) and AR(3). We also suggest an ARMA(2,2) as over fitting. To determine the best model to fit this Indian CPI data, we calculate the AIC values for all possible models and the RCA(1) model. Table 3.5 gives the AIC

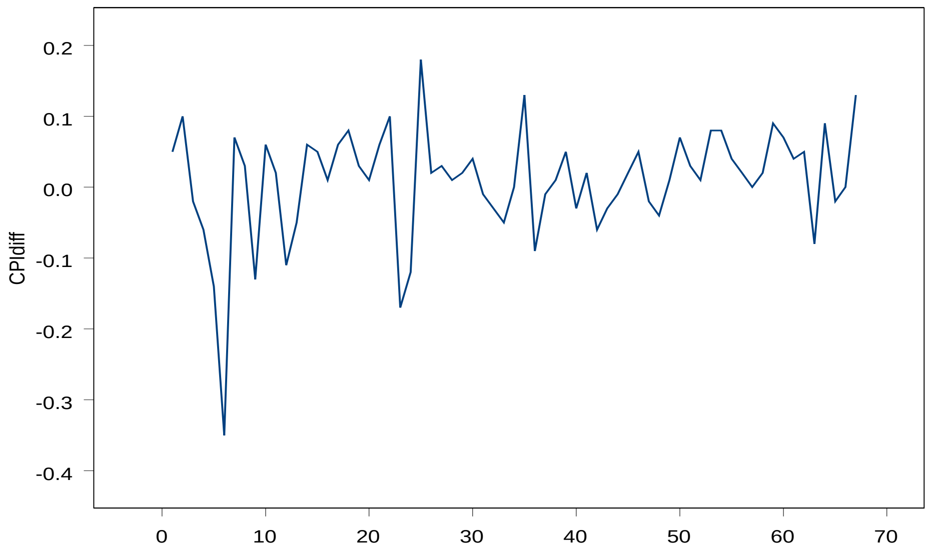

Figure 3.4: Time series plot of differenced Indian CPI data

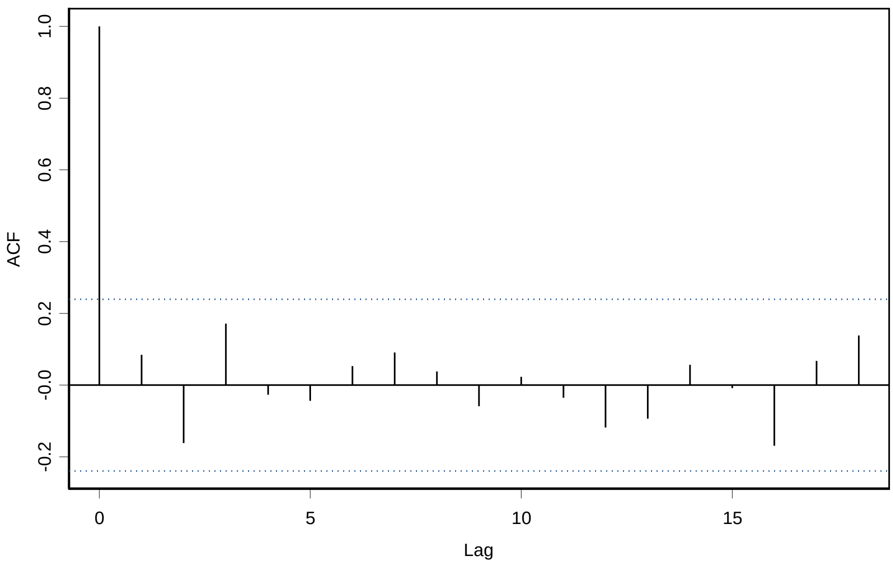

Figure 3.5: ACF plot of differenced Indian CPI data

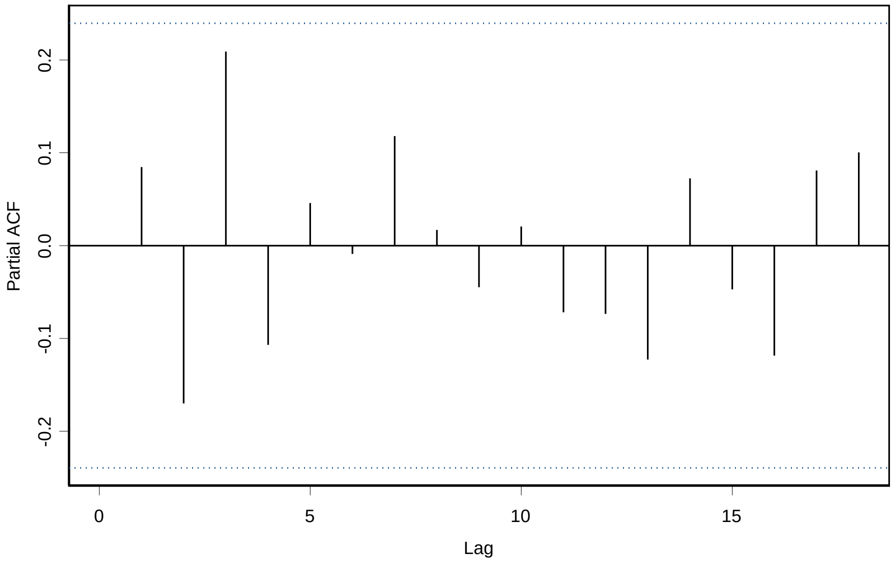

Figure 3.6: PACF plot of differenced Indian CPI data

Table 3.5: AIC values using different fitted models and methods for an Indian CPI data

Models	AIC
AR(2) using ML	-143.24
AR(3) using ML	-141.32
ARMA(2,2) using ML	-145.20
RCA(1) using LS	-146.80
RCA(1) using EF	-147.05
RCA(1) using IT	-148.32

values of an AR(2), AR(3) and ARMA(2,2) using ML method and RCA(1) using LS, EF and IT methods.

Based on the AIC values in Table 3.5, it is clear that the data is better fitted using RCA(1) compared to AR(2), AR(3) and ARMA(2,2). Other models have also been investigated but give poorer result. Further, the AIC values for RCA(1) using IT is smaller than that using LS and EF. This indicates that fitting the RCA(1) model using the IT is better compared to LS and EF.

To check the adequacy of the selected model, we produce three different residual plots obtained from the fitted RCA(1) model for LS, EF and IT methods as given in Figures 3.7 to 3.9 respectively. The residual plot is useful for diagnostic checking and spotting possible serial correlations, non-constant variance and outliers. It can be seen that the plots do not exhibit observable structure or serial correlation which shows any dependency. Possible outlier between time $t = 3$ to $t = 10$ can be observed. These plots suggest that the residuals are white noises if the outlier is removed.

We also perform the Box-Pierce test to confirm the validity of the fitted model, that is, by taking $\mathrm{H}_0 : y_t \sim \mathrm{RCA}(1)$. The critical value of the $\chi^2(20)$ is 27.58. The values of Box-Pierce statistics are given in Table 3.6.

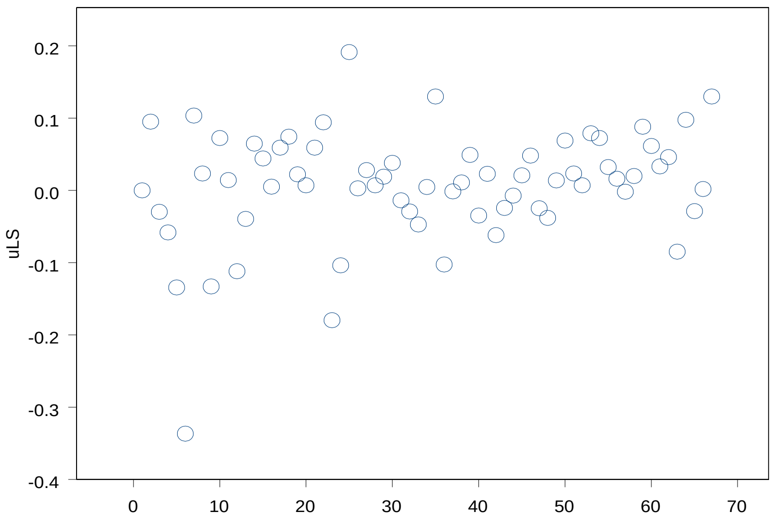

Figure 3.7: Residual plot of the fitted RCA(1) using LS method

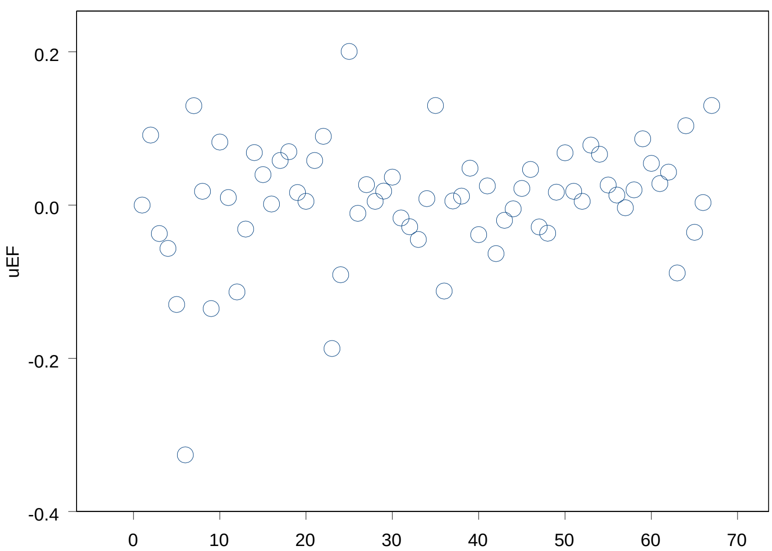

Figure 3.8: Residual plot of the fitted RCA(1) using EF method

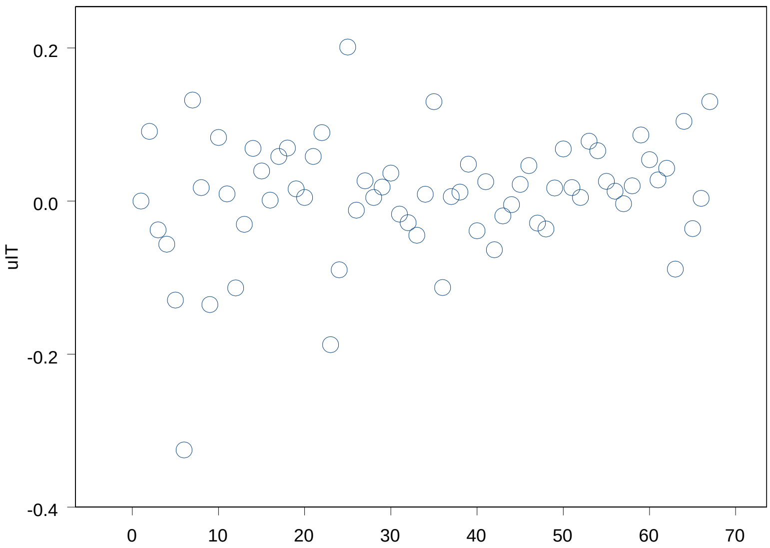

Figure 3.9: Residuals plot of the fitted RCA(1) using IT method

Table 3.6: Box-Pierce test for an Indian CPI data

Method	test statistics
LS	11.06
EF	11.74
IT	11.67

The Box-Pierce statistics confirm that all three methods fit the RCA(1) model satisfactorily and can be considered for further analysis, such as the detection of outliers. The parameter estimation of an RCA(1) using LS, EF and IT methods for an Indian CPI data are given in Table 3.7.

Table 3.7: Parameter estimation using different fitted methods for an Indian CPI data

Models	$\hat{\theta}$	$\hat{\sigma}_b^2$	$\hat{\sigma}_e^2$
RCA(1) using LS	0.0955	0.1704	0.0052
RCA(1) using EF	0.1709	0.1704	0.0052
RCA(1) using IT	0.1771	0.2138	0.0050

The next chapter will investigate the effects of outliers in RCA modeling.

Chapter 4

An Investigation of Outliers in RCA(1) Modeling

This chapter concerns with the understanding of outliers in RCA(1) modeling. This is an extension of the Chang et al. (1988) work on outliers in ARMA(p,q) models. The effects of additive outlier (AO) and innovational outlier (IO) are investigated and shown graphically.

An outlier free RCA(1) model is given by

$$y_t = (\theta + b_t)y_{t-1} + e_t, \tag{4.1}$$

where θ and b_t are as defined in Chapter 2 and satisfies the stationary condition $\theta^2 + \sigma_b^2 < 1$.

Let

$$u_t = e_t + b_t y_{t-1} = y_t - \theta y_{t-1}. \tag{4.2}$$

It is clear that

$$\begin{aligned} E[u_t|F_{t-1}] &= E[e_t + b_t y_{t-1}|F_{t-1}] \\ &= E[e_t|F_{t-1}] + y_{t-1}E[b_t|F_{t-1}] \\ &= 0 \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} E[u_t^2|F_{t-1}] &= E[e_t^2 + b_t^2 y_{t-1}^2 + 2e_t b_t y_{t-1}|F_{t-1}] \\ &= E[e_t^2] + E[b_t^2]y_{t-1}^2 + 2E[e_t b_t]y_{t-1} \\ &= \sigma_e^2 + \sigma_b^2 y_{t-1}^2 \\ &= Var[u_t], \end{aligned} \tag{4.4}$$

where F_{t-1} is the information set available up to time $t-1$.

Let y_t^* and u_t^* be the contaminated observation and residual at time t respectively, when an outlier exists in the data set. In the next Section, we formulate the effects of AO and IO on both observations and residuals.

4.1 Nature of Additive Outlier Effects in RCA(1) Model

From the definition of AO, observations with the presence of AO, y_t^*, will differ from the original observations according to the following rule:

$$y_t^* = \begin{cases} y_t & \text{for } t \neq d \\ y_t + \omega & \text{for } t = d. \end{cases}$$

The shock caused by AO with magnitude ω affect the observation at time $t = d$ only, while the rest remains unchanged.

As for the effect of AO on residuals, there are no changes for $t < d$. For $t = d, d+1, d+2, ...$ the effects are described as follows:

$$
\begin{aligned}
u_d^* &= y_d^* - \theta y_{d-1}^* = y_d + \omega - \theta y_{d-1} = u_d + \omega \\
u_{d+1}^* &= y_{d+1}^* - \theta y_d^* = y_{d+1} - \theta(y_d + \omega) = u_{d+1} - \theta\omega \\
u_{d+2}^* &= y_{d+2}^* - \theta y_{d+1}^* = y_{d+2} - \theta y_{d+1} = u_{d+2} \\
&\vdots \\
u_{d+k}^* &= u_{d+k}. \qquad (4.5)
\end{aligned}
$$

It can be summarized as

$$
u_{d+k}^* = \begin{cases} u_{d+k} + \omega & \text{for } k = 0 \\ u_{d+k} - \theta\omega & \text{for } k = 1 \\ u_{d+k} & \text{for } k = 2, 3, ... \end{cases}
$$

To illustrate the effects of AO on observations and residuals, we generate two series of contaminated and uncontaminated RCA(1) process and these are shown graphically. A set of data is generated using S-Plus package with sample size $n = 30$, $\theta = 0.5$ and $\sigma_b^2 = 0.3$ assuming e_t follows a standard normal distribution. Next, we add an AO with a magnitude ω=5 at time $t = 15$ into the series. The plots of contaminated and uncontaminated series for observations and residuals are shown in Figures 4.1 and 4.2 respectively. The plot of the simulated series without AO is represented by solid line, whereas the dash line represent the contaminated AO series.

Figures 4.1 and 4.2 illustrate the effects of AO on observations and residuals respectively. From Figure 4.1, it can be seen that there is a sudden shock at time $t = 15$ only, corresponding to the AO magnitude $\omega = 5$. The rest of the observations are left unaffected. For the effect on residuals as shown in Figure 4.2, the residual at time t=15 is changed according to the magnitude of AO while a number of subsequent residuals are also altered. From the formulation, it is expected that the fluctuation of the residual at $t = d + 1$ will be higher if larger coefficient of θ is used.

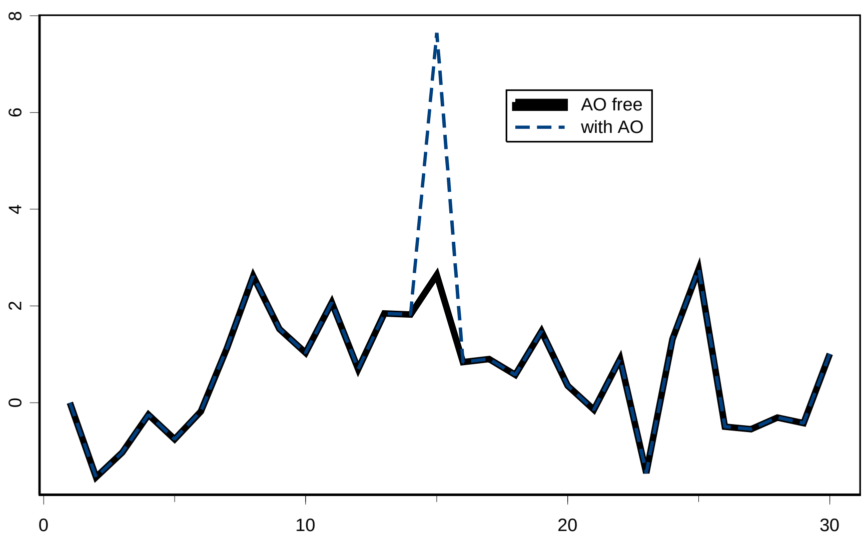

Figure 4.1: AO effect on observations

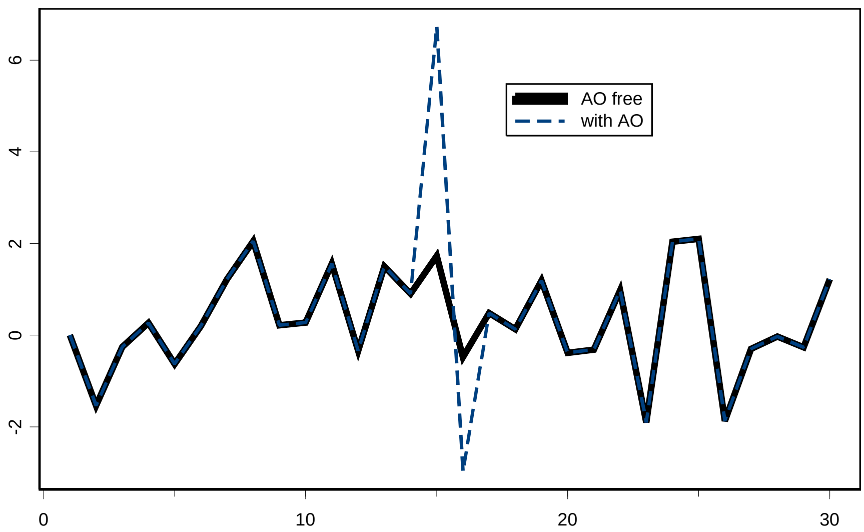

Figure 4.2: AO effect on residuals

4.2 Nature of Innovational Outlier Effects in RCA(1) Model

Balke and Fomby (1994) pointed out that IO appears in many real data sets, especially in data with high frequency. Battaglia and Orfei (2005) described the effect of IO at time $t = d$ on the residuals as

$$u_t^* = \begin{cases} u_t & \text{for } t \neq d \\ u_t + \omega & \text{for } t = d. \end{cases}$$

The IO will affects the observations y_t^* at $t = d,\ d+1,\ d+2, ..$ onwards as described below:

$$\begin{aligned} y_d^* &= \theta y_{d-1}^* + u_d^* \\ &= \theta y_{d-1} + (u_d + \omega) \\ &= y_d + \omega \end{aligned} \tag{4.6}$$

$$\begin{aligned} y_{d+1}^* &= \theta y_d^* + u_{d+1}^* \\ &= \theta(y_d + \omega) + u_{d+1} \\ &= y_{d+1} + \omega\theta \end{aligned} \tag{4.7}$$

$$\begin{aligned} y_{d+2}^* &= \theta y_{d+1}^* + u_{d+2}^* \\ &= \theta[y_{d+1} + \omega\theta] + u_{d+2} \\ &= y_{d+2} + \omega\theta^2 \end{aligned} \tag{4.8}$$

Thus we can write y_{d+k}^* as

$$y_{d+k}^* = y_{d+k} + \theta^k\omega \quad \text{for} \;\; k = 0, 1, 2, ..., n - d. \tag{4.9}$$

Now, it is clear that the IO influences not only the residual u_d, but also the observations $y_d, y_{d+1}, y_{d+2}, ...$. The IO effects on observations and residuals are illustrated in Figures 4.3 and 4.4 respectively. Further, the IO effects will eventually die out exponentially. For both cases of AO and IO, when ω is negative, the original observation at the time where outlier occurs will decrease corresponding to the magnitude ω. For example, the observation with IO of $\omega = -6$ is down by 6 units and the effects will eventually die out.

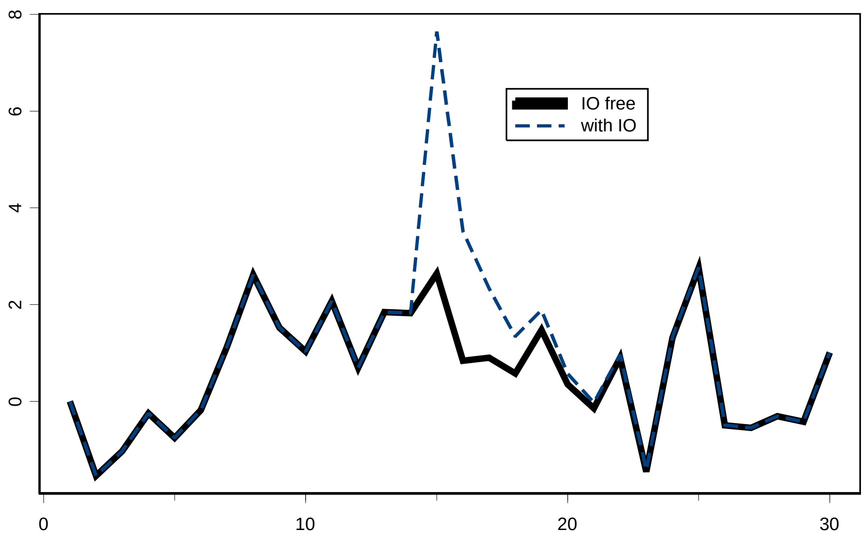

Figure 4.3: IO effect on observations

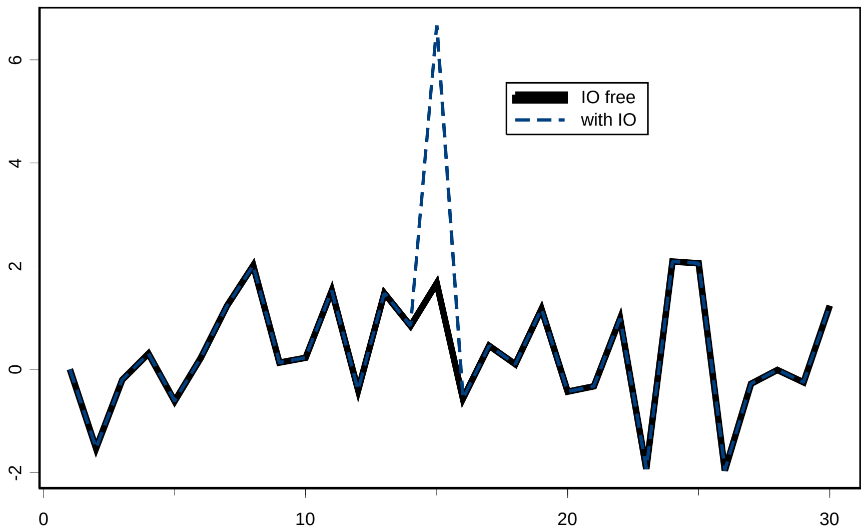

Figure 4.4: IO effect on residuals

The next section will investigate the effects of AO and IO on the parameter estimates via simulation study.

4.3 A Simulation Study

A simulation study in this section is designed to show the effects of AO and IO on the LS, EF and IT estimates. We provide detailed steps to investigate the robustness property of LS, EF and IT when AO occurs as follows:

1. We generate the RCA(1) series of length $n = 500$ with known parameter values of θ, σ_b^2 and σ_e^2. We assume that e_t and b_t follow a normal distribution with mean 0 and variance 1 and σ_b^2 respectively. The initial value y_0 is chosen to be 0 and the first 200 values from the series are ignored to remove the initial value effect.

2. We then obtain the estimates of θ_{free}, $\sigma^2_{b,free}$ and $\sigma^2_{e,free}$ using LS, EF and IT. The bias for each parameter is calculated. The subscript "*free*" here stands for outlier free data set.

3. This process is repeated $s = 1,000$ times. The bias for each parameter of the outlier free series using LS, EF and IT is calculated. They are denoted by Bias^{LS}_{free}, Bias^{EF}_{free} and Bias^{IT}_{free} respectively.

4. Using the original series generated in step 1, we introduce AO with magnitude ω at time $t = 250$. Now we have the contaminated AO series.

5. With the contaminated AO series, we estimate the parameters of θ_{AO}, $\sigma^2_{b,AO}$ and $\sigma^2_{e,AO}$ using LS, EF and IT methods.

6. This process is repeated $s = 1,000$ times. The bias for each parameter of the contaminated AO series using LS, EF and IT methods is calculated. They are denoted by Bias_{AO}^{LS}, Bias_{AO}^{EF} and Bias_{AO}^{IT} respectively.

Similar steps as above are taken for IO. The next two subsequent subsections report the results for AO and IO cases.

4.3.1 Additive Outlier

We investigate the effect of AO on the parameter estimation of the RCA(1) model. Table 4.1 gives the simulation results for the true parameter values of coefficient $\theta = 0.3$ and variance of random disturbance $\sigma_b^2 = 0.16$ with different AO magnitudes ω. The table presents the bias for different parameters; θ, σ_b^2, σ_e^2 and different methods; LS, EF and IT, in contaminated and uncontaminated series of AO. In the first three rows, it can be seen that the AO affects the estimation of θ using all three methods. The LS estimate is the most affected by AO, where the difference between the bias of contaminated and uncontaminated series using LS is larger compared to that using EF and IT estimates. Note that in both equations (2.6) and (3.4), the IT and EF estimators of θ differ from LS by the factor $\frac{1}{\sigma_e^2 + y_{t-1}^2 \sigma_b^2}$. This weighted factor has successfully reduced the effect of AO in the estimation of θ. As a result, the IT and EF give better estimates than the LS when AO exist in the data.

Meanwhile, the fourth and seventh rows of Table 4.1 give the bias for σ_b^2 and σ_e^2 respectively. It can be seen that the IT method has improved the estimations of $\hat{\sigma}_b^2$ and $\hat{\sigma}_e^2$ and gives smaller bias than that using LS and EF

Table 4.1: With and without the AO effect on parameter estimation for true parameter values θ=0.3 and σ_b^2=0.16

		Least squares		**Estimating Function**		**Iterative EF**	
Parameter		$Bias^{LS}_{free}$	$Bias^{LS}_{AO}$	$Bias^{EF}_{free}$	$Bias^{EF}_{AO}$	$Bias^{IT}_{free}$	$Bias^{IT}_{AO}$
θ	ω=4	-0.00064	-0.00740	-0.00042	-0.00374	-0.00043	-0.00365
	ω=8	-0.00157	-0.02753	-0.00084	-0.00680	-0.00081	-0.00597
	ω=12	-0.00090	-0.05392	-0.00004	-0.01027	-0.00004	-0.00748
σ_b^2	ω=4	-0.01742	-0.02692	-0.01742	-0.02692	-0.01513	-0.02423
	ω=8	-0.01391	-0.05480	-0.01391	-0.05480	-0.01124	-0.04372
	ω=12	-0.02027	-0.08807	-0.02027	-0.08807	-0.01764	-0.06172
σ_e^2	ω=4	0.01982	0.06317	0.01982	0.06317	0.01700	0.05978
	ω=8	0.01405	0.19390	0.01405	0.19390	0.01077	0.17886
	ω=12	0.02082	0.40294	0.02082	0.40294	0.01760	0.36441

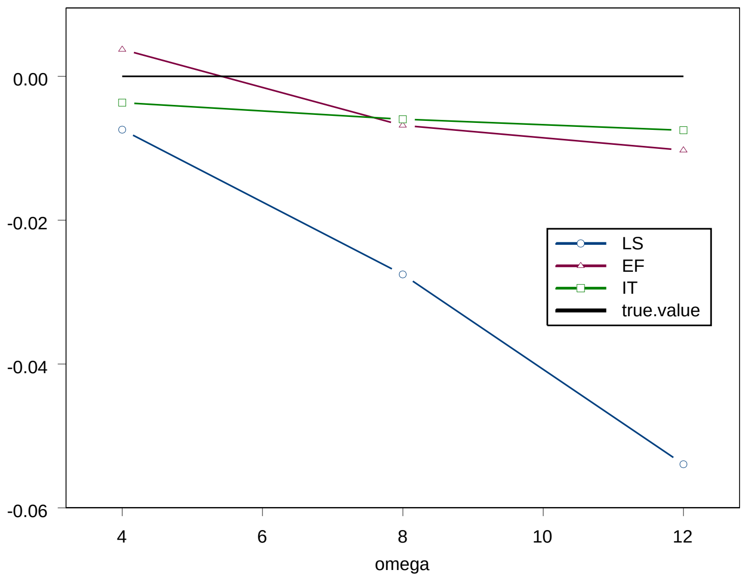

Figure 4.5: AO effect on parameter θ as ω increases

methods. Another noticeable and expected feature is that the bias using the three estimates increases as the values of ω increase. Figure 4.5 shows a clear comparison of the above information. The most affected estimate by AO is the LS while the least is the IT.

Table 4.2: With and without the AO effect on parameter estimation for true parameter values ω=8 and σ_b^2=0.16

		Least squares		**Estimating Function**		**Iterative EF**	
Parameter		$Bias^{LS}_{free}$	$Bias^{LS}_{AO}$	$Bias^{EF}_{free}$	$Bias^{EF}_{AO}$	$Bias^{IT}_{free}$	$Bias^{IT}_{AO}$
θ	θ=0.7	-0.00900	-0.03959	-0.00428	-0.00998	-0.00416	-0.00907
	θ=0.5	-0.00403	-0.03988	-0.00275	-0.00973	-0.00275	-0.00866
	θ=0.3	-0.00273	-0.02981	-0.00157	-0.00943	-0.00154	-0.00866
	θ=-0.3	0.00304	0.02901	0.00196	0.00898	0.00194	0.00818
	θ=-0.5	0.00361	0.03906	0.00102	0.00724	0.00096	0.00617
	θ=-0.7	0.00756	0.03837	0.00232	0.00780	0.00220	0.00681
σ_b^2	θ=0.7	-0.02153	0.03466	-0.02153	0.03466	-0.01779	0.05151
	θ=0.5	-0.02066	0.01162	-0.02066	0.01162	-0.01780	0.03068
	θ=0.3	-0.01774	-0.06157	-0.01774	-0.06157	-0.01565	-0.05114
	θ=-0.3	-0.01547	-0.06220	-0.01547	-0.06220	-0.01299	-0.05191
	θ=-0.5	-0.02274	0.01161	-0.02274	0.01161	-0.01982	0.03146
	θ=-0.7	-0.02463	0.03367	-0.02463	0.03367	-0.02139	0.05045
σ_e^2	θ=0.7	0.05805	0.07051	0.05805	0.07051	0.04988	0.02976
	θ=0.5	0.02781	0.11009	0.02781	0.11009	0.02344	0.07868
	θ=0.3	0.01842	0.20511	0.01842	0.20511	0.01588	0.19084
	θ=-0.3	0.01469	0.20460	0.01469	0.20460	0.01165	0.19061
	θ=-0.5	0.03235	0.11280	0.03235	0.11280	0.02781	0.07983
	θ=-0.7	0.05892	0.06700	0.05892	0.06700	0.05177	0.02614

In Table 4.2, we tabulate the results for the AO case with magnitude $\omega = 8$, random disturbance variance $\sigma_b^2 = 0.16$ and vary the coefficient θ of the RCA(1) model. It can be seen that the IT method gives better estimation for parameters; θ and σ_e^2. However for σ_b^2, the bias for IT is slightly larger when $|\theta| = 0.5$ and $|\theta| = 0.7$. We also present the bias for parameter θ graphically in Figure 4.6. The plot shows that the three methods over estimate the parameter θ when $\theta < 0$ and under estimate when $\theta > 0$. Furthermore, the estimated bias for LS, EF and IT get larger as $|\theta|$ gets larger. A reason for this is that, the value of stationary condition $\theta^2 + \sigma_b^2$ is near 1 and this is generally associated with the non-stationary series. However, this effect is minimal on EF and IT compared to LS estimate.

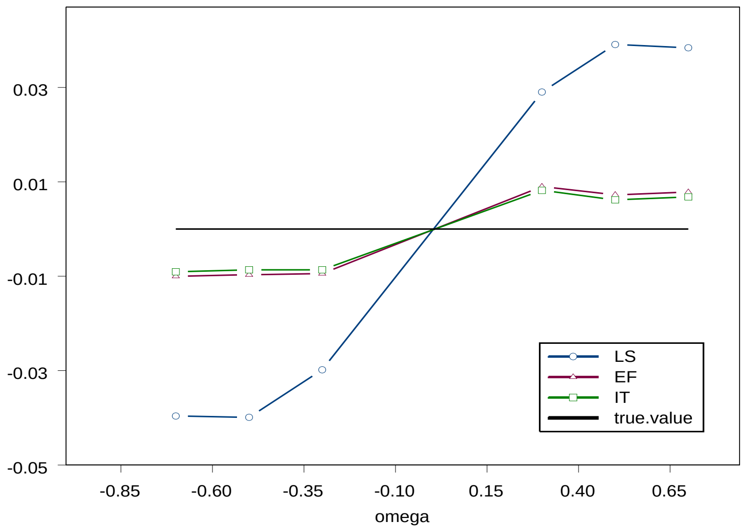

Figure 4.6: AO effect on parameter θ as θ increases

4.3.2 Innovational Outlier

We now investigate the effects of IO in a generated series from RCA(1) process. We use similar steps and set of parameters as in the AO case. Table 4.3 gives the bias of the estimated parameters when different IO magnitude ω are introduced in the series. Table 4.4 gives the bias of the estimated parameters when we vary the coefficient θ. In both tables, it can be seen that the bias of IT is smaller than that using LS and EF for all magnitudes of ω and for all different values of coefficient θ. Figures 4.7 and 4.8 show graphically the bias of parameter θ. The LS is most affected by IO compared to EF and IT.

Table 4.3: With and without the IO effect on parameter estimation for true parameter values θ=0.3 and σ_b^2=0.16

		Least squares		**Estimating Function**		**Iterative EF**	
Parameter		$Bias^{LS}_{free}$	$Bias^{LS}_{IO}$	$Bias^{EF}_{free}$	$Bias^{EF}_{IO}$	$Bias^{IT}_{free}$	$Bias^{IT}_{IO}$
θ	ω=4	-0.00257	-0.00221	-0.00173	-0.00173	-0.00172	-0.00172
	ω=8	-0.00515	-0.00669	-0.00398	-0.00545	-0.00395	-0.00541
	ω=12	0.00002	-0.00293	0.00085	0.00105	0.00085	0.00105
σ_b^2	ω=4	-0.01754	-0.02114	-0.01754	-0.02114	-0.01486	-0.01771
	ω=8	-0.01782	-0.04315	-0.01782	-0.04315	-0.01558	-0.02981
	ω=12	-0.01674	-0.06032	-0.01674	-0.06032	-0.01407	-0.02679
σ_e^2	ω=4	0.01706	0.05460	0.01706	0.05460	0.01382	0.05031
	ω=8	0.01877	0.18391	0.01877	0.18391	0.01605	0.16538
	ω=12	0.01830	0.37409	0.01830	0.37409	0.01501	0.32283

Table 4.4: With and without the IO effect on parameter estimation for true parameter values ω=8 and σ_b^2=0.16

		Least squares		**Estimating Function**		**Iterative EF**	
Parameter		$Bias_{free}^{LS}$	$Bias_{IO}^{LS}$	$Bias_{free}^{EF}$	$Bias_{IO}^{EF}$	$Bias_{free}^{IT}$	$Bias_{IO}^{IT}$
θ	θ=0.7	-0.00885	-0.00952	-0.00222	-0.00193	-0.00209	-0.00168
	θ=0.5	-0.00333	-0.00625	-0.00135	-0.00237	-0.00131	-0.00228
	θ=0.3	-0.00515	-0.00669	-0.00398	-0.00545	-0.00395	-0.00541
	θ=-0.3	0.00323	0.00426	0.001862	0.00238	0.001834	0.00238
	θ=-0.5	0.00460	0.00679	0.002491	0.00298	0.002457	0.00292
	θ=-0.7	0.00557	0.00792	-0.00013	0.00036	-0.00030	0.00016
σ_b^2	θ=0.7	-0.02295	-0.02601	-0.02295	-0.02601	-0.01932	-0.01979
	θ=0.5	-0.01763	-0.03106	-0.01763	-0.03106	-0.01460	-0.02132
	θ=0.3	-0.01782	-0.04315	-0.01782	-0.04315	-0.01558	-0.02981
	θ=-0.3	-0.01645	-0.03429	-0.01645	-0.03429	-0.01386	-0.01996
	θ=-0.5	-0.01902	-0.03274	-0.01902	-0.03274	-0.01604	-0.02321
	θ=-0.7	-0.02259	-0.02538	-0.02259	-0.02538	-0.01881	-0.01925
σ_e^2	θ=0.7	0.05720	0.19933	0.05720	0.19933	0.04890	0.18430
	θ=0.5	0.02092	0.17437	0.02092	0.17437	0.01632	0.15777
	θ=0.3	0.01877	0.18391	0.01877	0.18391	0.01605	0.16538
	θ=-0.3	0.02134	0.17736	0.02134	0.17736	0.01816	0.15720
	θ=-0.5	0.02393	0.17960	0.02393	0.17960	0.01941	0.16347
	θ=-0.7	0.05400	0.19485	0.05400	0.19485	0.04518	0.17931

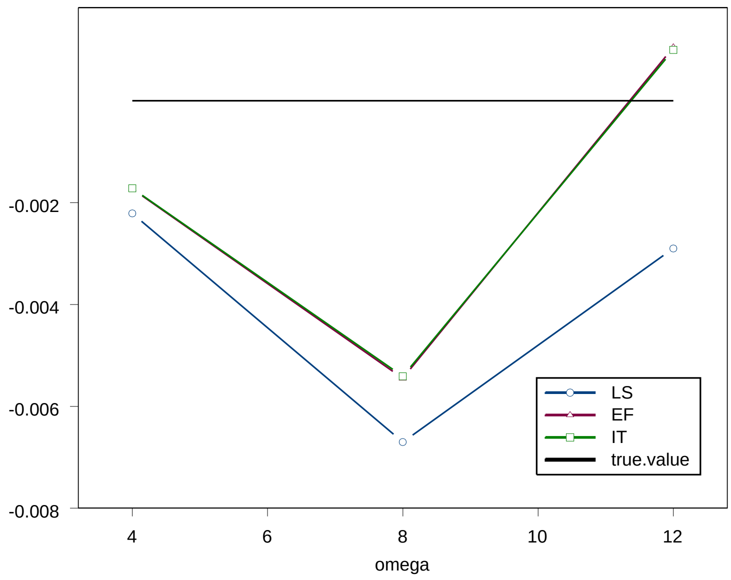

Figure 4.7: IO effect on parameter θ as ω increases

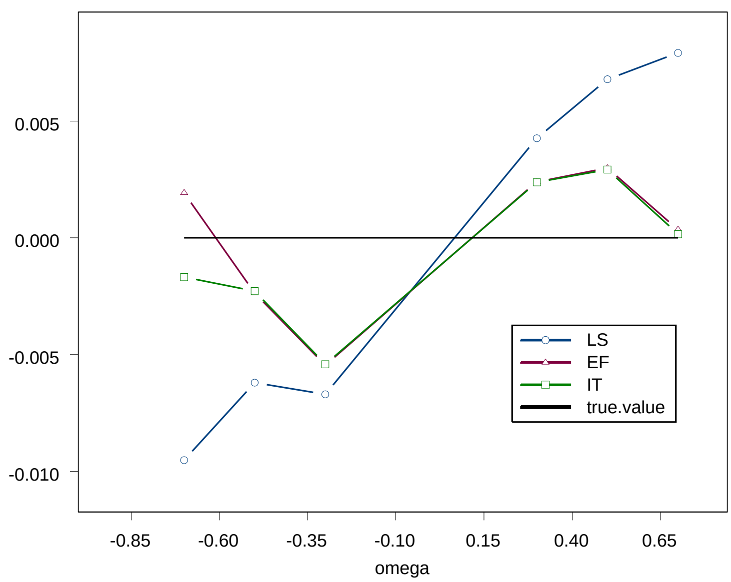

Figure 4.8: IO effect on parameter θ as θ increases

By looking at the results in both AO and IO cases, it is clear that the IT estimate is more robust and preferable to be used in practice, when AO and IO occur in the process.

Chapter 5

Outlier Detection in RCA(1) Models

In chapter 4, we have shown the effects of additive (AO) and innovational outliers (IO) to the parameter estimation. To overcome this problem, we need to detect these outliers and consequently remove their effects accordingly. In some cases, we might use visual inspection to detect outliers. However, such procedure is very subjective and suffers basic limitations of data visualization methods as pointed out by Last and Kendel (2001). That is, an analyst needs to apply his or her own subjective perception to judge whether an observation is "very outstanding" or "too far" from the rest of the observations.

In this chapter, we will develop the statistical procedures of AO and IO detection for the RCA(1) process. These suggested procedures are the extension of Chang et al. (1988) work on ARIMA models.

There are three stages in developing the AO and IO detection procedures. The first and most important stage is to derive the statistics to measure the AO and IO effects. We denote the statistics as ω_{AO} and ω_{IO} respectively. Secondly, the test statistics (i.e., the standardization of outlier effects) for AO and IO are defined. This test statistics will be used to decide whether the null hypothesis H_0: $\omega=0$ at known outlier position t should be rejected or not. The rejection of the null hypothesis indicates the existence of outlier at the particular point. However, the position of the outliers are usually unknown. Thus, the next stage of the outlier detection procedures is to identify the occurrence of single outlier in a data set one at a time, using the test criteria. If we have more than one outlier, these procedures need to be repeated iteratively until no outlier is left in the data.

The following Sections describe the outlier detection procedures in detail.

5.1 Estimation of Outlier Effects in RCA(1)

In this section, two statistics for measuring the AO and IO effects in the RCA(1) process are derived. The mean and variance of these statistics are also obtained.

5.1.1 Additive Outlier

To measure the AO effects, we follow the work of Chang et al. (1988), assuming that the parameters are known. From Section 4.1, the affected residuals are given by

$$u^*_{d+k} = \begin{cases} u_{d+k} + \omega & \text{for } k = 0 \\ u_{d+k} - \theta\omega & \text{for } k = 1 \\ u_{d+k} & \text{for } k = 2, 3, ... \end{cases}$$

Using the least squares method, the sum of squares of the residuals is

$$S = \sum_{t=1}^{n} {u_t}^2 = \sum_{t=1}^{d-1} {u_t}^2 + {u_d}^2 + {u_{d+1}}^2 + \sum_{t=d+2}^{n} {u_t}^2. \tag{5.1}$$

Equivalently, we have

$$S = \sum_{t=1}^{d-1} {u_t}^2 + (u^*_d - \omega)^2 + (u^*_{d+1} + \theta\omega)^2 + \sum_{t=d+2}^{n} {u_t}^2. \tag{5.2}$$

Differentiating S with respect to ω gives

$$\frac{\partial S}{\partial \omega} = 2(u^*_d - \omega)(-1) + 2(u^*_{d+1} + \theta\omega)(\theta). \tag{5.3}$$

Solving $\frac{\partial S}{\partial \omega} = 0$, one has

$$\omega - u^*_d + \theta u^*_{d+1} + \omega\theta^2 = 0$$

$$\omega(1 + \theta^2) - u^*_d + \theta u^*_{d+1} = 0. \tag{5.4}$$

Hence, an estimate of the AO effect at time $t = d$ is given by

$$\hat{\omega}_{AO,d} = \frac{u^*_d - \theta u^*_{d+1}}{1 + \theta^2}. \tag{5.5}$$

The conditional mean of the ω_{AO} is

$$\begin{aligned}
E[\hat{\omega}_{AO,d}|F_{t-1}^{y}] &= E[\frac{u_d^* - \theta u_{d+1}^*}{1+\theta^2}] \\
&= \frac{E[u_d^*] - \theta E[u_{d+1}^*]}{1+\theta^2} \\
&= \frac{E[u_d + \omega] - \theta E[u_{d+1} - \theta\omega]}{1+\theta^2} \\
&= \frac{\omega + \theta^2\omega}{1+\theta^2} \\
&= \omega, \qquad (5.6)
\end{aligned}$$

where $E[u_i] = 0$ for $i = 1, 2, ..., n$ and F_{t-1}^{y} denotes the series generated by y up to time $t-1$, namely $y_1, y_2, ..., y_{t-1}$. Since

$$\begin{aligned}
Cov[u_{d+1}, u_d] &= E\{(u_{d+1} - E[u_{d+1}](u_d - E[u_d])\} \\
&= E\{(u_{d+1} - 0)(u_d - 0)\} \\
&= E\{u_{d+1}u_d\} \\
&= 0, \qquad (5.7)
\end{aligned}$$

we have

$$Var[u_{d+1} + u_d] = Var[u_{d+1}] + Var[u_d]. \qquad (5.8)$$

And the conditional variance of $\hat{\omega}_{AO,d}$ is

$$\begin{aligned}
Var[\hat{\omega}_{AO,d}|F_{t-1}^{y}] &= Var[\frac{u_d^* - \theta u_{d+1}^*}{1+\theta^2}] \\
&= \frac{1}{(1+\theta^2)^2}\{Var[u_d^*] + \theta^2 Var[u_{d+1}^*]\} \\
&= \frac{1}{(1+\theta^2)^2}\{Var[u_d + \omega] + \theta^2 Var[u_{d+1} - \omega\theta]\} \\
&= \frac{1}{(1+\theta^2)^2}\{Var[u_d] + \theta^2 Var[u_{d+1}]\} \\
&= \frac{[\sigma_b^2 y_{d-1}^2 + \sigma_e^2] + \theta^2[\sigma_b^2 y_d^2 + \sigma_e^2]}{(1+\theta^2)^2} \quad \text{from (4.4)} \\
&= \frac{\sigma_e^2(\theta^2+1) + \sigma_b^2(\theta^2 y_d^2 + y_{d-1}^2)}{(1+\theta^2)^2}. \qquad (5.9)
\end{aligned}$$

5.1.2 Innovational Outlier

From Section 4.2, the residuals u_t^* affected by IO are given by

$$u_t^* = \begin{cases} u_t & \text{for } t \neq d \\ u_t + \omega & \text{for } t = d \end{cases}$$

Using the least squares method as in the AO case, we have

$$\begin{aligned} S &= \sum_{t=1}^{n} {u_t}^2 \\ &= \sum_{t=1}^{d-1} {u_t}^2 + {u_d}^2 + \sum_{t=d+1}^{n} {u_t}^2 \\ &= \sum_{t=1}^{d-1} {u_t}^2 + (u_d^* - \omega)^2 + \sum_{t=d+1}^{n} {u_t}^2. \end{aligned} \tag{5.10}$$

Differentiating S with respect to ω

$$\frac{\partial S}{\partial \omega} = 2(u_d^* - \omega) \tag{5.11}$$

and solving $\frac{\partial S}{\partial \omega} = 0$, an estimate of the IO effect at time $t = d$ is given as

$$\hat{\omega}_{IO,d} = u_d^*. \tag{5.12}$$

The conditional mean of $\hat{\omega}_{IO,d}$ is given by

$$\begin{aligned} E(\hat{\omega}_{IO,d}|F_{t-1}^y) &= E[u_d^*] \\ &= E[u_d + \omega] \\ &= \omega, \end{aligned} \tag{5.13}$$

since $E[u_i] = 0$ for $i = 1, 2, ..., n$. From (5.8), the conditional variance of $\hat{\omega}_{IO,d}$ is

$$\begin{aligned} Var(\hat{\omega}_{IO,d}|F_{t-1}^{y}) &= Var(u_d^*) \\ &= Var(u_d + \omega) \\ &= Var(u_d) \\ &= \sigma_e^2 + \sigma_b^2 y_{d-1}^2. \end{aligned} \tag{5.14}$$

5.2 Test Statistics and Test Criteria

Tsay (1986) and Chang et al. (1988) had derived the likelihood ratio criteria for testing the existence of AO and IO in linear time series. The same form of criteria has been extended to nonlinear models by Franses and Ghijsels (1999) and Charles and Darne (2005) for GARCH models, and by Mohamed (2005) and Zaharim et al. (2006) for bilinear models. We now consider this approach for the RCA(1) model.

Let H_0 be the null hypothesis of no outlier in the RCA(1) process and denoted by H_0: $\omega = 0$. An alternative hypothesis for the presence of AO and IO is denoted by H_1 and H_2 respectively. That is, we are testing for

$$H_1 : \omega_{AO} \neq 0 \tag{5.15}$$

or

$$H_2 : \omega_{IO} \neq 0. \tag{5.16}$$

Using the central limit theorem and assuming regularity conditions on moments, we have

$$\tau_{AO,t} = \frac{\hat{\omega}_{AO,t} - E[\hat{\omega}_{AO,t}]}{\sqrt{Var(\omega_{AO,t})}} \tag{5.17}$$

and

$$\tau_{IO,t} = \frac{\hat{\omega}_{IO,t} - E[\hat{\omega}_{IO,t}]}{\sqrt{Var(\omega_{IO,t})}}. \tag{5.18}$$

Thus, under the null hypothesis, the test statistics for AO and IO are given by

$$\begin{aligned} \tau_{AO,t} &= \frac{\hat{\omega}_{AO,t}}{\sqrt{Var(\omega_{AO,t})}} \\ &= \frac{(u_d^* + u_{d+1}^*)}{\sqrt{\sigma_e^2(\theta^2+1) + \sigma_b^2(\theta^2 y_d^2 + y_{d-1}^2)}} \end{aligned} \tag{5.19}$$

and

$$\begin{aligned} \tau_{IO,t} &= \frac{\hat{\omega}_{IO,t}}{\sqrt{Var(\omega_{IO,t})}} \\ &= \frac{u_d^*}{\sqrt{\sigma_e^2 + \sigma_b^2 y_{d-1}^2}}. \end{aligned}$$

respectively.

Note that:

1. If the position of the outlier is known to be at $t = d$, we use (5.19) or (5.20) to confirm the presence of AO or IO.
2. If the position of an outlier is unknown, we define the following test criteria to detect a single outlier at a time as

$$\eta_{AO} = max_{t=1}^{n}\{|\tau_{AO,t}|\} \tag{5.20}$$

and

$$\eta_{IO} = max_{t=1}^{n}\{|\tau_{IO,t}|\} \tag{5.21}$$

for AO and IO respectively.

To confirm the presence of AO and IO, we need a suitable critical value to accept the alternative hypothesis in (5.15) and (5.16) respectively. The critical value can be obtained by investigating the sampling behavior of the test criteria in a normal, uncontaminated series. Now, we investigate the sampling behavior of the test criteria in (5.20) and (5.21) for AO and IO respectively.

5.3 Sampling Behavior of the Test Criteria

The sampling behaviour of test criteria is studied to decide the pre-determined critical value C for AO and IO. We are looking for the maximum value of the standardized statistics τ_t in a normal, uncontaminated series to be the critical value. The sampling behavior of the test criteria are investigated based on different factors:

a) Sample size, n

b) Coefficient of the RCA(1) model, θ

c) Variance of random disturbance, σ_b^2

d) Methods used to estimate the test criteria (i.e., using LS, EF and IT)

e) Types of outliers (i.e., AO and IO)

Different models are considered. They represent a broad choice of RCA(1) coefficients; from $-1 < \theta < 1$ to values of σ_b^2 which are closer to the bound of stationarity condition ($\theta^2 + \sigma_b^2 \simeq 1$). For each model, three cases of different sample sizes $n = 60, n = 100$ and $n = 200$ are examined. The random errors, e_t are assumed to follow standard normal distribution and the variance of random disturbance is fixed to $\sigma_b^2 = 0.16$. One would ask why such restriction on σ_b^2 is enforced. We will show in the later simulation part that small changes of σ_b^2 would not affect the sampling behavior as much.

Below are the steps taken to investigate the sampling behavior of the test criteria:

1. Generate an outlier free series of RCA(1) model. Similar scheme in Section 3.5 is adopted.

2. Estimate the test statistics $\hat{\eta}_{AO}$ and $\hat{\eta}_{IO}$ using (5.20) and (5.21) for AO and IO respectively using LS, EF and IT methods.

3. Repeat steps 1 and 2 for 1,000 times giving $\hat{\eta}_{AO,1}$,$\hat{\eta}_{AO,2}$,...,$\hat{\eta}_{AO,1000}$ and $\hat{\eta}_{IO,1}$,$\hat{\eta}_{IO,2}$,...,$\hat{\eta}_{IO,1000}$ for LS, EF and IT methods.

4. For each estimation method, we calculate the 90^{th}, 95^{th} and 99^{th} percentile levels of the test criteria for AO and IO.

The next two sections report the results of sampling behavior for AO and IO cases.

5.3.1 Additive Outlier

Table 5.1 presents the sampling behavior of AO test criteria based on three estimation methods; LS, EF and IT, together with their 99^{th}, 95^{th} and 90^{th} percentile levels. Two different coefficients $\theta = 0.1$ and 0.7 are considered, with sample size $n = 100$ and variance of random disturbance $\sigma_b^2 = 0.16$. It is clear that the AO test criteria for each percentile level are almost the same for IT, EF and LS methods. The difference is too small ranging from 0 to 0.09. Thus, we will only focus on the test criteria using IT henceforth in describing the sampling behavior of AO test criteria.

Table 5.1: Sampling behavior of AO test criteria for different estimation methods

		Est. Methods		
	Percentile	IT	EF	LS
$\theta = 0.1$	90^{th}	3.13	3.14	3.14
	95^{th}	3.34	3.37	3.35
	99^{th}	3.71	3.80	3.72
$\theta = 0.7$	90^{th}	3.15	3.16	3.16
	95^{th}	3.35	3.35	3.35
	99^{th}	3.75	3.75	3.75

Next, we conduct another simulation study to investigate the sampling behaviour of AO test criteria for different random disturbance variance σ_b^2. We fix the sample size $n = 100$, coefficient θ=0.1 and vary the variance of random disturbance σ_b^2. The 99^{th}, 95^{th} and 90^{th} percentile values for different values of σ_b^2 (using IT) are presented in Table 5.2. The AO test criteria for each percentile level increases slightly as the value of σ_b^2 increases. With this, we fixed the random disturbance variance $\sigma_b^2 = 0.16$ to reduce the number of sampling factors for the convenience of our analysis.

Table 5.2: Sampling behavior of AO test criteria for different values of σ_b^2

	Percentile		
	90^{th}	95^{th}	99^{th}
$\sigma_b^2 = 0.16$	3.18	3.34	3.66
$\sigma_b^2 = 0.25$	3.18	3.41	3.80
$\sigma_b^2 = 0.30$	3.22	3.42	3.82

Further, a simulation study for the negative values of coefficient θ is performed. We fixed the sample size $n = 100$, variance of random disturbance $\sigma_b^2 = 0.16$ and vary the coefficient θ. Figure 5.1 gives the 95^{th} percentile plot of AO test criteria when θ ranges from -0.7 to 0.7. From the symmetrical plot at $\theta = 0$, we found out that the AO sampling behavior for negative coefficients are similar to that for positive coefficients. Hence, the remaining investigation will focus only on positive coefficient θ.

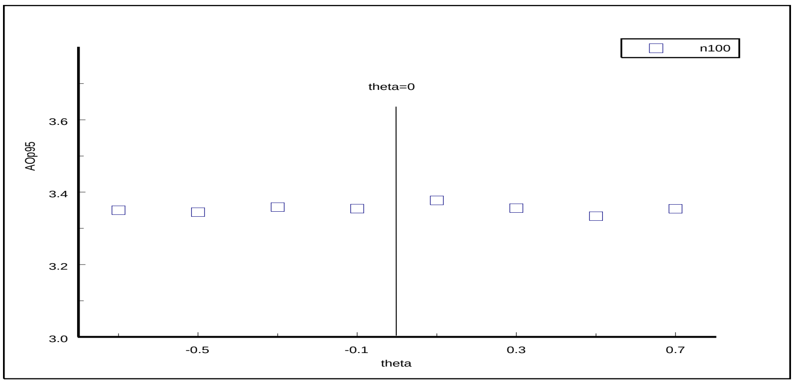

Figure 5.1: Sampling behavior of AO test criteria for positive and negative coefficient θ

Figures 5.2 to 5.4 consist of different combinations of AO test criteria at 99^{th}, 95^{th} and 90^{th} percentile levels. In each plot, we fixed $\sigma_b^2 = 0.16$ and vary the positive coefficient θ for three different sample sizes $n = 60, 100, 200$. All three percentile levels show a clear pattern of increasing with sample size n. As the sample size n increases, the percentile value increases. This is expected due to the extreme-value nature of the test criteria (they are maximum of a set of random variables). Thus, the tail probability of the test criteria is expected to increase as the sample size increases (see Chang et al.,1988). Further, the increments of percentile values are quite slight. For instance, when sample size n changes from $n = 60$ to $n = 200$, the 99^{th} percentile value of $\theta = 0.1$ changes

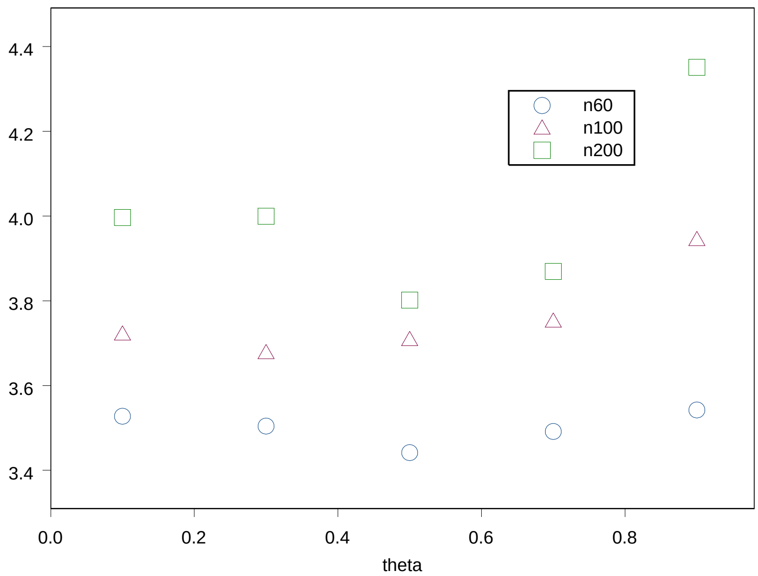

Figure 5.2: Sampling behavior of AO test criteria at 99^{th} percentile level

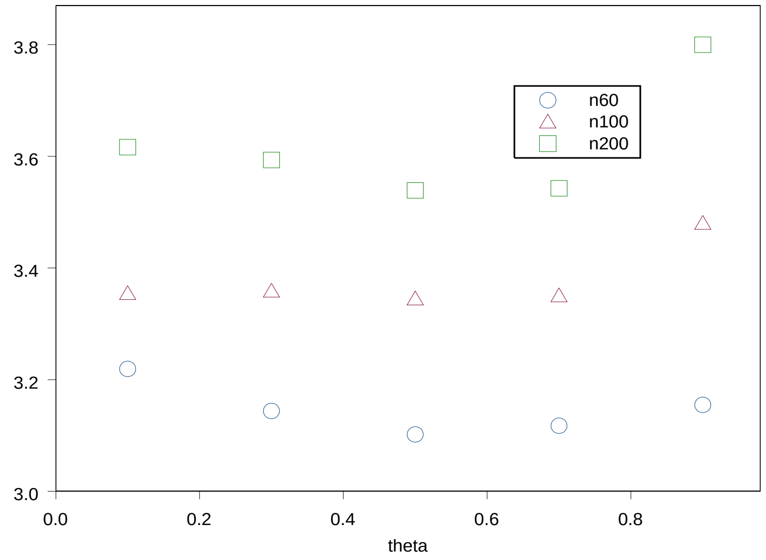

Figure 5.3: Sampling behavior of AO test criteria at 95^{th} percentile level

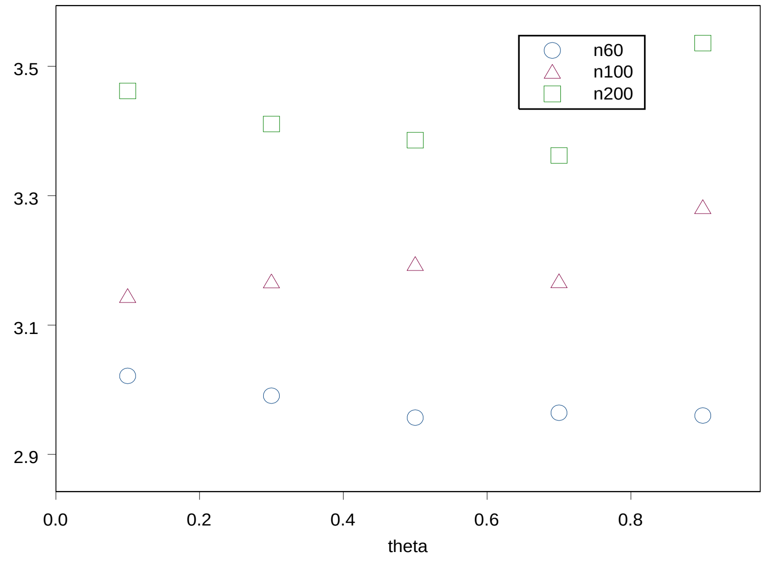

Figure 5.4: Sampling behavior of AO test criteria at 90^{th} percentile level

from 3.68 to 3.99. However, the changes are quite large for $\theta = 0.9$. A reason for this change is that, the stationarity condition $\theta^2 + \sigma_b^2$ is close to unity. The series tend to fluctuate largely and hence contribute to large percentile values.

The 90^{th} percentile value for all sample sizes and coefficient values ranged between 2.9 to 3.6, the 95^{th} percentile value ranged between 3.1 to 3.8 while the 99^{th} percentile value ranged between 3.4 to 4.4.

5.3.2 Innovational Outlier

Similar restrictions as in AO are used in describing the sampling behavior of the IO test criteria. We find out that the percentile value for IO test criteria does not depend on the estimation method (see Table 5.3), the values of random disturbance variance σ_b^2 (see Table 5.4) and the sign of coefficient values θ (see Figure 5.5).

Table 5.3: Sampling behavior of IO test criteria for different estimation methods

IO		**Est. Methods**		
	Percentile	IT	EF	LS
$\theta = 0.1$	90^{th}	2.68	2.67	2.68
	95^{th}	2.78	2.77	2.78
	99^{th}	2.99	2.99	3.00
$\theta = 0.7$	90^{th}	2.74	2.74	2.75
	95^{th}	2.87	2.86	2.89
	99^{th}	3.27	3.25	3.27

Table 5.4: Sampling behavior of IO test criteria for different values of σ_b^2

	Percentile		
	90^{th}	95^{th}	99^{th}
$\sigma_b^2 = 0.16$	2.68	2.78	2.99
$\sigma_b^2 = 0.25$	2.71	2.81	3.04
$\sigma_b^2 = 0.30$	2.69	2.85	3.02

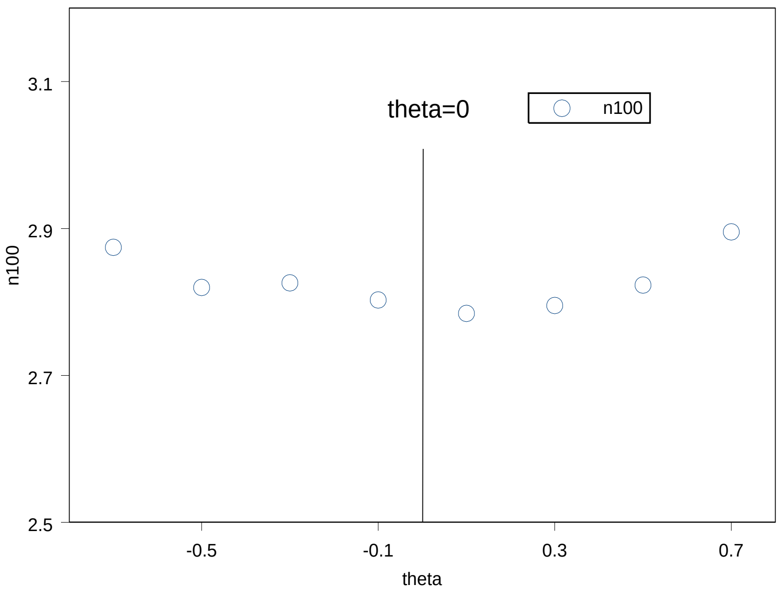

Figure 5.5: Sampling behavior of IO test criteria for positive and negative coefficient θ

Figures 5.6 to 5.8 plots the percentile values IO test criteria at the 99^{th}, 95^{th} and 90^{th} percentile levels respectively. Similar to the AO case, the percentile value increases as the sample size n and the coefficient θ increase. The increments of these percentile values are quite slight, except for large coeffi-

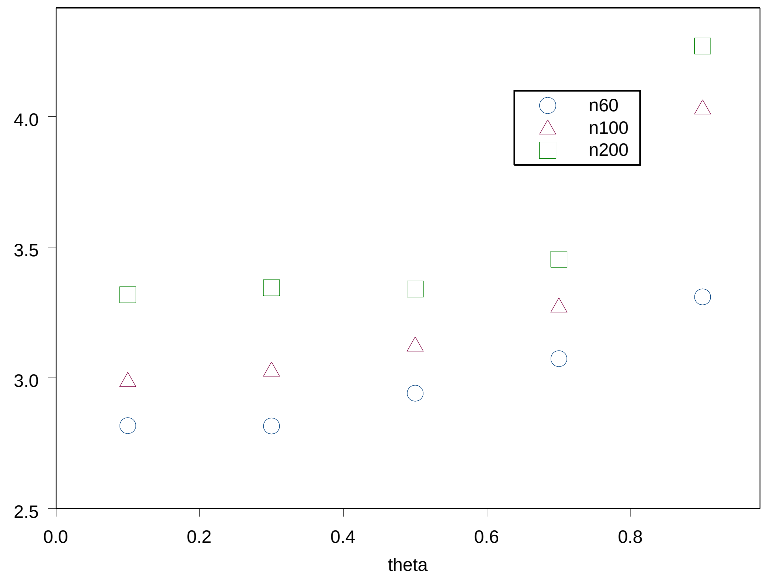

Figure 5.6: Sampling behavior of IO test criteria at 99^{th} percentile level

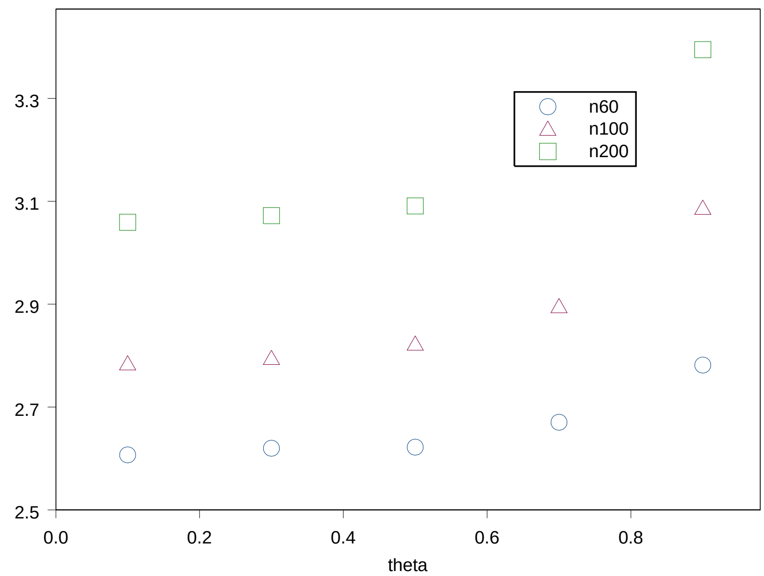

Figure 5.7: Sampling behavior of IO test criteria at 95^{th} percentile level

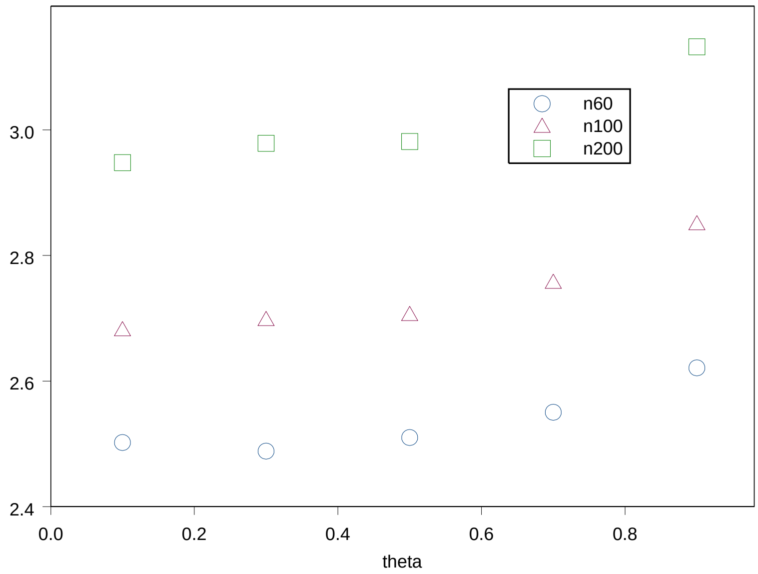

Figure 5.8: Sampling behavior of IO test criteria at 90^{th} percentile level

cient $\theta = 0.9$. The 90^{th} percentile values for all sample sizes and coefficients θ values ranged from 2.8 to 4.5, the 95^{th} percentile values ranged from 2.6 to 3.5 while the 99^{th} percentile values ranged from 3.2 to 4.1.

The percentile values will be used as the critical value C for the proposed detection procedures. In practice, it is recommended to use more than one critical value for analysis. Based on the presented $90th$ to $99th$ percentile values, critical values of 2.5 to 4.5 seem appropriate for use in order to identify the presence of AO and IO in (5.15) and (5.16) respectively.

Now, we propose the general single outlier detection procedures for AO and IO.

5.4 General Single Outlier Detection Procedures for AO and IO

The procedure begins with modeling the original time series data by supposing no outlier exists. Full procedure for detecting the AO is described below:

1. Estimate the parameters of the RCA(1) model using the original data. Hence, residuals of the model can be obtained.

2. Compute $\tau_{AO,t}$ for each time $t = 1, 2, ..., n$ using the residuals obtained in step 1.

3. Find the AO test criteria, $\eta_{AO} = \max_t \{|\tau_{AO,t}|\}$.

4. Given a pre-determined critical value C, if $\eta_{AO} = |\tau_{AO,d}| > C$, then there exists an AO at time d

The same steps are used for detecting an IO, with the AO replaced by IO. Through the suggested procedures, an AO or IO can be detected at the right time point t.

In the next section, a simulation study is carried out to investigate the performance of the outlier detection procedures in detecting AO and IO.

5.5 Performance of the Outlier Detection Procedures

A simulation study is conducted to examine the performance of the proposed general procedures for detecting AO and IO. It is applied to cases characterized by a combination of following factors:

1. One underlying RCA(1) model with different combination of coefficient θ=0.1, 0.3, 0.5 and 0.7.
2. Four values of outlier magnitude $\omega = 4, 6, 8$ and 10
3. Three different values of $\sigma_b^2 = 0.16, 0.25$ and 0.35
4. Three different sample sizes $n = 60, 100$ and 200
5. Five chosen critical values, C=2.5, 3.0, 3.5, 4.0, 4.5.

We first consider the AO case. Detailed steps to evaluate the AO detection performance are as follows:

1. Generate a series of RCA(1) and introduce an AO with magnitude ω at time $t = d$. The random errors, e_t, are assumed to follow standard normal distribution.
2. Estimate the test criteria $\hat{\eta}_{AO}$ using the IT method.
3. If the AO test criteria $\hat{\eta}_{AO}$ has the same value corresponding to $|\hat{\tau}_{AO,d}|$ and exceeds a pre-determined critical value C, then we reject the null hypothesis. Hence, an AO has occurred at time $t = d$.

4. Repeat the above steps 1,000 times and the proportion of correctly detecting AO is calculated.

The same steps as above are used for the IO case, based on the $\hat{\eta}_{IO}$. The performance of correctly detecting AO and IO are presented in Table 5.5 and 5.6 respectively.

Table 5.5: The performance of correctly detecting AO

				critical value				
ω	θ	σ_b^2	n	2.5	3.0	3.5	4.0	4.5
4	0.1	0.16	100	0.789	0.659	0.458	0.228	0.081
6				0.981	0.966	0.911	0.789	0.570
8				0.996	0.996	0.991	0.977	0.945
10				0.999	0.999	0.998	0.994	0.977
8	0.1	0.16	100	0.996	0.996	0.991	0.977	0.945
	0.3			0.989	0.982	0.962	0.905	0.818
	0.5			0.910	0.877	0.798	0.651	0.473
	0.7			0.611	0.570	0.467	0.304	0.135
	0.9			0.574	0.553	0.487	0.388	0.223
8	0.1	0.16	100	0.996	0.996	0.991	0.977	0.945
		0.25		0.992	0.986	0.973	0.949	0.885
		0.35		0.984	0.975	0.962	0.931	0.844
8	0.1	0.16	60	0.993	0.988	0.966	0.914	0.769
			100	0.996	0.996	0.991	0.977	0.945
			200	0.999	0.999	0.995	0.989	0.980

The performance of correctly detecting AO is given in Table 5.5. The table consists of different considered cases using five critical values. For the case of different values of AO magnitude ω, we fixed $\theta = 0.1$, σ_b^2=0.16, $n = 100$ and introduced an AO at time $t = 50$. Whilst for different coefficients θ, we fixed $n = 100$, $\omega = 8$ and σ_b^2=0.16. For the case of different σ_b^2, $n = 100$, $\omega = 8$ and $\theta = 0.1$ have been fixed. As for different case of n, we fixed $\theta = 0.1$, σ_b^2=0.16, $\omega = 8$ and introduced the AO to three different sample sizes $n = 60, 100$ and 200 at time $t = 30$, $t = 50$, and $t = 100$ respectively. The values in Table 5.5 represent the proportion of correctly detecting AO. For instance, in the first row, the proportion of correctly detecting the AO is 0.789 when using critical value 2.5. The proportion of correct detection can be interpreted as a power of the procedure in term of outlier detection.

Three main points can be observed from the AO detection performance results in Table 5.5. Firstly, in all cases, the power is a decreasing function of the critical value C. When the critical value is too large, fewer outliers will be detected. This is expected, since some of the test criteria $\hat{\eta}_{AO}$ may have lower value than the critical value C. Secondly, the performance of AO detection procedure improves when larger magnitudes of ω are introduced. At the lowest critical value C=2.5, this procedure can capture almost 100% of the introduced AO. Lastly, as coefficient θ, σ_b^2 and n increases, the performance of AO detection procedure decreases. The detection drops abruptly at the coefficient $\theta = 0.7$, where the proportion of detecting AO at C=2.5 is only 0.611. This is due to the formulation of AO test statistics τ_{AO} which depend on coefficient θ. However, in general, the AO detection procedure works well if the stationarity condition $\theta^2 + \sigma_b^2$ is far from unity.

Table 5.6: The performance of correctly detecting IO

				critical value				
ω	θ	σ_b^2	n	2.5	3.0	3.5	4.0	4.5
4	0.1	0.16	100	0.763	0.654	0.466	0.266	0.104
6				0.980	0.974	0.933	0.841	0.663
8				0.996	0.990	0.986	0.968	0.928
10				1.000	1.000	0.997	0.996	0.991
8	0.1	0.16	100	0.996	0.990	0.986	0.968	0.928
	0.3			0.998	0.997	0.996	0.993	0.986
	0.5			0.995	0.990	0.990	0.985	0.974
	0.7			0.993	0.992	0.987	0.977	0.963
	0.9			0.957	0.942	0.916	0.869	0.823
8	0.1	0.16	100	0.996	0.990	0.986	0.968	0.928
		0.25		0.999	0.999	0.998	0.986	0.978
		0.35		0.997	0.995	0.986	0.973	0.954
8	0.1	0.16	60	0.998	0.993	0.981	0.953	0.899
			100	0.996	0.990	0.986	0.968	0.928
			200	0.993	0.991	0.989	0.983	0.973

The performance of IO detection procedure is reported in Table 5.6. The same sets of parameters and critical values as in the AO case are used. In general, the power of IO detection procedure increases as IO magnitude ω increases and decreases as critical value C, coefficient θ, variance of random disturbance σ_b^2 and sample size n increases. When compared to the performance of AO, similar results can be observed, except that there is no abrupt drops at the largest coefficient $\theta = 0.7$. The IO detection performance does not depends on coefficient θ (see the formulation of IO test statistics, τ_{IO} in equation (5.12)).

We also present the results for the misdetection of outlier location for both AO and IO as below:

Table 5.7: Misdetection of AO detection procedure

				critical value				
ω	θ	σ_b^2	n	2.5	3.0	3.5	4.0	4.5
4	0.1	0.16	100	0.167	0.056	0.015	0.000	0.000
6				0.011	0.009	0.003	0.002	0.001
8				0.004	0.003	0.002	0.001	0.000
10				0.001	0.001	0.001	0.001	0.001
8	0.1	0.16	100	0.004	0.003	0.002	0.001	0.000
	0.3			0.011	0.006	0.004	0.001	0.000
	0.5			0.083	0.068	0.049	0.026	0.008
	0.7			0.372	0.328	0.263	0.172	0.093
	0.9			0.398	0.362	0.315	0.256	0.157
8	0.1	0.16	100	0.004	0.003	0.002	0.001	0.000
		0.25		0.007	0.004	0.002	0.000	0.000
		0.35		0.016	0.007	0.003	0.001	0.001
8	0.1	0.16	60	0.006	0.004	0.001	0.001	0.001
			100	0.004	0.003	0.002	0.001	0.000
			200	0.001	0.001	0.001	0.000	0.000

Tables 5.7 and 5.8 give the misdetection of the outlier location for AO and IO respectively. The same sets of parameters similar to AO and IO detection performance in Table 5.5 and 5.6 respectively, are considered. The proportion of misdetection for both AO and IO increases as θ,σ_b^2 and n increases and ω decreases. However, for AO case, the proportion of incorrect outlier location is the highest for largest coefficient θ. The AO detection procedure is not very efficient for large values of coefficient θ.

Table 5.8: Misdetection of IO detection procedure

				critical value				
ω	θ	σ_b^2	n	2.5	3.0	3.5	4.0	4.5
4	0.1	0.16	100	0.181	0.080	0.017	0.003	0.000
6				0.016	0.004	0.002	0.001	0.001
8				0.002	0.002	0.000	0.000	0.000
10				0.000	0.000	0.000	0.000	0.000
8	0.1	0.16	100	0.002	0.002	0.000	0.000	0.000
	0.3			0.002	0.002	0.001	0.000	0.000
	0.5			0.005	0.004	0.001	0.001	0.000
	0.7			0.005	0.003	0.001	0.001	0.001
	0.9			0.028	0.016	0.008	0.002	0.002
8	0.1	0.16	100	0.002	0.002	0.000	0.000	0.000
		0.25		0.001	0.001	0.000	0.000	0.000
		0.35		0.002	0.001	0.000	0.000	0.000
8	0.1	0.16	60	0.002	0.000	0.000	0.000	0.000
			100	0.002	0.002	0.000	0.000	0.000
			200	0.006	0.003	0.002	0.000	0.000

By looking at the above results, we conclude that the procedures work very well in detecting AO and IO in the RCA(1) process, especially when the stationarity condition $\theta^2 + \sigma_b^2$ is far away from unity.

The next chapter will apply these outlier detection procedures to a real data set and illustrate the importance of our study in practice.

Chapter 6

Data Analysis

In this chapter, the differenced Indian CPI data considered in Section 3.4 is further analyzed. The objective is to illustrate the application of the proposed outlier detection procedures to a real data set from the RCA(1) model. To detect any outliers in a real data set, we use the following steps:

1. Estimate both AO and IO test criteria, $\hat{\eta}_{AO}$ and $\hat{\eta}_{IO}$ respectively.

2. If $\hat{\eta}_{AO}$ or $\hat{\eta}_{IO}$ exceed the critical value C, then we may say that AO or IO has occurred at time $t = d$ corresponding to the maximum value of $|\hat{\tau}_{AO,t}|$ or $|\hat{\tau}_{IO,t}|$ respectively.

3. Once detected, remove the effect of AO and IO at time $t = d$. Below are the steps in detail to remove the AO effect from the series:

    - To get an uncontaminated AO series $\{y_t\}$, the AO definition in Section 4.1 is used. The observation with AO at detected time $t = d$, y_d^*, is subtracted by the estimated AO effect $\hat{\omega}_{AO,d}$.

- Using the above uncontaminated observations $\{y_t\}$, we calculate the corresponding residuals and AIC value.

Whilst the steps to remove IO effect from the series are as follows:

- To get an uncontaminated IO series $\{y_t\}$, the derived formula in equation (4.9) is used. The observation with IO at detected time $t = d$, y_d^*, is subtracted by the estimated IO effect $\hat{\omega}_{IO,d}$ and coefficient θ.

- Using the above uncontaminated observations $\{y_t\}$, we calculate the corresponding residuals and AIC value.

4. The process is repeated iteratively until no outlier is detected in the data. In other words, the process should be stopped when the values of $\hat{\eta}_{AO}$ and $\hat{\eta}_{IO}$ are less than the critical value C.

We apply the outlier detection procedures to the differenced Indian CPI data. The test criteria resulting from the first iteration of the outlier detection procedures is given in Table 6.1. When critical value C=3.0 is considered, both AO and IO procedures detected an AO and IO respectively at the same time point $t = 6$. It corresponds to the highest spike seen in the time series plot as given in Figure 6.1.

We remove the effects of both AO and IO at time $t = 6$. Then the outlier detection procedures are applied on the AO-adjusted and IO-adjusted data for the second iteration. Table 6.2 gives the results for the second iteration.

Table 6.1: Test criteria of AO and IO detection procedures on the first iteration

Test Criteria			
AO		IO	
t	$\hat{\eta}_{AO}$	t	$\hat{\eta}_{IO}$
6	3.45	6	3.38

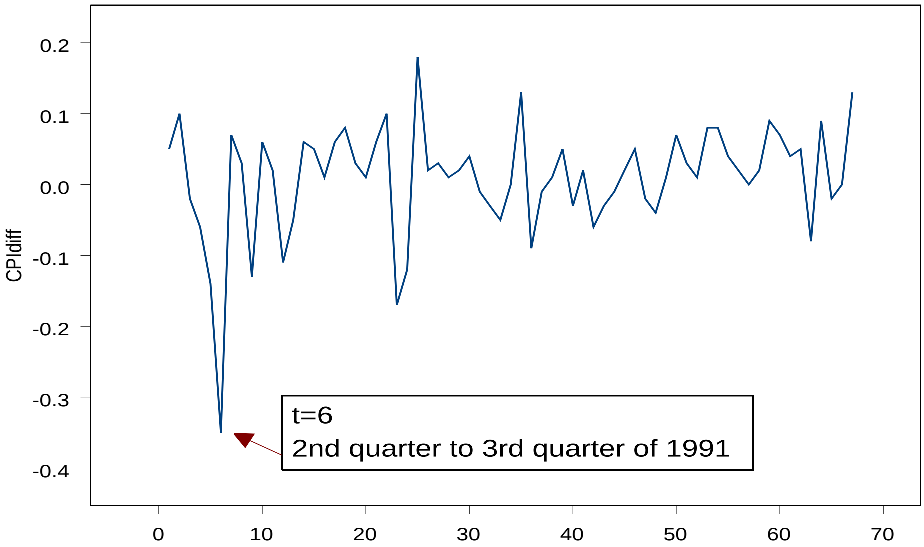

Figure 6.1: Detected outlier on the first iteration

When critical value C=3.0 is used, none of the test criteria exceeds the critical value. Thus, we stop the outlier detection procedure and conclude that no more outliers occurred in the differenced Indian CPI data.

The parameter estimations and its AIC values are calculated for the unadjusted, AO-adjusted and IO-adjusted data and are given in Table 6.3. The AIC values for the AO-adjusted and IO-adjusted data are significantly reduced compared to the AIC of the unadjusted data. The reduction of AIC is greater

Table 6.2: Test criteria of AO and IO detection procedures on the second iteration

Test Criteria			
AO-adjusted		**IO-adjusted**	
t	$\hat{\eta}_{AO}$	t	$\hat{\eta}_{IO}$
25	2.33	25	2.36

Table 6.3: Outlier effects in parameter estimations and AIC value for the Indian CPI data

	AIC	**Parameter**		
Removal at time $t = 6$		$\hat{\theta}$	$\hat{\sigma}_e^2$	$\hat{\sigma}_b^2$
unadjusted	-148.32	0.1771	0.0050	0.2138
AO-adjusted	-164.38	0.0834	0.0036	0.2014
IO-adjusted	-160.96	0.0679	0.0039	0.1671

after removing the AO effect compared to when removing the IO effect. Similarly, the parameter estimates also differ when the AO and IO effects are adjusted.

Further, we compute the next nine forecast values using the unadjusted, AO-adjusted and IO-adjusted model. As noted in Nicholls and Quinn (1982), a natural predictor of y_t from $\{y_{t-1}, y_{t-2}, ..\}$ is

$$y_t = sgn(\theta y_{t-1})[(\theta y_{t-1})^2 + \sigma_b^2 + \sigma_e^2 y_{t-1}^2]^{1/2}, \tag{6.1}$$

where

$$sgn(\theta y_{t-1}) = \begin{cases} 1 & \text{for} \quad \theta y_{t-1} \geq 0 \\ -1 & \text{for} \quad \theta y_{t-1} < 0. \end{cases}$$

Table 6.4: Forecasted values of Indian CPI data for fitted contaminated, uncontaminated AO and uncontaminated IO models

Period	**Observed values**	**Fitted Values**		
		unadjusted model	AO-adjusted model	IO-adjusted model
Q4 2006	0.09	0.0957	0.0845	0.0826
Q1 2007	0.26	0.0853	0.0715	0.0714
Q2 2007	0.11	0.0825	0.0685	0.0692
Q3 2007	0.15	0.0818	0.0678	0.0689
Q4 2007	-3.41	0.0816	0.0677	0.0688
$\sqrt{\frac{E}{5}}$		1.5638	1.5581	1.5586

Table 6.4 contains the observed values and forecast values using unadjusted, AO-adjusted and IO-adjusted models. The last row of Table 6.4 reports their error sum of squares, E. It can be seen that the fitted AO-adjusted and IO-adjusted model are better than that to the unadjusted model in a way that it has smaller AIC and error sum of squares, E. The procedures have detected an outlier at time $t = 6$ in the Indian CPI data set, and consequently improved the RCA(1) modeling and forecasting.

Chapter 7

Conclusion and Further Works

7.1 Summary of the Study

This study is aimed at developing outlier detection procedures for the RCA(1) model. We first proposed an iterative (IT) method based on estimating functions approach to estimate the RCA parameters. The robustness properties of the considered estimates, namely least squares, estimating functions and IT, when AO and IO exist in RCA(1) process have been investigated. We found out that the IT is the most robust estimate compared to the other two. We have also explored the nature of AO and IO effects on observations and residuals which were then used to derive the statistics for measuring these effects. Consequently, test statistics are defined to detect the presence of outliers in the RCA(1) process. A simulation study has been carried out to investigate the performance of the suggested procedures. In general, the procedures work well in detecting AO and IO. As an illustration, a real data set of differenced

Indian consumer price index is considered. These procedures have detected one outlier in the data set. It is further shown that the modeling of the data has improved by removing the effects of AO and IO from the data set.

7.2 Significance of the Study

This study has focused on the improvement of parameter estimation, the robustness of estimating functions and outlier detection procedures in RCA(1) model. It has contributed to the time series analysis in following ways:

1. Proposed an iterative estimation method based on the estimating functions to improve the RCA(1) parameter estimation.

2. Studied the robustness property of the least squares, estimating functions and the iterative estimates when outlier occurs in the data set.

3. Proposed the outlier detection procedures for the RCA(1) model.

7.3 Further Research

This research can be extended in many ways. We had studied the detection of AO and IO in the RCA(1) model. It can be extended to two other types of outliers, namely temporary change (TC) and level change (LC) and also the detection for non stationary RCA(1) process. We may also come up with a more comprehensive procedure which can identify the type of outliers by comparing the value of test statistics for all types of outliers. Moreover, an extension to higher order RCA models based on this approach can be further explored.

References

1. Abraham, B. and Box, G. E. P. (1979). Bayesian analysis of some outlier problems in time series , *Biometrika,* **66**, 229-236.

2. Abraham, B. and Chuang, A. (1989). Outlier Detection and Time Series Modeling. *Technometrics*, **31**, 241-248.

3. Andel, J. (1976). Autoregressive Series with Random Parameters. *Math. Operationsforsch. u. Statist.* **7**, 735-741.

4. Aue, A., Horváth, L., Steinebach, J. (2006). Estimation in Random Coefficient Autoregressive Models. *Journal of Time Series Analysis.* **27** (1), 61-76.

5. Bera, A. K., Garcia, P. and Roh, J. S. (1997). Estimation of Time-varying Hedge Ratios for Corn and Soybeans: BGARCH and Random Coefficient Approaches. [Online]. Available at http://www.ace.uiuc.edu /ofor/OFOR9706.pdf [accessed 11 March 2009].

6. Balke, N. S. and Fomby, T. B. (1994). Large Shocks, Small Shocks, and Economic Fluctuations: Outliers in Macroeconomic Time series. *Journal of Applied Econometrics* , **9** , 181-200.

7. Barnett, V. and Lewis, T. (1978). *Outliers in Statistical Data.* Wiley, Chichester.

8. Battaglia, F. and Orfei, L. (2005). Outlier Detection and Estimation in Nonlinear Time Series. *Journal of time series analysis* **26** (1), 107-121.

9. Boente, G., Fraiman,R., and Yohai, V. (1982). Qualitative robustness for general stochastic processes. *Annals of Statistics*, **15(3)**, 1293-1312.

10. Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, **31**, 307-327.

11. Box, G. E. P. and Tiao, G. C (1965). A Change in Level of Stationary Time Series. *Biometrika*, **52**, 181-192.

12. Campbell, M.J. and Walker, A.M (1977). A survey of statistical work on the Mackenzie River series of Annual Canadian lynx trappings for the years 1921-1934 and a new analysis. *J. Roy. Statist. Soc. A*, **140**, 411-431.

13. Chandra, S.A. and Taniguchi, M. (2001). Estimating Functions for nonlinear time series models. *Ann. Inst. Statist. Math.*, **53**(1), 125-136.

14. Chatfield, C. (1975). *The Analysis of Time Series : Theory and Practice.* Chapman and Hall, London.

15. Chang, I., Tiao, G. C. and Chen, C. (1988). Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics*, **40** (2), 193-204.

16. Charles, A. and Darne, O. (2005). Outliers and GARCH models in Financial Data. *Economics Letter*, **86**, 347-352.

17. 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**(421), 284-297.

18. Chick, Z. (1994). Comparing short time series using resampling methods. PhD Thesis, Loughborough University of Technology.

19. Conlisk, J. (1974). Stability in a Random Coefficient Model. *International Economic Review*, **15** (2), 529-533.

20. Conlisk, J. (1976). A Further Note on Stability in a Random Coefficient Model. *International Economic Review*, **17** (3), 759-764.

21. Durbin, J. (1960). Estimation of Parameters in Time Series Regression Models. *Journal of Royal Statistician Society* **22**, 139-153.

22. Engle, R. F (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation. *Econometrica*, 50, 987-1007.

23. Ghahramani, M. and Thavaneswaran, A. (2008). A Combining Estimating Functions for Volatility. *Journal of Statistical Planning and Inference*, **139** (4), 1449-1461.

24. Godambe, V.P. (1960). An Optimum Property of Regular Maximum Likelihood Estimation. *Annals of Mathematical Statistics*, **31**, 1208-1212.

25. Godambe, V.P. (1985). The foundations of finite sample estimation in stochastic processes. *Biometrika*, **72**, 419-428.

26. Granger, C. W. J. and Andersen, A. P. (1978). Introduction to Bilinear Time Series Model. Gottinge: Vandenhoeck and Ruprecht.

27. Feigin, P. D. and Tweedie, R. L. (1985) . Random coefficient autoregressive processes: A Markov chain analysis of stationarity and finiteness of moments. *Journal of Time Series Analysis*,**6**, 1-14.

28. Franses, P. H. and Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting*, **15**, 1-9.

29. Hampel, F. (1971). A general qualitative definition of robustness. *Annals of Mathematics and Statistics*, **24**, 1887-1896.

30. Harvey, D. I., L., Stephen J. and Newbold, P. (2001) . Innovational outlier unit root tests with an endogenously determined break in level .*Oxford Bulletin of Economics and Statistics*, **63**, 559-576.

31. Heyde, C. C. and Hannan, E. J. (1972). On limit theorems for quadratic functions of discrete series. *Ann. Math. Statist.* **43**, 2058-2066.

32. Hogg, R. V. (1979). An introduction to robust estimation. In *Robustness in Statistics: Proceeding of a Workshop* (ed. R. L. Launer and G. N. Wilkinson), New York: Academic Press, 1-17.

33. Huber, P.J. (2004). *Robust Statistics.* Wiley, New Jersey.

34. Hwang, S. Y. and Basawa, I. V. (1993). Parameter Estimation in a Regression Model with Random Coefficient Autoregressive Errors. *Journal of Statistical Planning and Inference* **36**, 57-67.

35. Ismail, M. I., Mohamed, I. and Yahya, M. S. (2008). Improvement on Additive Outlier Detection Procedure in Bilinear Model. *Malaysian Jurnal of Science*, **27**(2) , 107-117.

36. Last M. and Kendel A. (2001). Automated Detection of Outliers in Real-World Data. [Online]. Available at http://citeseerx.ist.psu.edu/viewdoc /summary?doi=10.1.1.15.8466 [accessed 11 March 2009].

37. McCulloch, R. E. and Tsay, R. S. (1994) . Bayesian analysis of autoregressive time series via the Gibbs sampler . *Journal of Time Series Analysis*, **15**, 235-250.

38. Martin, R. D. (1980). Robust estimation of autoregressive models. In *Direction in Time Series* (ed. Brillinger, D. R. and Tiao, G. C.), Hayward, CA: Institute of Mathematical Statistics, 228-254.

39. Mohamed, I. (2005). *Outliers in Bilinear time series model.* Ph.D. Universiti Teknologi MARA.

40. Moran, P. A. P. (1953). The Statistical Analysis of the Canadian Lynx Cycle *Aust. J. Zool.*, **1**, 163-173; 291-298.

41. Nicholls, D.F. and Quinn, B.G. (1980). The estimation of random coefficient autoregressive models. I. *Journal of time series analysis*, **1(1)**, 37-46.

42. Nicholls, D.F. and Quinn, B.G. (1981). The estimation of random coefficient autoregressive models. II. *Journal of time series analysis*, **2**, 185-203.

43. Nicholls, D.F. and Quinn, B.G. (1982). *Random coefficient autoregressive models: An introduction.* Lecture Notes in Statistics, 11, Springer-Verlag, New York.

44. Nicholls, D. F. (1986). The Box-Jenkins approach to random coefficient autoregressive modeling . *Journal of Applied Probability* **23(A)**, 231-240.

45. Papantoni-Kazakos, P., and Gary, R. M. (1979). Robustness of estimates and stationary observation. *Annals of Probability*, **7**, 989-1002.

46. Peirce, B. (1852). Criterion for the rejection of doubtful observations. *Astronomical Journal* **2**, 161-163.

47. Pena, D. (1990). Influenctial observations in time series. *Journal of Business Econom. Statist.* **8**, 235-241.

48. Perron, P. and Rodriguez, G. (2000). Searching for Additive Outliers In Stationary Time Series. *Journal of Time Series Analysis* **24** (2), 193-220.

49. Rahiala, M. (1999). Random Coefficient Autoregressive model for longitudinal data. *Biometrika*, **86(3)**, 718-722.

50. Schick, A. (1996). $\sqrt{n}$-Consistent Estimation in a Random Coefficient Autoregressive Model. *Austral. J. Statist.*, **38**(2), 155-160.

51. Thavaneswaran, A. and Abraham, B. (1988). Estimation for nonlinear time series model using estimating function. *J. Time Ser. Anal.*, **9**, 99-108.

52. Thavaneswaran, A. and Peiris, S. (1996). Nonparametric estimation for some nonlinear models. *Statistics and Probability Letters*, **28**, 227-233.

53. Thavaneswaran, A. and Peiris, S. (2001). Inference for some time series models with random coefficients and infinite variance innovations. *Mathematical and Computer Modeling*, **33**, 843-849.

54. Thavaneswaran, A. and Peiris, S. (2003). Generalized smoothed estimating Function for nonlinear time series. *Statistics and Probability Letters*, **65**, 51-56.

55. Thavaneswaran, A. and Peiris, S. (2004). Smoothed estimates for models with random coefficients and infinite variance innovations. *Mathematical and Computer Modeling*, **39**, 363-372.

56. Tong, H. (1977). Some Comments on the Canadian Lynx Data. *J. Roy. Statist. Soc.*, A, **140**, 432-436.

57. Tong, H. (1983). *Threshold Models in non-linear time series analysis.* New York: Springer Verlag.

58. Tiao, G. C. and Tsay, R. S. (1994). Some advances in non-linear and adaptive modeling in time-series. *Journal of Forecasting* **13**, 109-131 .

59. Tjøstheim, D. (1986). Estimation in Nonlinear Time Series Models. *Stochastic Processes and their Applications* **21**, 251-273.

60. Tsay, R.S. (1986). Time Series Model Specification In the Presence of Outliers. *Journal of the American Statistical Association*, **81**, 132-141.

61. Tukey, J.W. (1976). The present state of robust/resistant techniques. Talk given at the annual meeting of the Psychometric Society, bell Laboratories, Murray Hill, N. J.

62. Turnovsky, S. J. (1968). Stochastic Stability of Short-run Market Equilibrium Under Variation in Supply. *Quarterly Journal of Economics* **LXXXII**, 666-681.

63. Wang, D. and Ghosh, S. K. (2002). Bayesian Analysis of Random Coefficient Autoregressive Models. *Model Assisted Statistics and Applications.*, **3**(2), 281-295.

64. Wright, T. W. (1884). A treastise on the adjustment of observations by the method of least squares. New York: Van Nostrand.

65. Vogelsang, T. J. (1999). Two Simple Procedures for Testing for a Unit Root when there are Additive Outliers. *Journal of Time Series Analysis* **20**, 237-252.

66. Zaharim, A., Mohamed, I., Ahmad, I., Abdullah, S. and Omar, M. Z, (2006). Performances test statistics for single outliers detection in Bilinear (1,1,1,1) models. *Wseas Transactions on Mathematics* **5** (12), 1359-1364.

Appendix A

Indian Consumer Price Index (CPI) Data

Period	CPI	Period	CPI	Period	CPI	Period	CPI	Period	CPI
Q1 1990	2.23	Q3 1993	1.76	Q1 1997	2.11	Q3 2000	2.10	Q1 2004	2.41
Q2 1990	2.28	Q4 1993	1.81	Q2 1997	2.13	Q4 2000	2.07	Q2 2004	2.41
Q3 1990	2.38	Q1 1994	1.82	Q3 1997	2.17	Q1 2001	2.06	Q3 2004	2.43
Q4 1990	2.36	Q2 1994	1.88	Q4 1997	2.16	Q2 2001	2.08	Q4 2004	2.52
Q1 1991	2.3	Q3 1994	1.96	Q1 1998	2.13	Q3 2001	2.13	Q1 2005	2.59
Q2 1991	2.16	Q4 1994	1.99	Q2 1998	2.08	Q4 2001	2.11	Q2 2005	2.63
Q3 1991	1.81	Q1 1995	2.00	Q3 1998	2.08	Q1 2002	2.07	Q3 2005	2.68
Q4 1991	1.88	Q2 1995	2.06	Q4 1998	2.21	Q2 2002	2.08	Q4 2005	2.6
Q1 1992	1.91	Q3 1995	2.16	Q1 1999	2.12	Q3 2002	2.15	Q1 2006	2.69
Q2 1992	1.78	Q4 1995	1.99	Q2 1999	2.11	Q4 200	2.18	Q2 2006	2.67
Q3 1992	1.84	Q1 1996	1.87	Q3 1999	2.12	Q1 2003	2.19	Q3 2006	2.67
Q4 1992	1.86	Q2 1996	2.05	Q4 1999	2.17	Q2 2003	2.27	Q4 2006	2.8
Q1 1993	1.75	Q3 1996	2.07	Q1 2000	2.14	Q3 2003	2.35		
Q2 1993	1.70	Q4 1996	2.1	Q2 2000	2.16	Q4 2003	2.39		

Appendix B

S-plus Program

a) Parameter estimation, aic and box-pierce for real data

```
EstAicBoxNew<-function(y,iter){
    #EstAicBoxNew(india$CPIdiff,50)
    n<-length(y)
#LS
    est<-EstLSIT(n,y,iter)
    thetaLS<-est$thetaLS
    var.b<-est$var.b
    var.e<-est$var.e
#EF
    EFplease<-EstEF(n,y,thetaLS)
    thetaEF<-EFplease$thetaEF
    var.bEF<-EFplease$var.bEF
    var.eEF<-EFplease$var.eEF
```

```
#IT
    thetaIT<-est$thetaIT
    var.bIT<-est$var.bIT
    var.eIT<-est$var.eIT
    Theta<-cbind(thetaLS,thetaEF,thetaIT)
    Var.b<-cbind(var.b,var.bEF,var.bIT)
    Var.e<-cbind(var.e,var.eEF,var.eIT)
#aic
    lnLS<-matrix(0,nrow=n)
    lnEF<-matrix(0,nrow=n)
    lnIT<-matrix(0,nrow=n)
    aicLS<-matrix(0,nrow=n)
    aicEF<-matrix(0,nrow=n)
    aicIT<-matrix(0,nrow=n)
        for(i in 2:n){
            lnLS[i]<-log(var.e+(var.b*y[i-1]^2))
            lnEF[i]<-log(var.eEF+(var.bEF*y[i-1]^2))
            lnIT[i]<-log(var.eIT+(var.bIT*y[i-1]^2))
            aicLS[i]<-((y[i]-(thetaLS*y[i-1]))^2)/(var.e+(var.b*y[i-1]^2))
            aicEF[i]<-((y[i]-(thetaEF*y[i-1]))^2)/(var.eEF+(var.bEF*y[i-1]^2))
            aicIT[i]<-((y[i]-(thetaIT*y[i-1]))^2)/(var.eIT+(var.bIT*y[i-1]^2))
        }
    AICls<-(n*log(2*pi))+sum(lnLS)+sum(aicLS)+(2*3)
    AICef<-(n*log(2*pi))+sum(lnEF)+sum(aicEF)+(2*3)
```

```
    AICit<-(n*log(2*pi))+sum(lnIT)+sum(aicIT)+(2*3)
    aic<-cbind(AICls,AICef,AICit)
#box-pierce
    uLS<-matrix(0,nrow=n)
    uEF<-matrix(0,nrow=n)
    uIT<-matrix(0,nrow=n)
    hLS<-matrix(0,nrow=n)
    hEF<-matrix(0,nrow=n)
    hIT<-matrix(0,nrow=n)
    eLS<-matrix(0,nrow=n)
    eEF<-matrix(0,nrow=n)
    eIT<-matrix(0,nrow=n)
      for(i in 2:n){
        uLS[i]<-y[i]-(thetaLS*y[i-1])
        uEF[i]<-y[i]-(thetaEF*y[i-1])
        uIT[i]<-y[i]-(thetaIT*y[i-1])
        hLS[i]<-var.e+(var.b*y[i-1]^2)
        hEF[i]<-var.eEF+(var.bEF*y[i-1]^2)
        hIT[i]<-var.eIT+(var.bIT*y[i-1]^2)
        eLS[i]<-uLS[i]/sqrt(hLS[i])
        eEF[i]<-uEF[i]/sqrt(hEF[i])
        eIT[i]<-uIT[i]/sqrt(hIT[i])
      }
    autocorrLS<-acf(eLS,20,"correlation")$acf
```

```
autocorrEF<-acf(eEF,20,"correlation")$acf
autocorrIT<-acf(eIT,20,"correlation")$acf
autocorrLS2<-autocorrLS^2
autocorrEF2<-autocorrEF^2
autocorrIT2<-autocorrIT^2
#minus 1 bcz lag0=1
boxPls<-n*(sum(autocorrLS2)-1)
boxPef<-n*(sum(autocorrEF2)-1)
boxPit<-n*(sum(autocorrIT2)-1)
boxP<-cbind(boxPls,boxPef,boxPit)
criticalValue<-qchisq(0.95, 20-3)
#residual plot
uLSplot<-data.frame(uLS)
uEFplot<-data.frame(uEF)
uITplot<-data.frame(uIT)
list(Theta=Theta,Var.b=Var.b,Var.e=Var.e,aic=aic,boxP=boxP,
criticalValue=criticalValue,uLSplot=uLSplot,uEFplot=uEFplot,uITplot=uITplot)
}
#to estimate LS and IT
EstLSIT<-function(n,y,iter){
#LS
thetaLS<- sum( y[2:n]*y[1:(n-1)] ) / sum( (y[1:(n-1)]^2) )
u<-matrix(0,nrow=n)
for(i in 2:n){
```

```
          u[i]<-y[i]- (thetaLS*y[i-1])
     }
     z<- sum(y[1:(n-1)]^2)/(n-1)
     var.b<- sum( (u[2:n]^2)*( (y[1:(n-1)]^2)-z ) )/ sum(((y[1:(n-1)]^2)-z)^2)
     var.e<- ( sum(u[2:n]^2)/(n-1) )- (var.b*z)
#EF
     a<-matrix(0,nrow=n)
        for(i in 2:n)
          a[i-1]<–y[i-1]/ ( var.e+ (var.b*(y[i-1]^2)) )
        }
     thetaEF<-sum( a[1:(n-1)]*y[2:n] )/sum( a[1:(n-1)]*y[1:(n-1)] )
#IT
     thetaIT<-matrix(0,nrow=iter)
     var.bIT<matrix(0,nrow=iter)
     var.eIT<-matrix(0,nrow=iter)
     thetaIT[1]<-thetaEF
        for(j in 2:iter){
           ITplease<-EstEF(n,y,thetaIT[j-1])
           thetaIT[j]<-ITplease$thetaEF
           var.bIT[j]<-ITplease$var.bEF
           var.eIT[j]<-ITplease$var.eEF
           final<-j
          if( (abs(thetaIT[j]-thetaIT[j-1])<=0.000001) && (abs(var.bIT[j]-
```

```
var.bIT[j-1])<=0.000001) && (abs(var.eIT[j]-var.eIT[j-1])<=0.000001) )

break

if(final==iter)thetaIT[j]<-NA

if(final==iter)var.bIT[j]<-NA

if(final==iter)var.eIT[j]<-NA

}

thetaIT<-thetaIT[final]

var.bIT<-var.bIT[final]

var.eIT<-var.eIT[final]

list(thetaLS=thetaLS,var.b=var.b,

var.e=var.e,thetaIT=thetaIT,var.bIT=var.bIT,var.eIT=var.eIT)

}

#to estimate EF and iteration for IT

EstEF<-function(n,y,thetaLS){

u<-matrix(0,nrow=n)

for(i in 2:n){

u[i]<-y[i]- (thetaLS*y[i-1])

}

z<- sum(y[1:(n-1)]^2)/(n-1)

var.bEF<- sum( (u[2:n]^2)*( (y[1:(n-1)]^2)-z ) )/

sum(((y[1:(n-1)]^2)-z)^2)

var.eEF<- ( sum(u[2:n]^2)/(n-1) )- (var.bEF*z)

#EF

a<-matrix(0,nrow=n)
```

```
        for(i in 2:n){
            a[i-1]<-y[i-1]/ ( var.eEF+ (var.bEF*(y[i-1]^2)) )
        }
    thetaEF<-sum( a[1:(n-1)]*y[2:n] )/sum( a[1:(n-1)]*y[1:(n-1)] )
    list(thetaEF=thetaEF,var.bEF=var.bEF,var.eEF=var.eEF)
}
```

b) AO and IO detection

```
#ao and io detection for real data
detectionAOIO¡-function(y,iter)
        #detectionAOIO(india$CPIdiff,50)
        n¡-length(y)
#LS
        est<-EstLSIT(n,y,iter)
        thetaLS<-est$thetaLS
        var.b<-est$var.b
        var.e<-est$var.e
#EF
        EFplease<-EstEF(n,y,thetaLS)
        thetaEF<-EFplease$thetaEF
        var.bEF<-EFplease$var.bEF
        var.eEF<-EFplease$var.eEF
#IT
        thetaIT<-est$thetaIT
        var.bIT<-est$var.bIT
        var.eIT<-est$var.eIT
#detection
        uIT<-matrix(0,nrow=n)
            for(i in 2:n){
                uIT[i]<-y[i]-(thetaIT*y[i-1])
```

```
    }
omegaAO<-matrix(0,nrow=n)
varOmegaAO<-matrix(0,nrow=n)
tauAO<-matrix(0,nrow=n)
omegaIO<-matrix(0,nrow=n)
varOmegaIO<-matrix(0,nrow=n)
tauIO<-matrix(0,nrow=n)
    for(i in 1:n){
        omegaAO[i]<-(uIT[i]-(thetaIT*uIT[i+1]))/(1+(thetaIT^2))
        varOmegaAO[i]<-( (var.eIT*((thetaIT^2)+1)) + (var.bIT*(((thetaIT^2)
        *(y[i]^2))+(y[i-1]^2))) )/ ((1+(thetaIT^2))^2)
        tauAO[i]<- abs(omegaAO[i]/sqrt(varOmegaAO[i]))
        omegaIO[i]<-uIT[i]
        varOmegaIO[i]<-var.eIT+(var.bIT*(y[i-1]^2))
        tauIO[i]<- abs(omegaIO[i]/sqrt(varOmegaIO[i]))
    }
tauAO<-as.numeric(tauAO)
tt<-rep(1:n,1)
a<-cbind(tauAO,tt)
a<-na.exclude(a)
aSort<-sort.col(a,c("<ALL>"),"tauAO",F)
maxTauAO<-aSort[1,1]
timeMaxTauAO<-aSort[1,2]
tauIO<-as.numeric(tauIO)
```

```
        b<-cbind(tauIO,tt)

        b<-na.exclude(b)

        bSort<-sort.col(b,c("<ALL>"),"tauIO",F)

        maxTauIO<-bSort[1,1]

        timeMaxTauIO<-bSort[1,2]

        ao<-cbind(maxTauAO,timeMaxTauAO)

        io<-cbind(maxTauIO,timeMaxTauIO)

        omegaAO<-omegaAO[timeMaxTauAO]

        omegaIO<-omegaIO[timeMaxTauIO]

        list(ao=ao,io=io,omegaAO=omegaAO,omegaIO=omegaIO)

    }

#to estimate LS and IT

EstLSIT<-function(n,y,iter){

#LS

        thetaLS<- sum( y[2:n]*y[1:(n-1)] ) / sum( (y[1:(n-1)]^2) )

        u<-matrix(0,nrow=n)

            for(i in 2:n){

                u[i]<-y[i]- (thetaLS*y[i-1])

        }

        z<- sum(y[1:(n-1)]^2)/(n-1)

        var.b<- sum( (u[2:n]^2)*( (y[1:(n-1)]^2)-z ) )/ sum(((y[1:(n-1)]^2)-z)^2)

        var.e<- ( sum(u[2:n]^2)/(n-1) )- (var.b*z)

#EF

        a<-matrix(0,nrow=n)
```

```
        for(i in 2:n)
            a[i-1]<-y[i-1]/ ( var.e+ (var.b*(y[i-1]^2)) )
        }
    thetaEF<-sum( a[1:(n-1)]*y[2:n] )/sum( a[1:(n-1)]*y[1:(n-1)] )
#IT
    thetaIT<-matrix(0,nrow=iter)
    var.bIT<matrix(0,nrow=iter)
    var.eIT<-matrix(0,nrow=iter)
    thetaIT[1]<-thetaEF
        for(j in 2:iter){
            ITplease<-EstEF(n,y,thetaIT[j-1])
            thetaIT[j]<-ITplease$thetaEF
            var.bIT[j]<-ITplease$var.bEF
            var.eIT[j]<-ITplease$var.eEF
            final<-j
            if( (abs(thetaIT[j]-thetaIT[j-1])<=0.000001) && (abs(var.bIT[j]-

            var.bIT[j-1])<=0.000001) && (abs(var.eIT[j]-var.eIT[j-1])<=0.000001) )
            break
            if(final==iter)thetaIT[j]<-NA
            if(final==iter)var.bIT[j]<-NA
            if(final==iter)var.eIT[j]<-NA
        }
    thetaIT<-thetaIT[final]
```

```
    var.bIT<-var.bIT[final]

    var.eIT<-var.eIT[final]

    list(thetaLS=thetaLS,var.b=var.b,

    var.e=var.e,thetaIT=thetaIT,var.bIT=var.bIT,var.eIT=var.eIT)

}

#to estimate EF and iteration for IT

EstEF<-function(n,y,thetaLS){

    u<-matrix(0,nrow=n)

        for(i in 2:n){

            u[i]<-y[i]- (thetaLS*y[i-1])

        }

    z<- sum(y[1:(n-1)]^2)/(n-1)

    var.bEF<- sum( (u[2:n]^2)*( (y[1:(n-1)]^2)-z ) )/

    sum(((y[1:(n-1)]^2)-z)^2)

    var.eEF<- ( sum(u[2:n]^2)/(n-1) )- (var.bEF*z)

#EF

    a<-matrix(0,nrow=n)

        for(i in 2:n){

            a[i-1]<–y[i-1]/ ( var.eEF+ (var.bEF*(y[i-1]^2)) )

        }

    thetaEF<-sum( a[1:(n-1)]*y[2:n] )/sum( a[1:(n-1)]*y[1:(n-1)] )

    list(thetaEF=thetaEF,var.bEF=var.bEF,var.eEF=var.eEF)

  }
```