Application of Several Time Series Methods to Three Important Financial Time Series

Bryan O'Connell

Supervisor: Prof C. Koen

February 2007

Acknowledgement

I would like to thank the following people:
Professor Chris Koen, my parents Judith and Bryan O'Connell, my sister Amanda-Leigh and Leigh Phillips.

Abstract

This study is concerned with three different financial time series over an eight year period, namely: the government repurchase rate, the Rand-Dollar exchange rate and the Allshare Index. The aim is to better understand the statistical nature of the time series. The theory employed will be discussed briefly and then the results will be reported. Different methods are employed to model the different time series. The following topics are discussed: unit root tests, autoregressive integrated moving average models, outlier tests, transformations, generalised autoregressive conditional heteroscedasticity models, cointegration, transfer function models and vector autoregressive models.

Declaration

I declare that Application of Several Time Series Methods to Three Important Financial Time Series is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Full name Date

Signed

Contents

Chapter 1

Introduction

Financial time series generally exhibit what appears to be random behaviour. The purpose of this study is to try and develop models to adequately describe this behaviour in a few example cases. Three financial time series will be analysed, namely the Allshare Index (Alsi Index), the government Repurchase Rate (Repo Rate) and the Rand-Dollar (USDZAR) exchange rate.

The data were obtained from the Data Stream computer at the University of Cape Town. The time series were recorded over an eight year period from the 9_{th} of March, 1998 to the 31_{st} of December, 2005. The Alsi and USDZAR data were recorded at the close of each trading day for that period. The Repo series used here consists of the changes in the Repo rate. The time intervals between changes are not constant so the Repo series is not indexed by time.

Each of the three time series plays an important role in the South African economy. The ability to accurately model the behaviour of these time series is of great value to economic policy makers, traders, business enterprises as well as local and foreign investors.

The stock market is the best leading indicator of the state of the economy. It is a direct measure of the level of investor confidence in the economy. The Johannesburg Stock Exchange (JSE) is a highly liquid market. When economic indicators convey perceived negative news, the stock market is the first to react as traders try to sell their shares. The Alsi Index consists of the shares of each of the companies on the JSE, each given a weighting determined by the company's market value. It is therefore the best indicator of the performance of the JSE (Bodie *et al* 2005 [8]). When the Alsi Index performs well it instills confidence in foreign investors, on the state of the economy. This leads to increased foreign direct investment which strengthens the Rand-Dollar exchange rate.

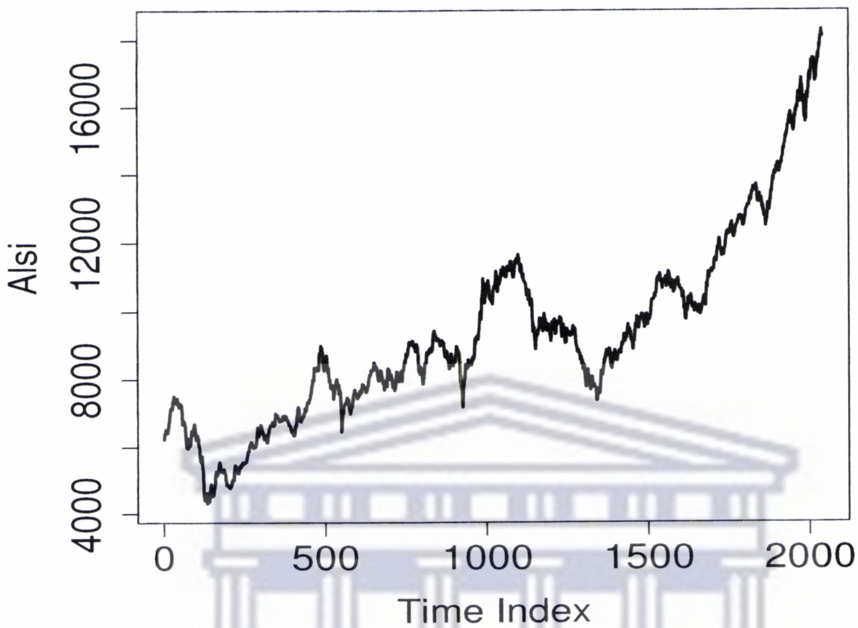

Figure 1.1: Daily closing values of Alsi

Figure 1.1 shows that the Alsi data contain a main uptrend which persists for the duration of the time series. This is partly due the steady increase in the price of gold over the last 5 years. Also, consumer demand has been growing at record levels due to a steady decline in the borrowing rates offered to businesses and individuals. The decline in the Repo rate increases the price of bonds. This has a positive effect on the stock market as investors look to obtain higher yields [8].

The Rand-Dollar exchange rate shown in figure 1.2 below, is the most important currency exchange rate pair that is monitored in South Africa. The South African economy is dominated by mining exports. All minerals are sold for U.S. Dollars and all expenses are incurred in Rands. A weaker Rand-Dollar exchange rate has a positive effect on the Alsi Index as mining stocks have the largest weighting on the Index.

The Rand-Dollar exchange rate also affects the South African current account. If the Rand-Dollar exchange rate is weaker South African exports are more attractive to for-

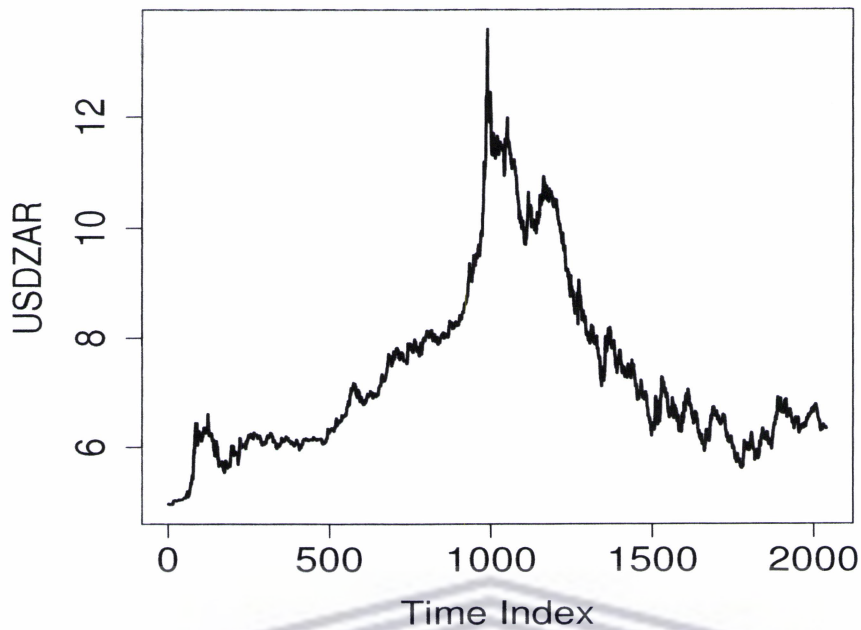

Figure 1.2: Daily closing values of USDZAR

eigners. Conversely foreign imports are relatively more expensive to South Africans. This leads to increased exports and decreased imports that shrink the current account deficit. A current account deficit is seen as an indication of weak macroeconomic policies (Nattrass 1997 [9]). Thus a large current account deficit often has a negative effect on foreign investor confidence, which is an important economic driver. From the sequence plot we can see that the Rand-Dollar exchange rate contains an initial uptrend followed by a downtrend. The change in trend can be attributed to the strengthening of the South African economy with respect to inflation targeting, consumer demand and the gold price, which has attained record highs in the last 5 years. Also, since 2002 the United States of America (U.S.A.) has had to spend billions on the war in Iraq. The U.S.A. has had a trade deficit and a budget deficit. This means the U.S.A. is borrowing to finance its expenditure and spending most of its income on foreign goods. These two factors have helped to weaken the U.S. Dollar against nearly every currency, the Rand included.

The Repo rate plays an important role in determining the level of consumer spending in the South African economy. This rate is set by the South African Reserve Bank (SARB). It is the rate at which commercial banks can borrow money from the SARB. Therefore, the Repo rate governs the rate at which commercial banks can lend money to businesses and private individuals. The rate at which individuals and businesses can obtain credit directly influences their ability to spend money in the economy. The Repo rate also influences exchange rates (Lipsey *et al* 1993 [10]). Interest rate differentials between South Africa and foreign countries either cause capital to flow into the South African economy or South African capital to flow into foreign economies, depending on whether the differential is positive or negative. Capital inflows strengthen the Rand against foreign currencies while capital outflows weaken it. It would be of great value if this relationship could be modelled. This would enable the SARB to test the effect a change in the Repo rate would have on the Rand-Dollar exchange rate and therefore the Alsi Index and the economy as a whole.

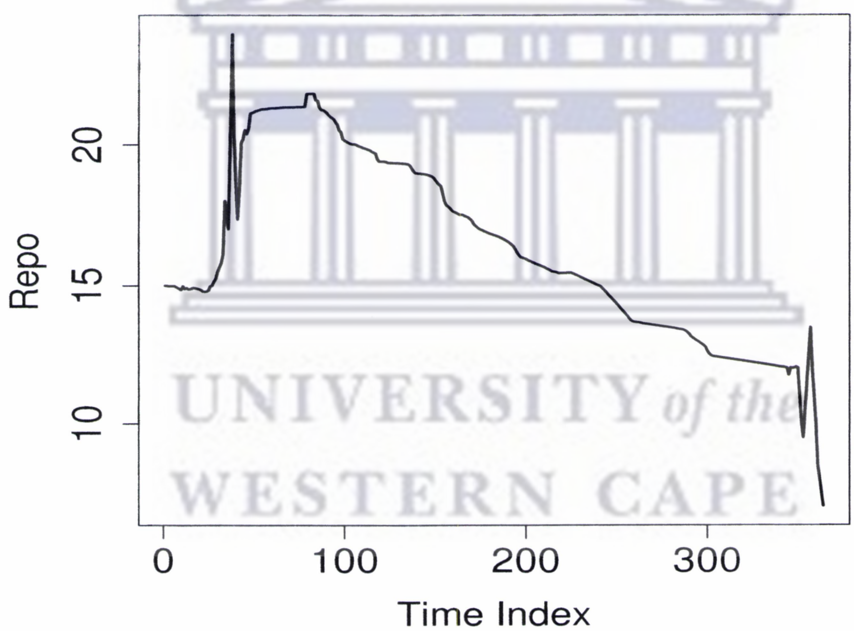

Figure 1.3: The value of the Repo series, indexed by times of change

Figure 1.3 shows that the general trend of the Repo Rate has been downward. This is partly due to the government's aim to boost consumption demand. It is also partly due to the benign inflation environment. The Repo rate is the main tool at SARB's disposal to combat inflation.

It is clear that the forces driving the level of each of the three time series are both varied and immense. These forces include the inflation, money supply, business confidence, consumer confidence, supply and demand, the unemployment rate and many more. Furthermore, the time series are expected to be interrelated. The USDZAR and Alsi time series were both recorded at the close of the trade on a daily basis. These time series are of equal length so the relationship between them can be investigated using, for example, cointegration analysis. As mentioned before, the Repo time series is different from the other two series because the time between observations is not constant. The government repurchase rate is announced at irregular intervals and often there is no change in it for several months. So the time series which is modelled in this study consists only of these changes in the repurchase rate. As a result the series is much shorter than the other two. This poses a particular problem when trying to analyse the relationship between the Repo and the USDZAR, or the Alsi time series. With so many variables at work it is no wonder that one cannot find simple deterministic relationships between the possible input variables and the time series. This study seeks to investigate the structure and interrelationships of the three time series described above. In general, time series are varied so it is not uncommon to find that methods used to model one time series do not work when applied to another. Each time series is unique and provides its own set of challenges to the modelling process. This study aims to engage these challenges, where possible, and to suggest possible methods that may be researched when it is not possible.

1.1 General Methodology

According to Mabert *et al* 1974 [18], the Box-Jenkins methodology applied in this study outperforms traditional forecasting methods because there is a structured approach to model building. In the latter case model forms are often arbitrarily chosen, with a great emphasis placed on modelling experience. In the former case a general model form is developed through various analyses and is then filtered using various diagnostic measures to obtain the specific model used for forecasting.

The Box-Jenkins methodology applies only to data that are at least weakly stationery. If the data contain a unit root, i.e. the mean follows a random walk, then it must be transformed before analysis can be performed. Franses 1998 [1] recommends performing the Augmented Dickey Fuller (ADF) test for the presence of a unit root. This requires the calculation of a t-statistic for the parameter ρ in the auxiliary regression

$$\Delta_1 y_t = \rho y_{t-1} + \alpha_1^* \Delta_1 y_{t-2} + ... + \alpha_{p-1}^* \Delta_1 y_{t-(p-1)} + \epsilon_t \tag{1.1}$$

to test the null hypothesis that $\rho = 0$, which means a unit root is present, versus the alternate hypothesis of $\rho < 0$, corresponding to no unit root present in the data. The method of unit root testing proposed by Enders 1948 [5], which modifies the ADF statistic to allow for correlation and heteroscedastic variance in the data, is used in this study.

Once the data are transformed to stationarity the modelling process may be hampered further by the presence of aberrant observations in the data. Franses [1] believes that outliers convey important information about the time series. Any outliers should be modelled using appropriate intervention analysis techniques. Box *et al* 1994 [6] state that outliers do not form part of the general time series for which an appropriate model needs to be created. In this study the outliers are removed and analysis conducted on the series containing missing values.

In financial time series it is not uncommon to observe clusters of outliers. This, of course, points to the presence of autoregressive heteroscedastic variance or ARCH. This can adequately be modelled using generalised ARCH (G(ARCH)) models. According to Bollerslev *et al* 1992 [19], G(ARCH)(1,1) models are fairly successful at modelling the ARCH in financial time series.

As was previously mentioned the interrelationship between financial time series and, in particular, these three time series, is of great importance in any well functioning economy. Pierce 1977 [20] found that many economic variables perceived to be interrelated can be regarded as weakly interrelated or even independent based on empirical studies. He goes on to state that due to limitations in the data, even when relationships do exist, econometric or empirical means can not ascertain their existence.

A summary of the steps taken in the analysis is given below.

(1) Consider possible transformations (typically logarithmic) of the time series.

(2) Generate the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the series.

(3) Perform a unit root test.

(4) Transform time series to mean stationarity if necessary. Plot the sequence diagram, ACF and PACF of transformed series.

(5) Fit possible Autoregressive Integrated Moving Average (ARIMA) models. Perform diagnostic checks.

(6) Test for outliers.

(7) Test for Autoregressive Conditional Heteroscedasticity (ARCH). Fit Generalised ARCH (G(ARCH)) models.

(8) Perform cointegration analysis.

(9) Fit possible Transfer Function Models. Perform diagnostic checks.

(10) Fit possible VAR Models.

Box-Jenkins methodology (see Box *et al* 1994 [6]) is used when fitting ARIMA models. A list of the software used in this study is provided in Appendix 1.
Once a group of tentative models has been identified, various techniques, which will be discussed shortly, will be used to determine and compare the adequacy of the models. Of course there is no guarantee that an adequate model will be obtained using these procedures. Time series each have their own unique characteristics, such as outliers and nonstationarity, which cause difficulty in the model building procedure. In this case various tools will have to be utilised to overcome the problems posed by the particular time series.

1.2 Some Useful Definitions

An ARMA(p,q) time series is defined by

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \tag{1.2}$$

where ϕ_i, θ_i and μ are constants and ϵ_t is a zero mean white noise process with constant variance σ_ϵ^2.

A differenced time series z_t is defined as

$$z_t = \Delta y_t = y_t - y_{t-1} = y_t - By_t = (1-B)y_t \tag{1.3}$$

where B is a backshift operator that satisfies the following equation

$$B^j y_t = y_{t-j} \tag{1.4}$$

Equation (1.2) can also be written more concisely, using the backshift operator, as

$$\Phi(B)y_t = \mu + \Theta(B)\epsilon_t \tag{1.5}$$

where $\Phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\Theta(B) = 1 + \sum_{i=1}^{q} \theta_i B^i$.

A particular nonstationary variant of (1.2) is also useful:

$$y_t = \mu + \delta t + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \tag{1.6}$$

where δ is a constant. This form may be necessary because the differenced series may require a non-zero constant.

A G(ARCH)(p,q) model is defined by:

$$E(\epsilon_t^2 | I_{t-1}) = \zeta_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \zeta_{t-i} \tag{1.7}$$

where $\epsilon_t = \upsilon_t \sqrt{\zeta_t}$ and υ_t is independently distributed with zero mean and variance of one, I_{t-1} is all the information up to time t-1 and α_i and β_i are constants satisfying certain restrictions (see e.g. Franses 1998 [1]).

The transfer function model is defined by

$$y_t = C + (\phi(B)/\theta(B))x_{t-\tau} + N_t \tag{1.8}$$

where y_t is the output series, C is a constant, $\phi(B)/\theta(B)$ is an ARMA model, $x_{t-\tau}$ is the input series with lag interval τ and N_t is a noise term which may or may not be autocorrelated (see Pankratz 1991 [14]).

A two dimensional VAR model of order p is of the following form:

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + \Sigma_t \tag{1.9}$$

where Y_t is a 2×1 matrix consisting of the t_{th} observation of each of the two time series being analysed, Φ_i is a 2×2 coefficient matrix and Σ_t is a 2×1 matrix consisting of the noise terms that are uncorrelated for different times - see Hamilton 1994 [7].

The Jarque-Bera test statistic, which measures the degree of non-normality in the residuals, is defined as

$$JB = n/6(s^2 + (k-3)^2/2) \tag{1.10}$$

where s is the sample skewness, k is the kurtosis and n is the time series length (see Jarque and Bera 1980 [25]).

The Akaike (AIC) and Schwartz's Bayes (BIC) Information criteria, which are useful for model selection, are given by:

$$AIC = -2 * (LL) + 2 * (K) \tag{1.11}$$

$$BIC = -2 * (LL) + log(n) * (K) \tag{1.12}$$

where K = the number of parameters in the model, n = the number of observations and LL is the log-likelihood of the model (see Burnham and Anderson 2002 [12]).

The ACF of a time series is defined as

$$\rho(k) = \gamma(k)/\gamma(0) \tag{1.13}$$

where $\gamma(k)$ is the autocovariance function

$$\gamma(k) = E[(y_t - E(y_t))(y_{t+k} - E(y_{t+k}))] \tag{1.14}$$

The sample autocovariance is calculated as follows

$$\widehat{\gamma}(k) = (1/N) * \sum_{j=1}^{N-k} [(y_j - \overline{y}) * (y_{j+k} - \overline{y})] \tag{1.15}$$

and the sample ACF is given by

$$\widehat{\rho}(k) = \widehat{\gamma}(k)/\widehat{\gamma}(0) \tag{1.16}$$

For the duration of this study the sample ACF will be referred to as simply the ACF.

Values obtained for autocorrelation estimates are deemed significant if they exceed $2/\sqrt{T}$ in absolute value, i.e. two times the standard error of the estimate. Simultaneous evaluation of a collection of estimated autocorrelations, by so-called portmanteau Q statistics, is also useful. The formula for calculating Box-Ljung Q statistics is given by

$$Q^* = T(T+2) * \sum_{i=1}^{L} (T-i)^{-1} \widehat{\rho}^{\,2}(i) \tag{1.17}$$

where L is the number of lags over which the Q^* is calculated. In this study, given the considerable length of the time series, the first 50 lags are analysed.

The sample PACF is also used throughout this study. It is henceforth referred to as simply the PACF. According to Shumway 1988 [23], the PACF provides the corrected autocorrelation between two observations in a given time series, conditional on all the intervening values of the series. The PACF provides information on the possible number of autoregressive components to include in the model.

1.3 Transformations

It is often simpler to model a given time series once it has been transformed. Common transformations include taking the square root or the logarithm of each observation. Transformations aim to homogenise or reduce the heteroscedasticity in the series and therefore assist in the modelling process. Often financial time series are log-transformed prior to analysis. This is because the differenced log transformed price of an asset has meaning, namely the continuously compounded rate of return for holding that asset (see Baillie *et al* 2001 [21]). In this study the log-transformed series will be used unless otherwise stated.

Before proceeding it must be stated that the above mentioned Box-Jenkins methodology only applies to weakly stationary time series, that is time series with a constant mean and covariance structure. So, before any ARMA modelling can take place, we will have to test whether or not the data are at least weakly stationary. Minimum requirements are testing for the presence of a unit root in the autoregressive polynomial

$$(1 - \Sigma_{j=1}^{p} \phi_j B^j) y_j \tag{1.18}$$

where B is the backshift operator (e.g. Graupe 1989 [4]). As was mentioned before, the Phillips-Perron (PP) test for unit roots is used in this study. This test is based on the Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) tests for unit roots in the sense that the PP test statistic is a DF test statistic modified to make the test more general [5]. The DF and ADF tests are not used because they require that certain assumptions be made about the residuals, namely that the residuals are uncorrelated and that they have homoscedastic variances. The PP test allows residuals to be serially correlated and to be heteroscedastic. Financial time series often exhibit heteroscedastic variance. According to Maddala and Kim 1998 [2] the PP test modifies the t-statistic from the DF test, in order to correct for the aforementioned serial autocorrelation and heteroscedasticity of the variance. For a detailed description of how this is done see Hamilton [7].
All three tests start with the fitting of an AR(1) model,

$$y_t = \phi_1 y_{t-1} + \epsilon_t \tag{1.19}$$

to the time series. They then proceed in different manners to ultimately test the hypothesis $\phi_1 = 1$ versus the hypothesis that $\phi_1 < 1$. It is noteworthy that the tests are one-sided. If it is decided that a unit root is present in the time series, we have to transform the series to stationarity. The simplest way is to take first differences. The t^{th} observation, z_t, of the differenced series is defined as follows:

$$z_t = y_t - y_{t-1} \tag{1.20}$$

where y_t is the level series. If the differenced series still contains a unit root it should be differenced again. If it does not contain a unit root ARMA models can, in principle, be fitted to it: see [4].

1.3.1 Unit Root Test Results

The ACF of the Alsi series reveals significant values for autocorrelation at all 50 lags. The partial autocorrelation at lag 1 is also highly significant - see figures 1.4 and 1.5. The PACF at other lags was insignificant. (The broken lines found on the graphs for the ACF and PACF indicate the significance cutoffs equal to ± 2 times the standard error of the estimate.) The ACFs and PACFs for the Repo and USDZAR series were similar to this. This type of ACF/PACF pattern is associated with the presence of a unit root.

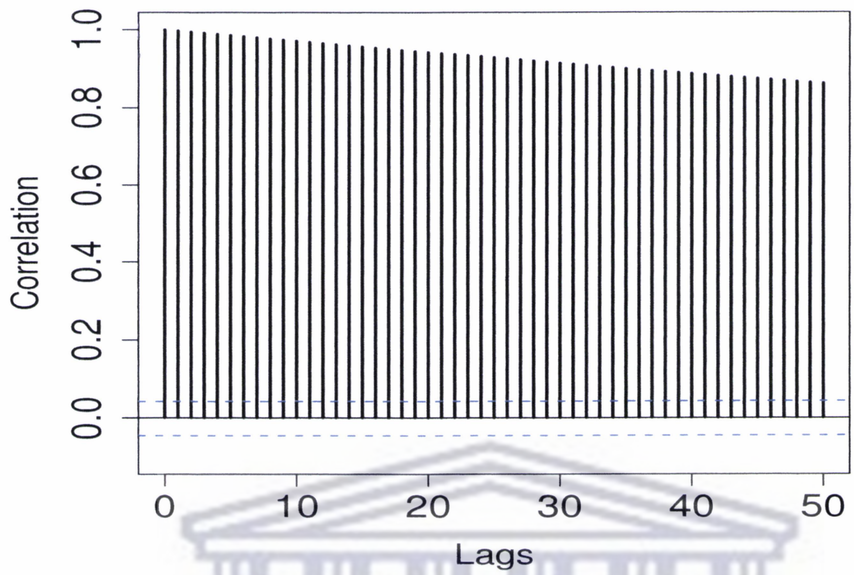

Figure 1.4: ACF of log-transformed Alsi

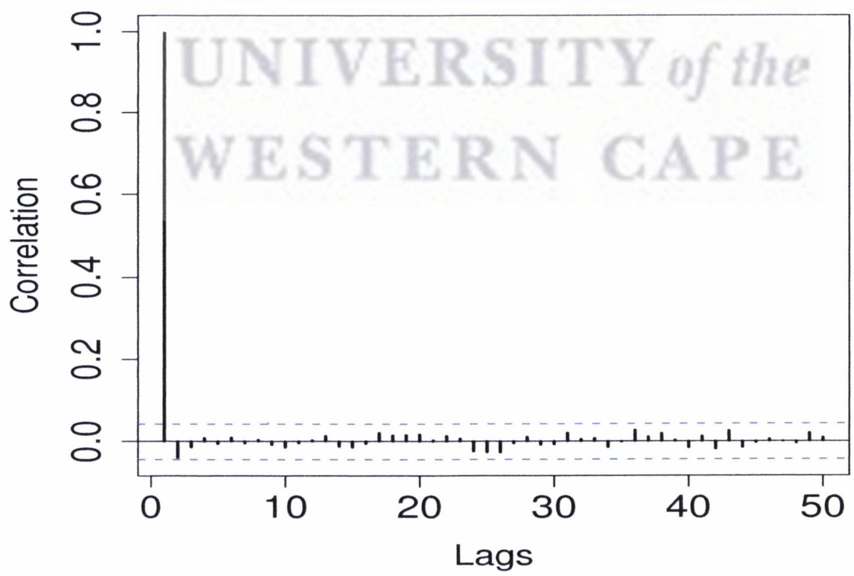

Figure 1.5: PACF of log-transformed Alsi

The results of the Phillips-Perron test are displayed in table 1.1.

Time Series	Phillips-Perron Statistic	p-value
Alsi	-9.8813	0.5586
Repo	-4.3701	0.8655
USDZAR	-3.3356	0.9191

Table 1.1: Phillips-Perron unit root test results

None of these test results are significant, which implies that the null hypothesis of a unit root cannot be rejected.

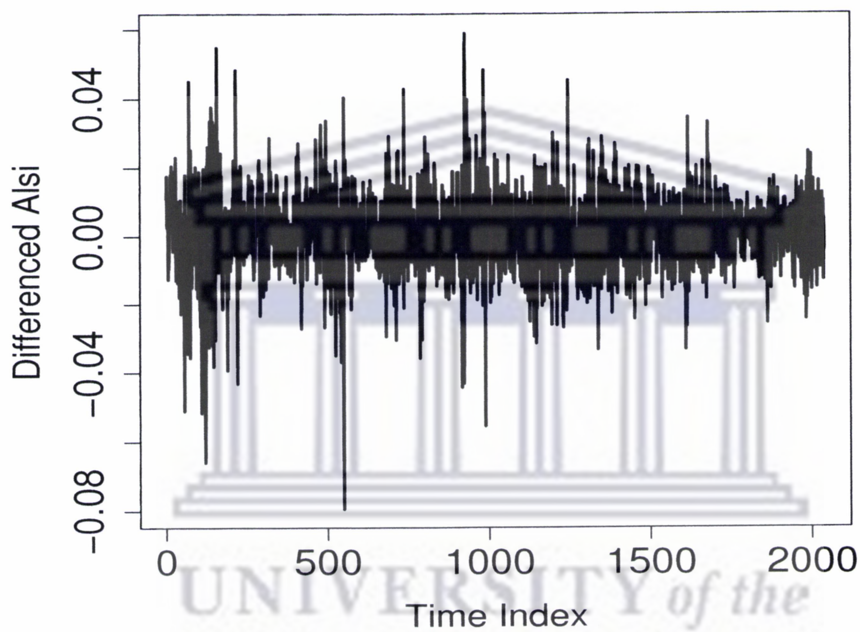

Figure 1.6: Sequence plot of differenced Alsi

The trends present in all three series are removed by taking first differences: see figures 1.6, 1.7 and 1.8. The sequence plots clearly show that the three time series are heteroscedastic. The heteroscedasticity is more pronounced in the Repo and USDZAR series. The unit root test was applied to the differenced log-transformed series and the results show that there are no unit roots present in the series: see table 1.2. The differenced log-transformed series will be used in the remainder of this study, unless otherwise stated.

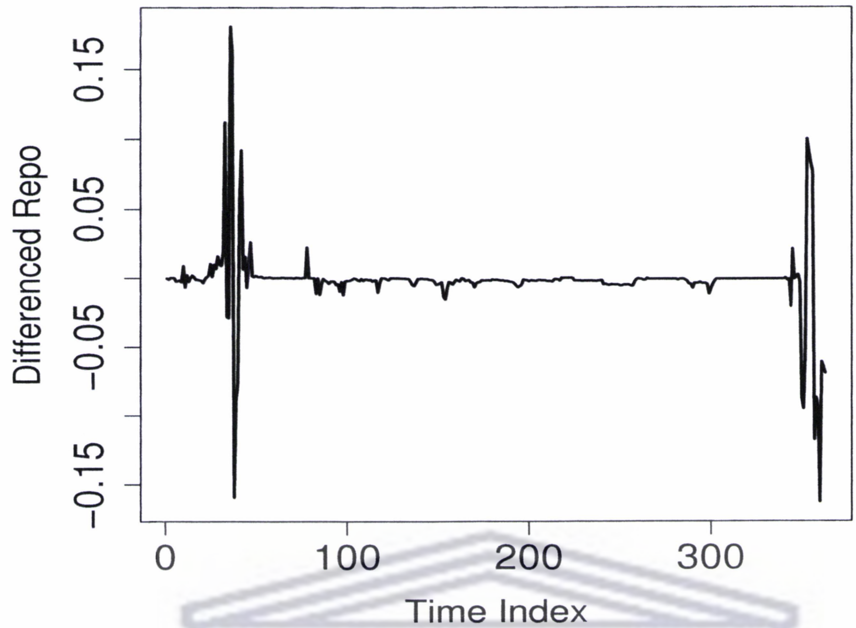

Figure 1.7: Sequence plot of differenced Repo

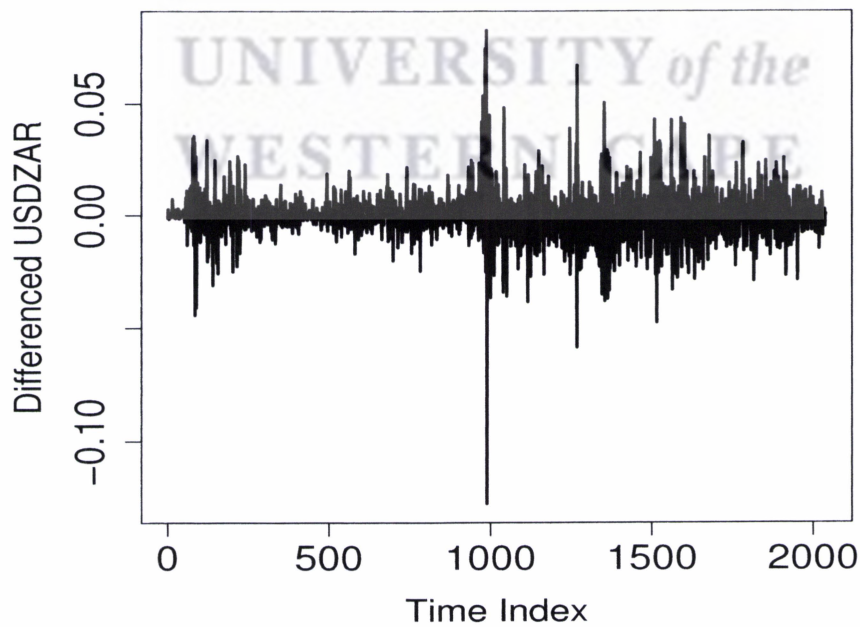

Figure 1.8: Sequence plot of differenced USDZAR

Time Series	Phillips-Perron Statistic	p-value
Alsi	-1809.471	0.01
Repo	-184.8492	0.01
USDZAR	-2104.377	0.01

Table 1.2: Results of the Phillips-Perron unit root test applied to the differenced series

Chapter 2

Univariate Analysis - Standard Approach

2.1 ARIMA Models

2.1.1 ARIMA Methodology

Once the series has been transformed to stationarity the next step is to inspect the ACF and PACF of the time series and to choose a range of possible ARMA models that might fit the data. There are no set rules to fitting ARMA models to time series. Each time series will exhibit its own unique behaviour, which may or may not conform to the behaviour of a theoretical model. Even if the theory suggests that a particular model would be best suited to the time series that we are trying to model, in practice this model is not necessarily optimal. We shall see evidence of this later in the case studies.

Once a model has been fitted there are several diagnostic checks that must be performed to check the adequacy of the model. First, a good model would have residuals that are white noise (stationary and uncorrelated)[6]. If the residuals of the ARMA model are not white noise then the model should be modified as some aspects of the data are not being incorporated into the model. In order to test for correlation in the residuals the ACF, with corresponding Box-Ljung statistics, and the PACF, in conjunction with the standard errors, are used. If more than one adequate model is obtained information criteria are used to select the best model. The AIC and BIC of each model are compared. These information criteria penalise models for each additional parameter estimated. The model with the lowest values for the AIC and BIC should be chosen, all other things being equal. If the AIC and BIC give conflicting results the model with the lowest BIC should be chosen, since it is generally thought that use of the AIC may lead to overparameterisation (see Harvey 1990 [24]).

Models can also be compared using the Box-Ljung statistics. Quantile-Quantile (QQ) plots of the residuals can be analysed in order to determine whether residuals are normally distributed. Finally, the significance of the parameters can be compared in order to see which model yields parameters with the highest significance levels. This should not, however, be a necessary condition for a good model.

2.1.2 ARIMA Results

Alsi

The plot of the ACF for the Alsi series reveals significant autocorrelation at the first three lags. The first two values are significant enough that they cannot be ignored. The high Box-Ljung statistics confirm that there is autocorrelation present, for example at lag 30 a significance level nearly equal to zero, is obtained. The value at the third lag is only just significant as shown in figure 2.1. The PACF returned significant values for the first three lags. In this case the value at the third lag is only just significant - see figure 2.2 below.

The ACF and PACF of the Alsi series suggest that an ARMA(3,3) model, with a constant, may describe the differenced log-transformed Alsi series. This model yielded small Box-Ljung statistics that were not significant, i.e. the residuals seem to be uncorrelated. The partial autocorrelations are also insignificant up till lag 16. None of the model parameters are significant at the 5 percent level though, except the constant. All the models nested in the ARMA(3,3) were therefore also fitted. The models with uncorrelated residuals are displayed in table 2.1. The standard errors of the estimates are given in brackets below each parameter.

The ARMA(2,1), ARMA(1,2) models, both without a constant, showed the best residual ACF and PACF results. Both models returned highly significant estimates for each of the parameters. The ARMA(1,2) model was selected as the most suitable model as it has the

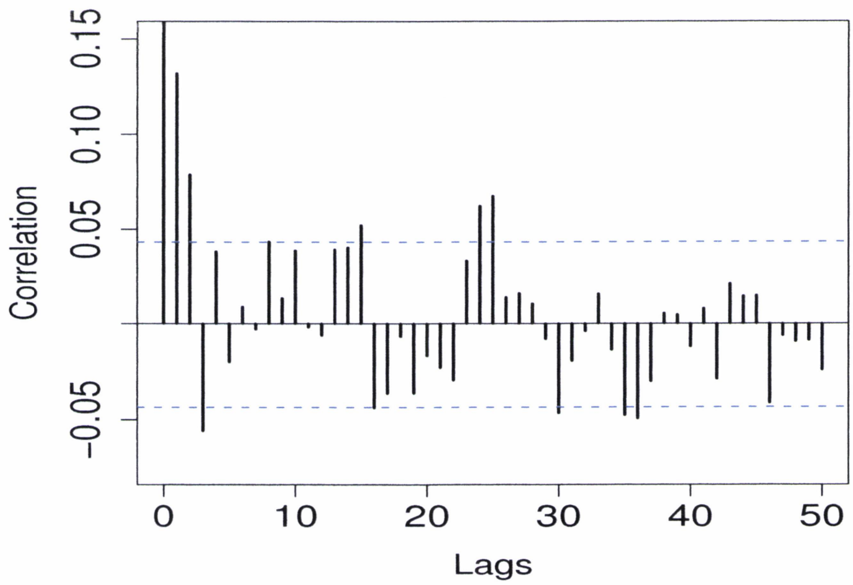

Figure 2.1: ACF of differenced Alsi

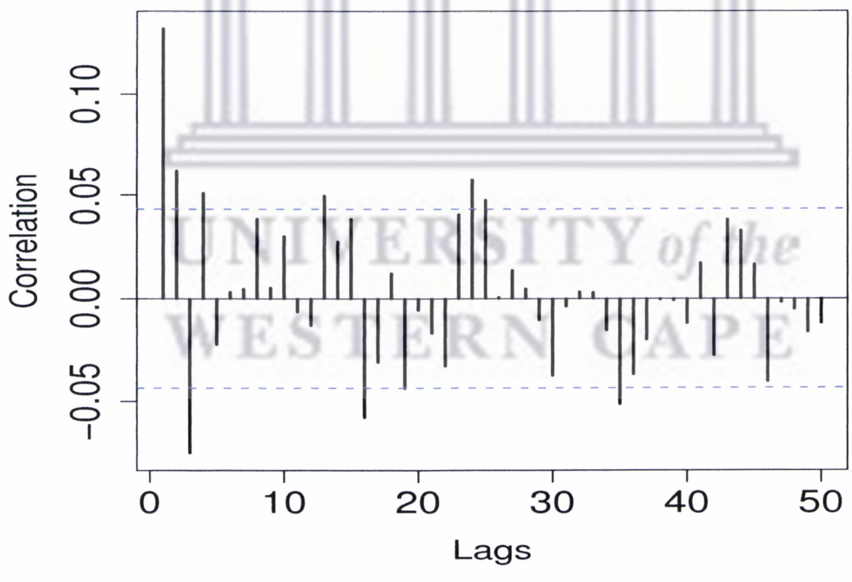

Figure 2.2: PACF of differenced Alsi

lowest BIC. The residual ACF and PACF for this model are given in figure 2.3 and figure 2.4. The ACF returned low Box-Ljung statistics, which were not significant, for the first 24 lags. The PACF showed no significant values in the first 14 lags. A QQ plot shows that the residuals of the ARMA(1,2) model are non-normal: see figure 2.5.

δ	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	σ_ϵ^2	AIC	BIC
0.001	-1.08	-0.33	-0.08	-1.21	-0.55	-0.14	1.37e-004	-12334.205	-12294.863
(0.00)	(1.41)	(0.95)	(0.13)	(1.41)	(1.13)	(0.25)			
0.001	-0.48	0.12	-0.03	-0.62	-0.03		1.37e-004	-12335.923	-12302.202
(0.00)	(0.73)	(0.44)	(0.13)	(0.73)	(0.54)				
0.001	-1.44	-0.50		-1.57	-0.78	-0.13	1.37e-004	-12335.699	-12301.978
(0.00)	(1.23)	(0.91)		(1.23)	(1.06)	(0.21)			
	-0.66	0.02		-0.80	-0.15		1.37e-004	-12336.933	-12314.452
	(0.13)	(0.13)		(0.13)	(0.13)				
	-0.58	0.17		-0.71			1.37e-004	-12337.366	-12320.505
	(0.10)	(0.02)		(0.10)					
	-0.67			-0.80	-0.17		1.37e-004	**-12338.91**	**-12322.04**
	(0.11)			(0.11)	(0.02)				

Table 2.1: Comparison of Alsi ARMA models

Repo

The ACF of the differenced Repo series reveals that the values at lags 1 and 4 are significant. The Box-Ljung statistics were significant for all 50 lags. The PACF is significant at the first two lags. This is shown in figures 2.6 and 2.7 respectively. An ARMA(2,1) and all nested models were fitted. The results are shown in table 2.2.

Only the ARMA(2,0) and the ARMA(0,1) models, both fitted without a constant, gave significant estimates for every parameter. The ARMA(0,1) model had fewer significant values in the residual ACF and PACF and only returned one marginally significantly value for both the ACF and the PACF - see figures 2.8 and 2.9.
The residuals were non-normal according to the QQ plot: see figures 2.10 and 2.11.
Every model fitted to the Repo series yielded non-normal residuals. This could be due to

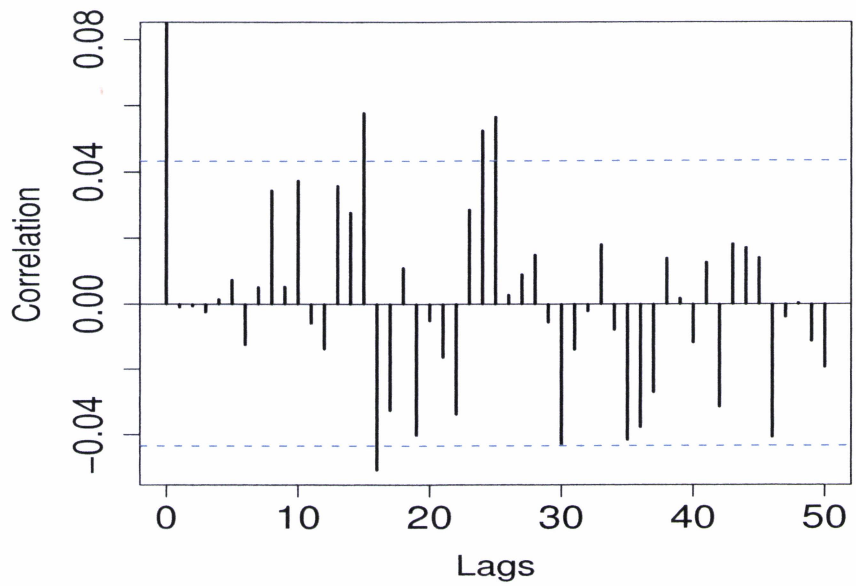

Figure 2.3: ACF of Alsi ARMA(1,2) residuals

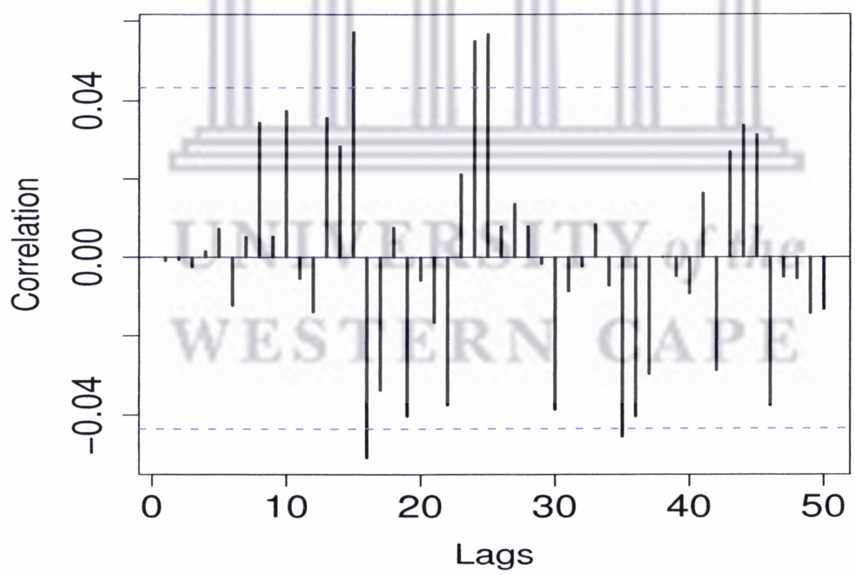

Figure 2.4: PACF of Alsi ARMA(1,2) residuals

the manner in which the Repo rate is determined. As mentioned previously, the Repo rate is determined by the South African Reserve Bank (SARB). Before the 24_{th} of November 1999 the Repo series was continuous. On the 24_{th} of November 1999 the SARB's Monetary Policy Committee decided to fix the Repo rate at 12 percent (Nattrass *et al* 2002 [11]).

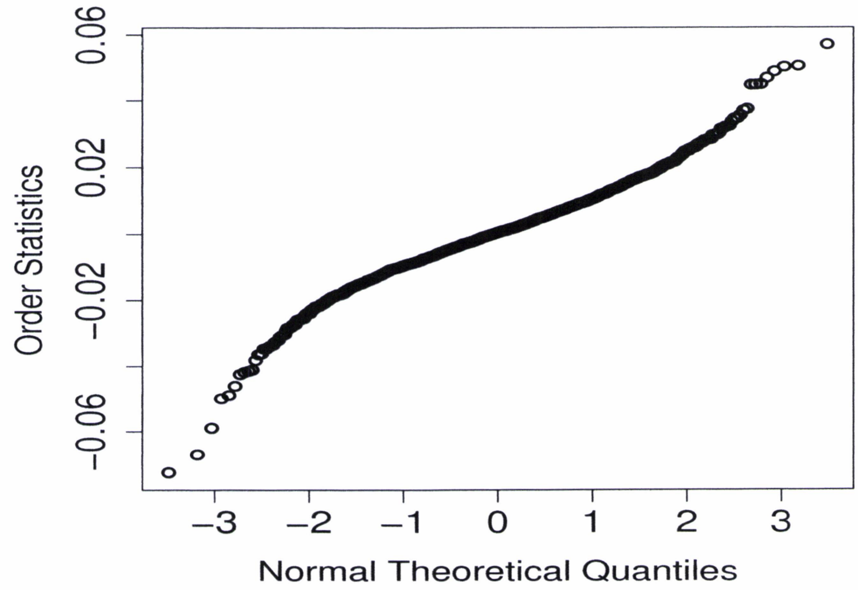

Figure 2.5: QQ plot of Alsi ARMA(1,2) residuals

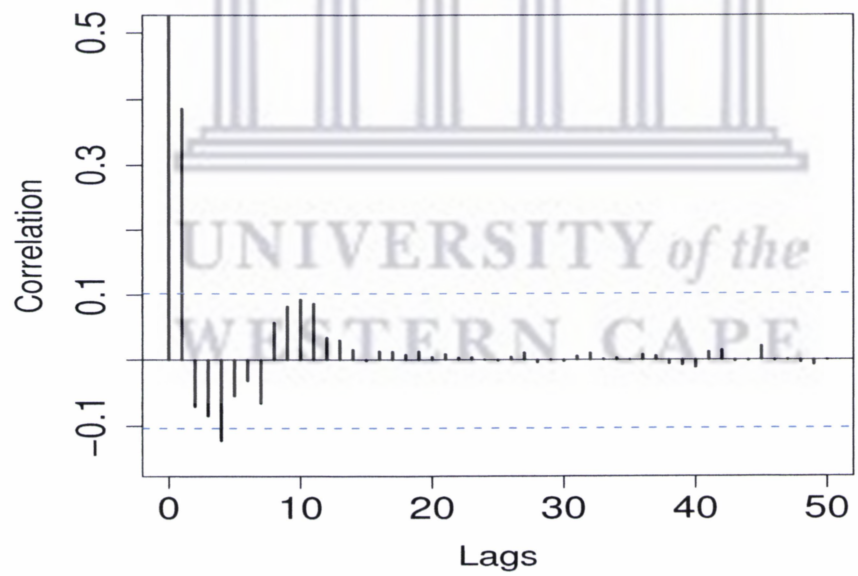

Figure 2.6: ACF of differenced Repo

Thereafter the SARB has modified the Repo rate in fixed increments, usually half a percentage point. The Repo series only takes on distinct discrete values, hence the differenced log series is also discrete. Figures 2.10 and 2.11 show that the normal distribution is a poor approximation for this discrete distribution.

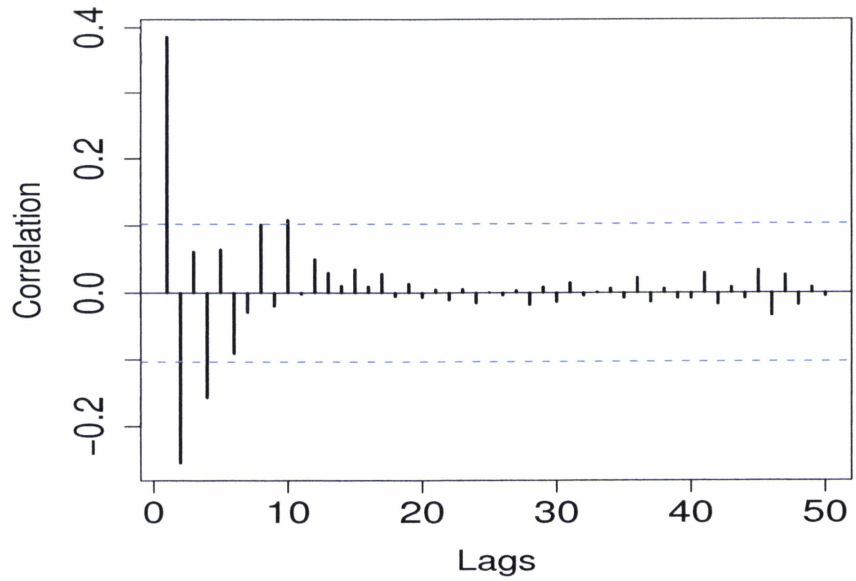

Figure 2.7: PACF of differenced Repo

δ	ϕ_1	ϕ_2	θ_1	σ^2_ϵ	AIC	BIC
-0.002	-0.028	-0.046	-0.560	0.01	-1697.233	-1681.655
(0.002)	(0.138)	(0.087)	(0.129)			
	0.493	0.493		0.01	-1692.564	-1684.77
	(0.051)	(0.052)				
-0.002	-0.072		-0.606	0.01	-1698.903	-1687.220
(0.002)	(0.095)		(0.076)			
			-0.561	0.01	**-1700.95**	**-1697.06**
			(0.044)			

Table 2.2: Comparison of Repo ARMA models

USDZAR

The ACF of the USDZAR series is significant at 4 different lags that are interspersed among the first 12 lags - see figure 2.12. The Box-Ljung statistics are large and significant

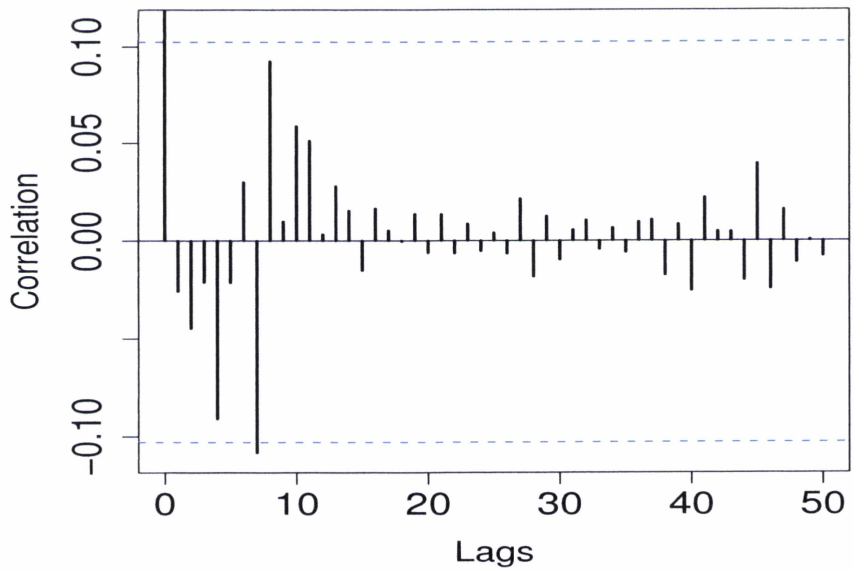

Figure 2.8: ACF of Repo ARMA(0,1) residuals

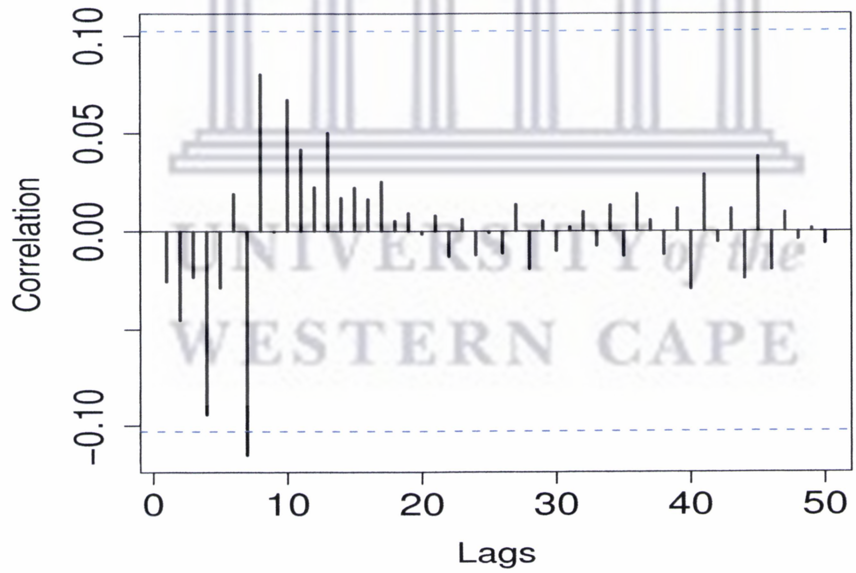

Figure 2.9: PACF of Repo ARMA(0,1) residuals

for lags 1-50.

The PACF returned five significant coefficient values in the first 12 lags. An ARMA(3,3), as well as all nested models, were fitted. It was found that the ARMA(3,2) and ARMA(2,3) models, both without a constant, managed to model the autocorrelation in the first 10

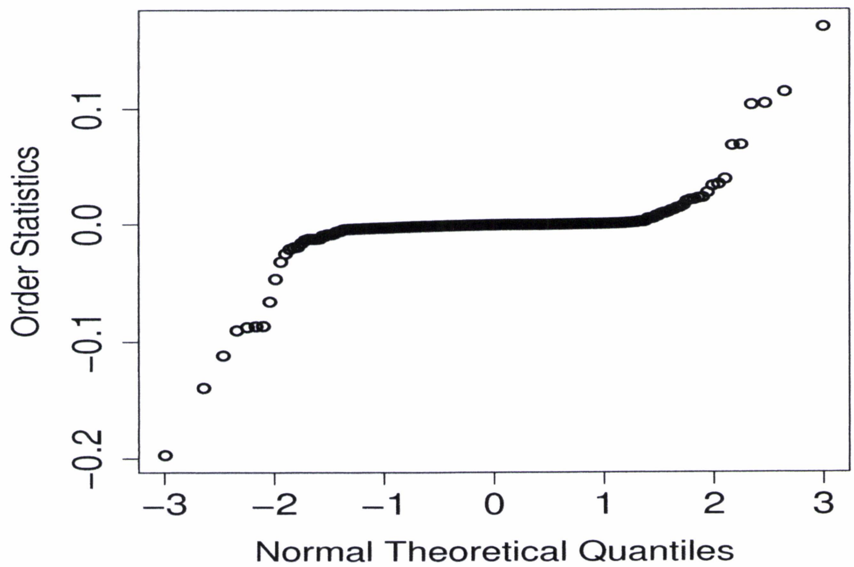

Figure 2.10: QQ plot of Repo ARMA(0,1) residuals

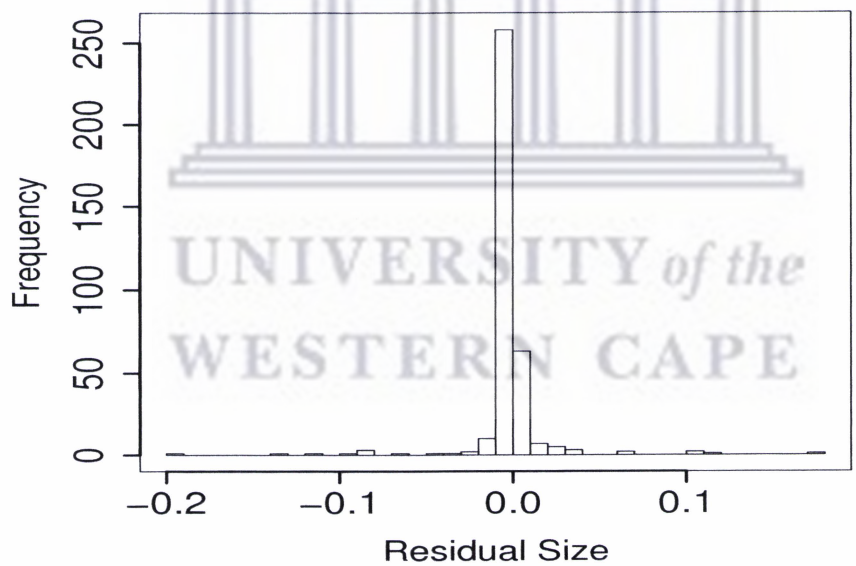

Figure 2.11: Histogram of Repo ARMA(0,1) residuals

lags: see figures 2.14 and 2.15. The details of these two models are shown in table 2.3 below.

The residual ACFs and PACFs of the two models are virtually identical - see figures 2.14 to 2.17. Neither model produced residuals that are normally distributed, as shown in

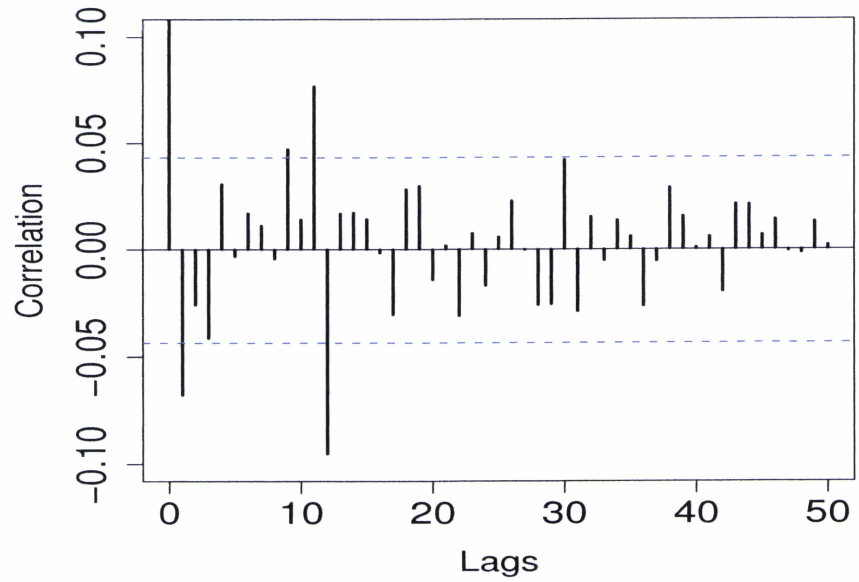

Figure 2.12: ACF of differenced USDZAR

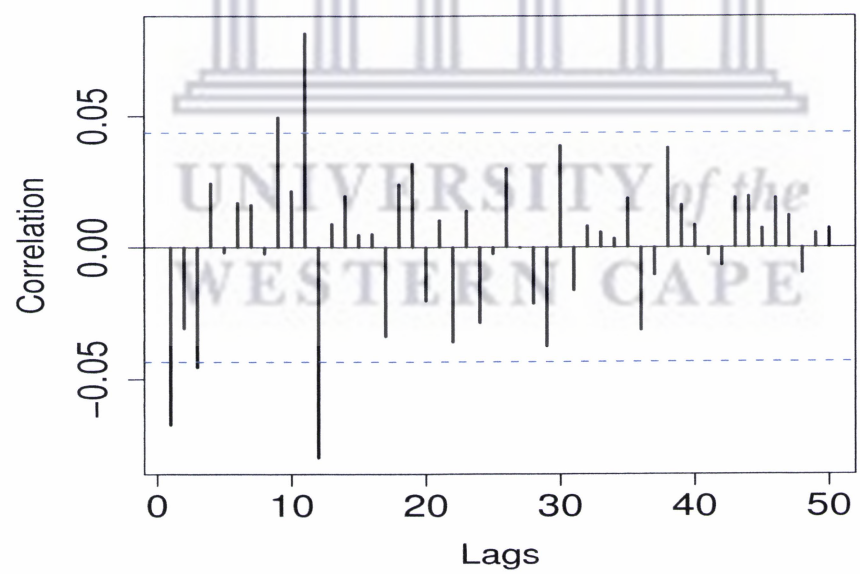

Figure 2.13: PACF of differenced USDZAR

figures 2.18 and 2.19. The ARMA(2,3) model was selected as it has the lowest value for the BIC.

ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	σ_ϵ^2	AIC	BIC
-0.528	-0.163	-0.066	-0.458	-0.098		1.2150e-004	-12586.95	-12558.85
(0.369)	(0.375)	(0.031)	(0.370)	(0.368)				
-0.661	-0.153		-0.592	-0.079	0.068	1.2150e-004	**-12587.10**	**-12559.00**
(0.372)	(0.372)		(0.371)	(0.358)	(0.032)			

Table 2.3: Comparison of USDZAR ARMA Models

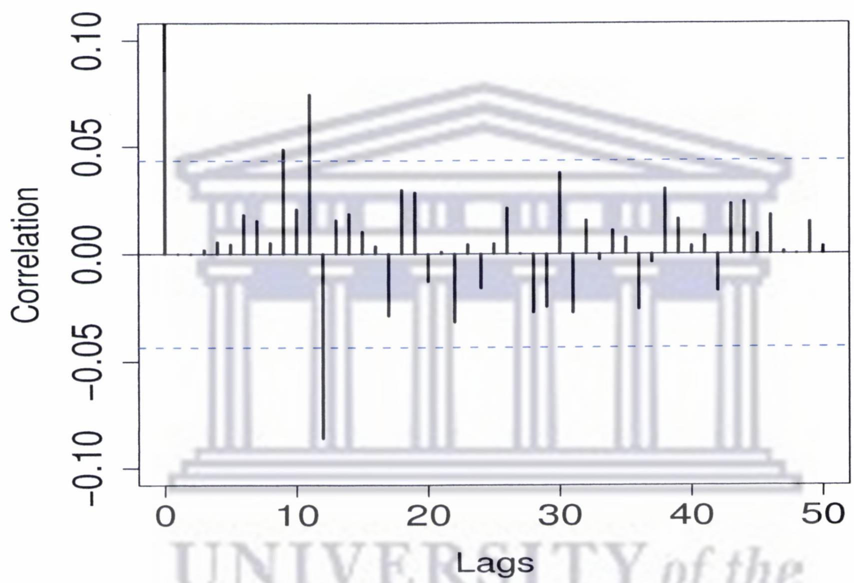

Figure 2.14: ACF of USDZAR ARMA(3,2) residuals

2.1.3 Discussion

It is interesting to note that the Repo is a short-memory series as it is best modelled with a single moving average term. The Alsi and USDZAR series are both long-memory series, both models requiring autoregressive terms. All three models successfully described the autocorrelation in the respective series.

The dramatic change in the USDZAR trend around index number 1000 (Figure 1.2) is bound to influence the details of the fitted models. An alternative approach to modelling this series is discussed later.

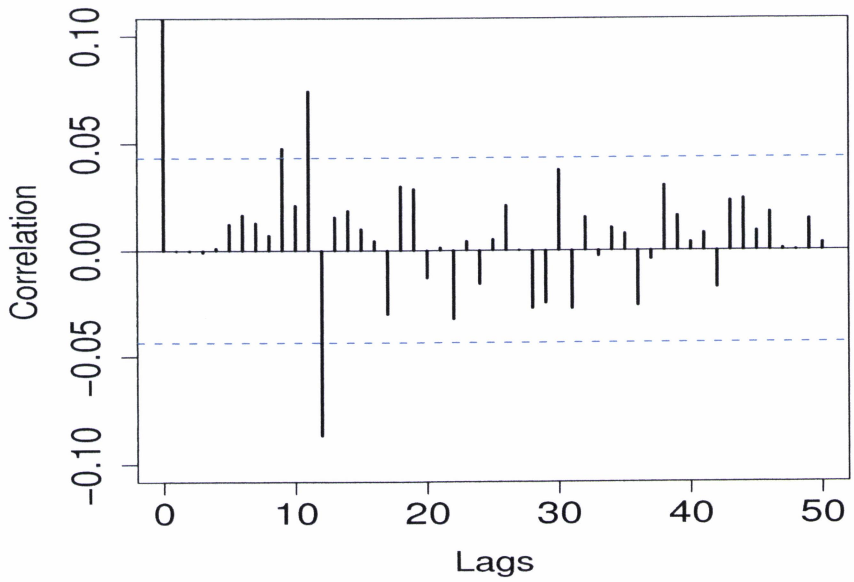

Figure 2.15: ACF of USDZAR ARMA(2,3) residuals

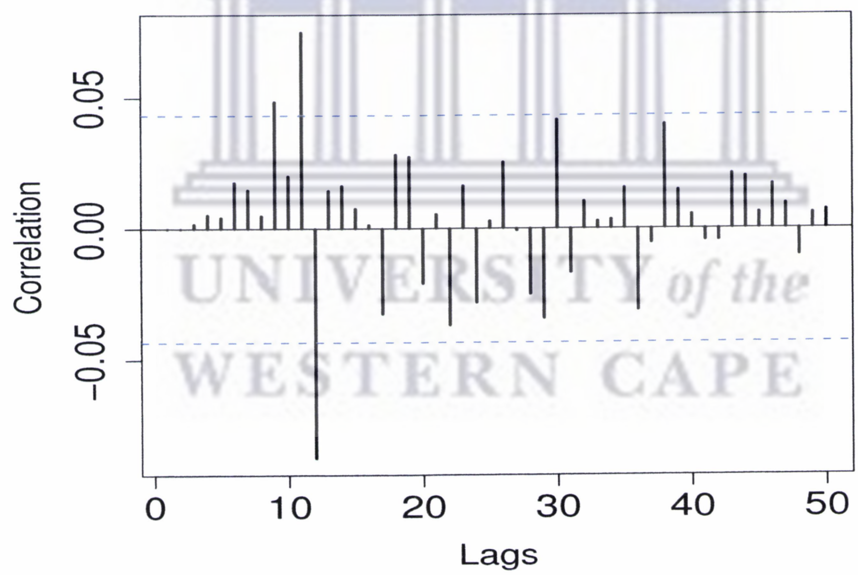

Figure 2.16: PACF of USDZAR ARMA(3,2) residuals

2.2 G(ARCH) Models

2.2.1 G(ARCH) Methodology

Financial time series often show variable variance. The sequence plots of the differenced, log-transformed series are particularly illustrative of this point. In foreign currency mar-

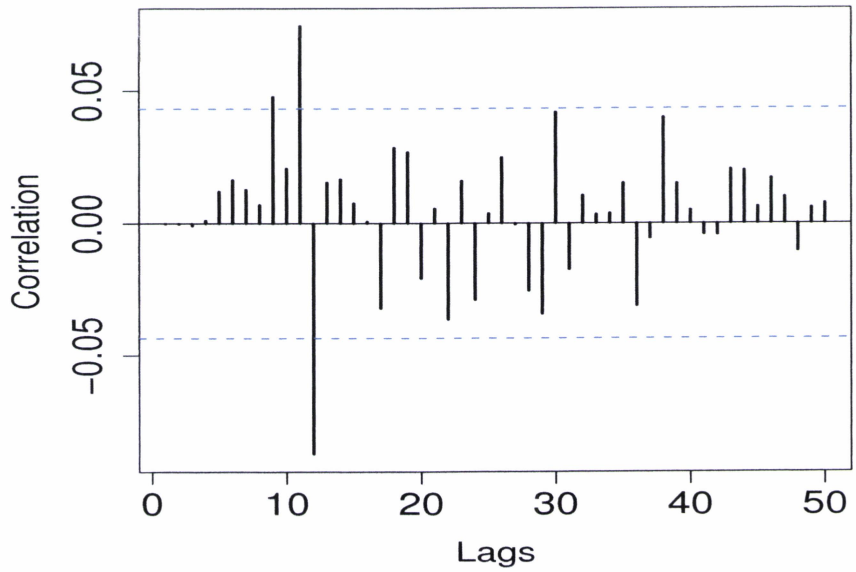

Figure 2.17: PACF of USDZAR ARMA(2,3) residuals

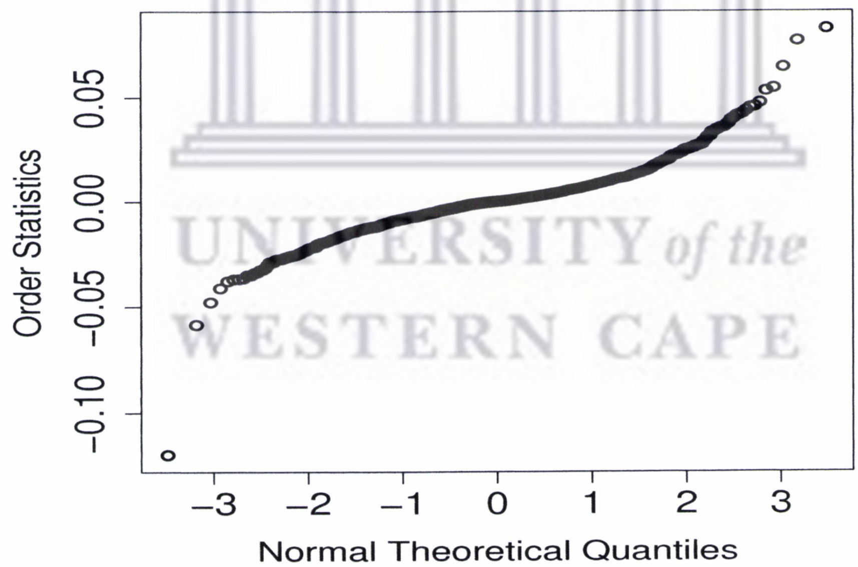

Figure 2.18: QQ plot of USDZAR ARMA(3,2) residuals

kets it is not uncommon to witness sustained periods of extreme volatility, particularly when there is an exogenous variable, such as supply concerns, influencing the market. An example can be seen in the oil markets. The price per barrel of Brent Crude oil has fluctuated from below forty dollars to above seventy dollars in the last two years alone,

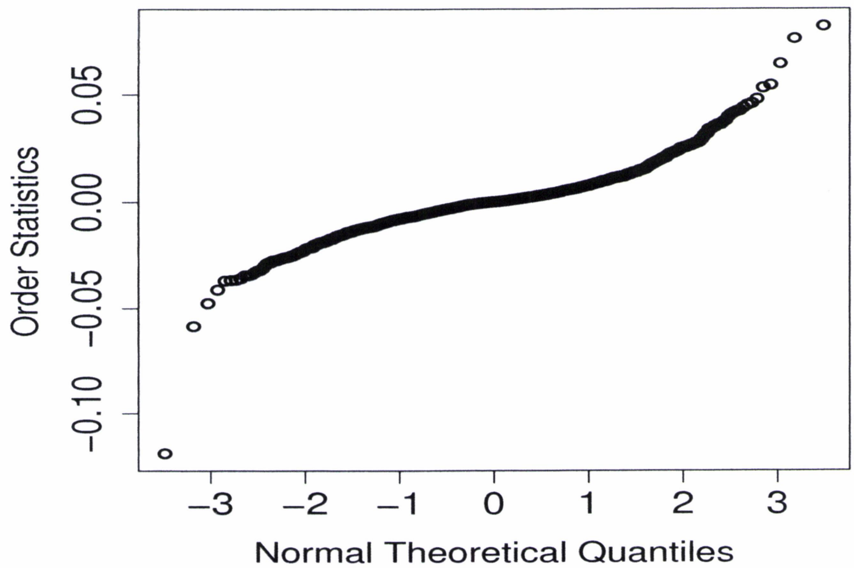

Figure 2.19: QQ plot of USDZAR ARMA(2,3) residuals

due to supply disruptions such as hurricanes in the Persian Gulf, war in Iraq and attacks on Nigerian pipelines. It seems that the assumption that the volatility of the oil price during times of supply unrest is equal to that during times of relatively secure supply, cannot hold. The same can be said of the volatility of the three time series discussed in this study.

The Pearson correlation statistic was used to test for autocorrelation in the model residuals. These values can be misleading if the data are non-Normal or have non-linear dependence. The correlation estimates are also influenced by outliers (Lange and Grubmüller 2006 [13]). Pearson correlation is a measure of linear association. Several methods of autocorrelation estimation exist that are not restricted to linear association. Examples are Kendall's Tau and the Spearman Rank correlation estimates, both of which are non-parametric measures of association. Both coefficients require ranking of the data. Kendall's Tau and the Spearman Rank correlation estimates were generated for the residuals of the final ARMA models, for the first 50 lags. These functions may shed some light on possible autocorrelation that is due to autoregressive conditional heteroscedasticity (ARCH), since this gives rise to nonlinear association between innovations.

If heteroscedastic variance is present, the next question is how to model it. Generalised

ARCH (G(ARCH)) modelling is used. G(ARCH) models are fitted in order to model the conditional expectation of the squared residuals of the ARMA models. In this case G(ARCH) models should be fitted to the residuals of the best fitting ARMA model for each time series. The G(ARCH) model should capture all correlation in the squared residuals. If more than one G(ARCH) model is adequate, then the one with the lowest BIC and AIC should be selected. The final model for the financial time series consists of two parts: the first describing the transformed level series, and the second the variance structure in the residuals.

2.2.2 G(ARCH) Results

The autocorrelation in the residuals of best fitting ARMA models was tested using Kendall's τ and Spearman's ρ. The results for Kendall's τ are displayed in figures 2.20 to 2.22. The results for Spearman's ρ are very similar. Figures 2.20 to 2.22 may be compared to the corresponding standard Pearson's ACFs in figures 2.3, 2.8 and 2.15. It is noteworthy that the Repo and USDZAR residuals do indeed show strong evidence of autocorrelation, whereas none is visible in the Pearson correlations.

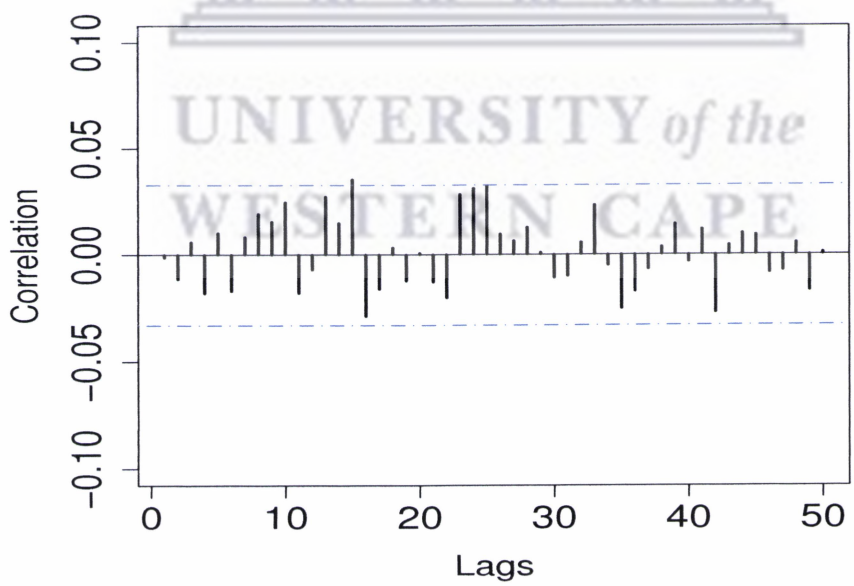

Figure 2.20: ACF of Alsi ARMA(1,2) residuals using Kendall's Tau

For each of the models the conventional (Pearson) ACF of the squared residuals was also generated. These are shown in figures 2.23 to 2.25. There is significant autocorrelation

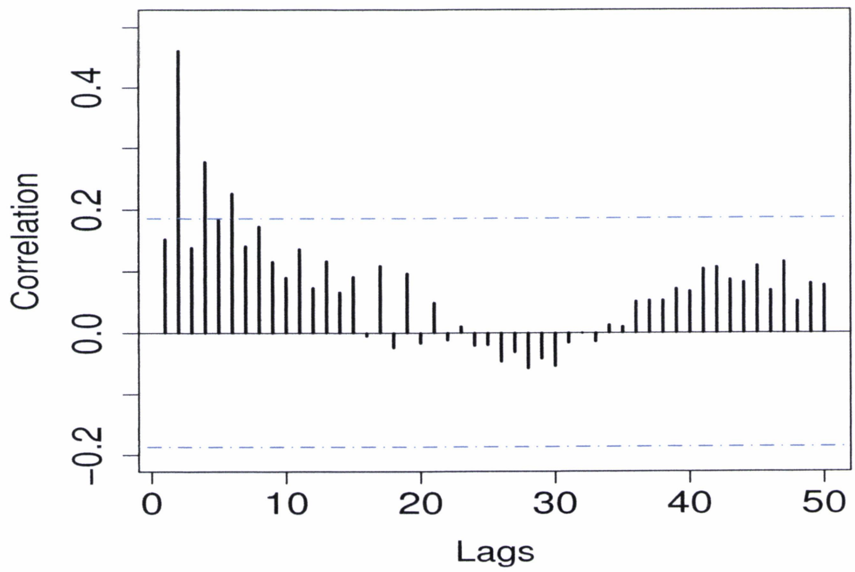

Figure 2.21: ACF of Repo ARMA(0,1) residuals using Kendall's Tau

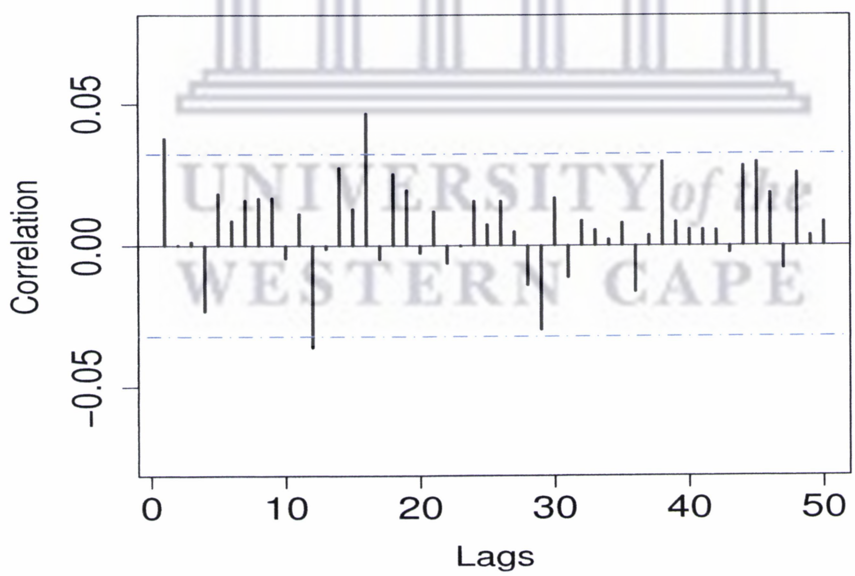

Figure 2.22: ACF of USDZAR ARMA(2,3) residuals using Kendall's Tau

in the squared residuals of all three series. G(ARCH) models were fitted to the residuals in an attempt to describe the ARCH.

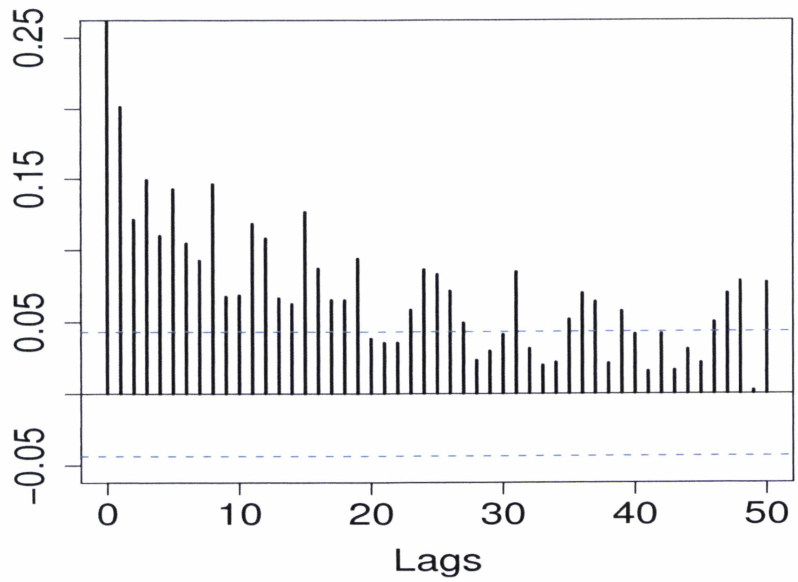

Figure 2.23: ARCH in Alsi ARMA(1,2) squared residuals

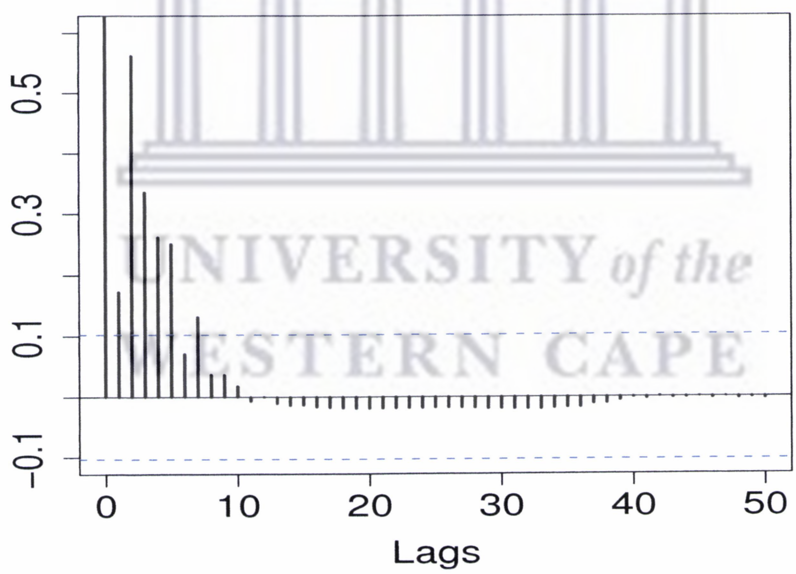

Figure 2.24: ARCH in Repo ARMA(0,1) squared residuals

Alsi

A G(ARCH)(1,1) model yielded the best results for the Alsi series. All ARCH effects have been captured - see figures 2.26 and 2.27. All parameter estimates are significant at below the 1 percent level. The Jarque Bera test yielded a high statistic of 179.48 that is

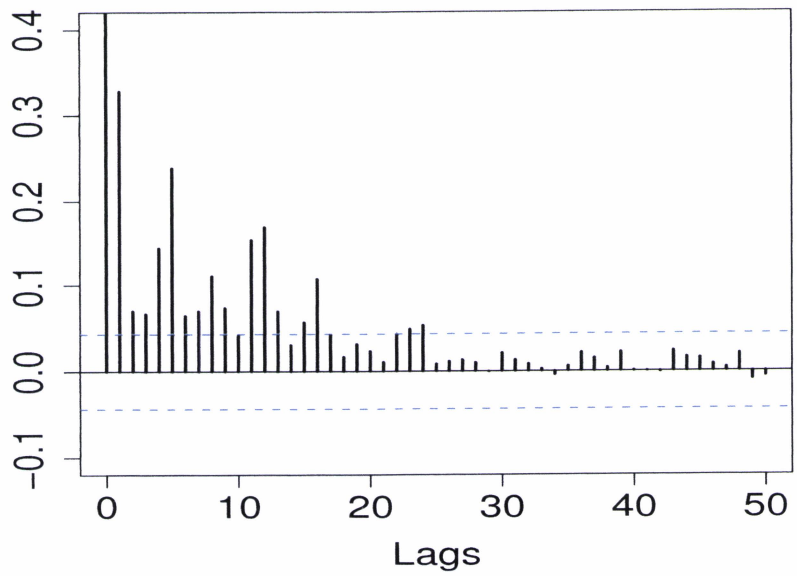

Figure 2.25: ARCH in USDZAR ARMA(2,3) squared residuals

highly significant. This implies that the G(ARCH) residuals are non-normal. The model summary is presented in table 2.4.

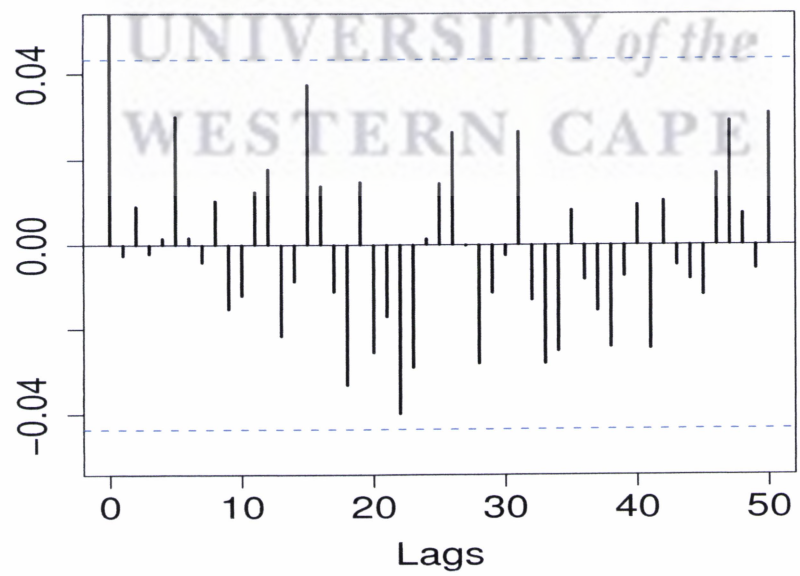

Figure 2.26: ARCH in Alsi squared GARCH residuals

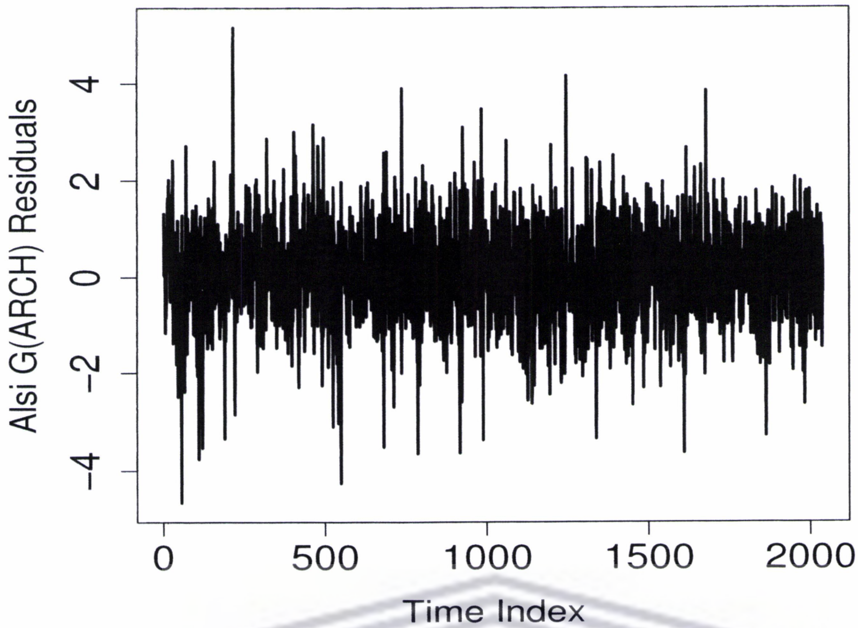

Figure 2.27: Sequence plot of G(ARCH) residuals for Alsi

Parameter	Estimate	Standard Error
α_0	2.295e-06	6.326e-07
α_1	7.369e-02	7.264e-03
β_1	9.111e-01	8.537e-03

Table 2.4: Alsi: G(ARCH)(1,1) model coefficients

USDZAR

The best model was found to be a G(ARCH)(1,1) once again. All the parameter estimates were highly significant. All ARCH effects have been accounted for - see figures 2.28 and 2.29. The model summary is displayed in table 2.5.

Parameter	Estimate	Standard Error
α_0	7.491e-07	1.535e-07
α_1	2.107e-01	1.426e-02
β_1	8.309e-01	7.720e-03

Table 2.5: USDZAR: G(ARCH)(1,1) model coefficients

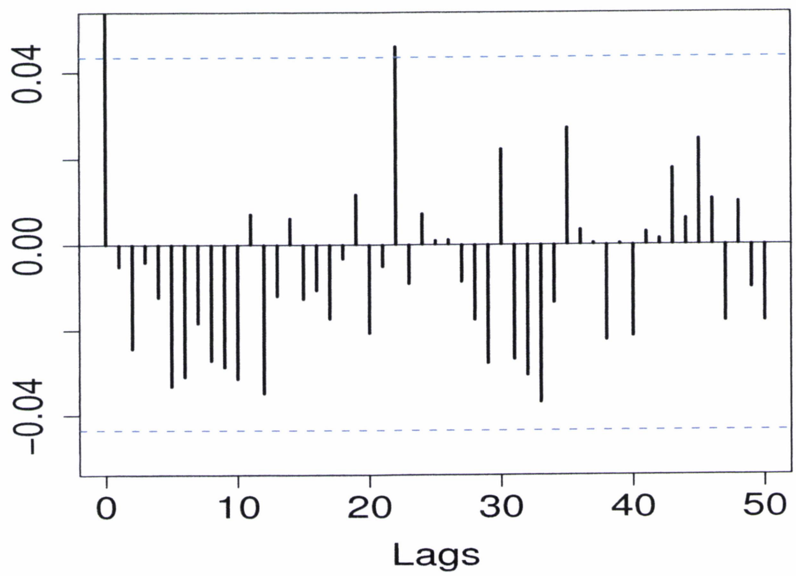

Figure 2.28: USDZAR ARIMA residuals: ARCH in squared G(ARCH) residuals

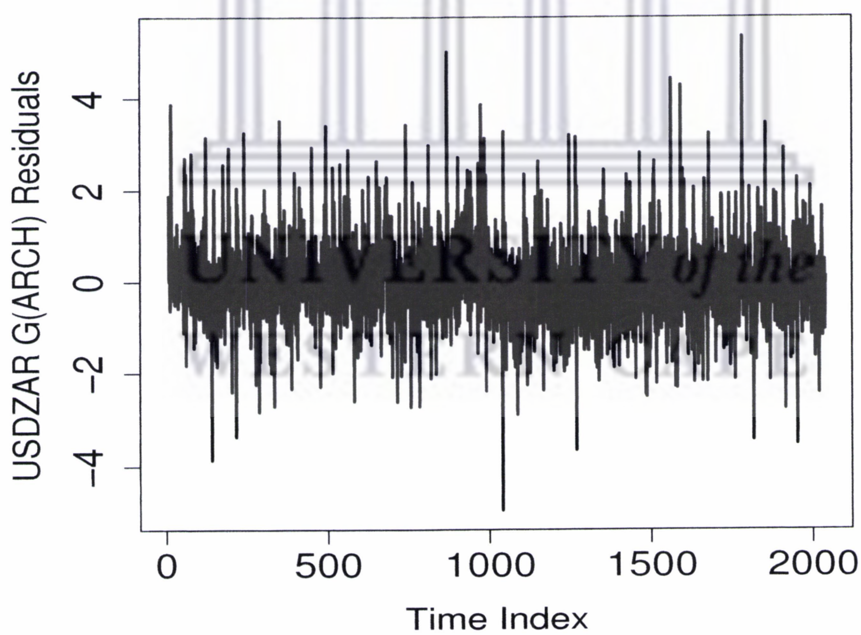

Figure 2.29: Sequence plot of G(ARCH) residuals for USDZAR

Repo

The only model that described most of the ARCH effects was the G(ARCH)(2,2) - see figure 2.31. The ACF of the squared residuals contains a large spike at lag 45. Because the lag length is so large and the significant value occurs at a lag without special significance, this spike can be ignored. It is probably specific to this sample rather than a behavioural characteristic that should be modelled. Not all of the parameter estimates are significant - see table 2.6. The estimate for α_2 is very small and non-significant yet leads to ARCH in lower lags if omitted. It appears from the sequence plot in figure 2.31 that the model is inadequate. Since all other models fitted failed to capture the ARCH in the lower lags, it seems G(ARCH) modelling is not a suitable means of describing the behaviour of the Repo ARMA residuals.

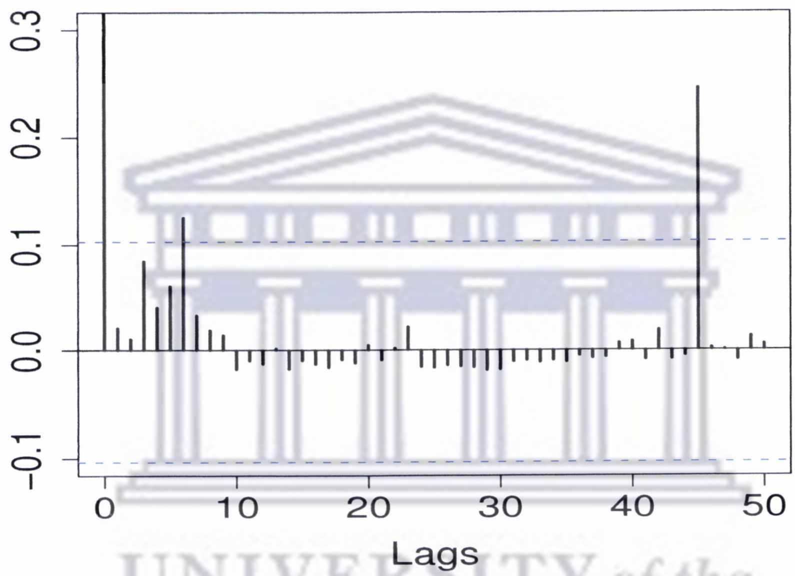

Figure 2.30: ARCH in Repo squared GARCH residuals

Table 2.6 displays the G(ARCH)(2,2) model summary for the Repo series.

Parameter	Estimate	Standard Error	p-values
α_0	3.565e-06	4.618e-07	1.15e-14
α_1	2.410e-01	1.699e-02	$<$ 2e-16
α_2	6.172e-09	2.305e-02	1.0000
β_1	2.138e-01	8.798e-02	0.0151
β_2	3.744e-01	4.263e-02	$<$ 2e-16

Table 2.6: Repo: G(ARCH)(2,2) model coefficients

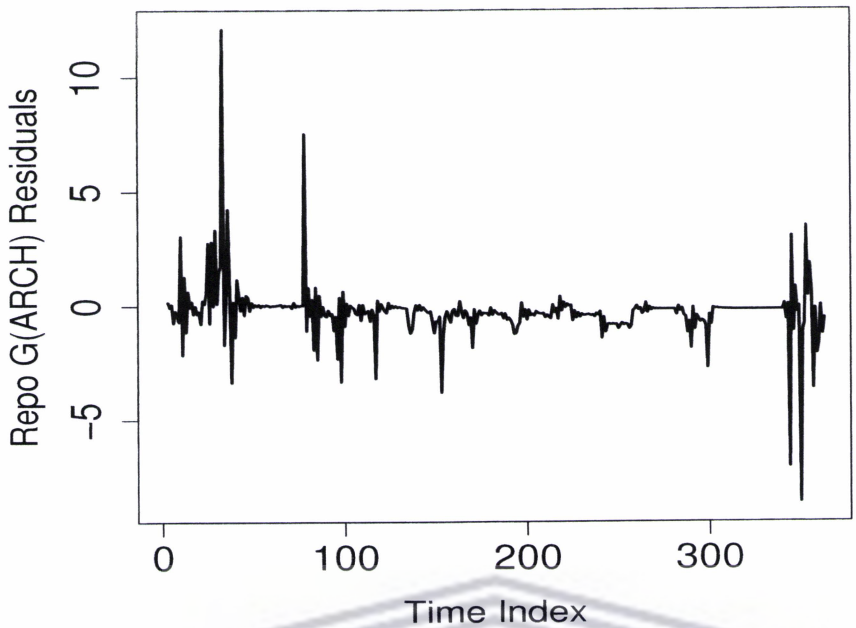

Figure 2.31: Sequence plot of G(ARCH) residuals for Repo

2.3 Outliers

2.3.1 Method of Testing for Outliers

Once the most suitable model has been identified it may be necessary to test for any outliers that may be affecting the model parameters and model accuracy. The method used for identifying time series outliers is taken from Wei 1989 [3] and Box *et al* [6] The technique is a four step iterative procedure that detects both additive outliers (AO) and innovation outliers (IO).

AOs affect the level series only at the time they occur: that is, the distortion is restricted to a single observation. IOs affect the level series at the observation when the distortion occurs and at every observation thereafter through the memory process of the series. If the model is written as:

$$\phi(B)y_t = \theta(B)\epsilon_t \tag{2.1}$$

then an IO affects the time series through $\theta(\mathrm{B})/\phi(\mathrm{B})$.

The method requires that an ARMA model be fitted to the data on the assumption that there are no outliers present in the data. The sequences

$$\Pi(B) = \phi(B)/\theta(B) = (1 - \sum_{i=1}^{n} \pi_i B^i) \tag{2.2}$$

and

$$\Psi(B) = \theta(B)/\phi(B) = (1 + \sum_{i=1}^{n} \psi_i B^i) \tag{2.3}$$

where n is the length of the time series, are calculated. These operators are applied to the estimated residuals to generate test statistics, for each observation of the time series, in the following manner:

$$\omega_{A,\tau} = (1 - \Pi(F))\widehat{\epsilon}_\tau/\gamma^2 \tag{2.4}$$

for AOs, where F is a forward shift operator, $\gamma^2 = \sum_{j=0}^{n-\tau} \pi_j^2$ and τ is the time of the outlier occurrence, and

$$\omega_{I,\tau} = \widehat{\epsilon}_\tau \tag{2.5}$$

for IOs. These statistics are scaled to create likelihood ratio test statistics

$$\lambda_{A,\tau} = \tau\omega_{A,\tau}/\sigma_\epsilon \tag{2.6}$$

to test the alternative hypothesis that y_τ is an AO and

$$\lambda_{I,\tau} = \omega_{I,\tau}/\sigma_\epsilon \tag{2.7}$$

to test the alternative hypothesis that y_τ is an IO against the null hypothesis that y_τ is neither an AO nor an IO (where y_τ is the series being modelled). These statistics are calculated for each time period. The test statistic that is greatest in absolute value is used to determine if there is an AO or an IO at that specific observation.

If y_τ is found to be an outlier then the series y_t is modified in the following manner

$$y_t^* = y_t - \omega_{A,\tau} I_t^{(\tau)} \tag{2.8}$$

if the outlier is an AO, and

$$y_t^* = y_t - \Psi(B)\omega_{I,\tau} I_t^{(\tau)} \tag{2.9}$$

if it is an IO, where

$$I_t^{(\tau)} = \begin{cases} 0 & t < \tau \\ 1 & t \geq \tau \end{cases} \tag{2.10}$$

is an indicator variable.

The residual series is modified in the following manner:

$$\widehat{\epsilon}_t^* = \widehat{\epsilon}_t - \omega_{A,\tau}\Pi(B) I_t^{(\tau)} \tag{2.11}$$

for an AO, and

$$\widehat{\epsilon}_t^* = \widehat{\epsilon}_t - \omega_{I,\tau} I_t^{(\tau)} \tag{2.12}$$

for an IO.

This completes one iteration. The likelihood ratio test statistics are re-estimated given the modified series with $\Pi(B)$ and $\Psi(B)$ kept intact. The whole process is repeated until no additional outliers are found. Once all the outliers have been detected then these observations can be removed and the model re-estimated. The program written to implement the method is given in the Appendix 2.

2.3.2 Outlier Results

Alsi

The ARMA(1,2) model residuals were used to test for outliers. The model returned just one IO at observation 550. This observation of the differenced Alsi series was deleted and the parameters re-estimated. The parameter estimates and their standard errors were virtually unaffected by the removal of the outlier. However, since further investigation of the properties of the time series may be affected by the outlier it is best removed.

Table 2.7 shows that the parameter estimates have slightly larger standard errors. Every ARMA and GARCH model was refitted to the time series with the outlier removed. The

Parameter	Estimate	Standard Error
ϕ_1	-0.657	0.112
θ_1	-0.797	0.111
θ_2	-0.174	0.022

Table 2.7: ARMA(1,2) model coefficients: Alsi with outlier removed.

ARMA(1,2) model was still found to be best. The G(ARCH)(1,1) model parameters changed very little and remained highly significant.

USDZAR

The ARMA(2,3) model was used to test for outliers. The program returned just one IO, namely the 989^{th} observation. Once this observation had been removed and the model re-estimated, nearly all the parameter estimates were significant at the 5 percent level. Indeed all of the parameters were significant at the 6 percent level. More importantly, the standard errors of the parameter estimates were dramatically reduced. The new model summary is given in table 2.8.

Parameter	Estimate	Standard Error	p-values
ϕ_1	0.328	0.170	0.075
ϕ_2	-0.770	0.156	0.681
θ_1	0.357	0.171	0.111
θ_2	-0.761	0.162	0.826
θ_3	0.053	0.025	0.032

Table 2.8: ARMA(2,3) model coefficients: USDZAR with outlier removed.

The ARMA(3,2) and ARMA(2,3) models were found to be the best to describe the time series with the outlier removed. Interestingly, the ARMA(2,3) model captured more of the correlation structure than the ARMA(3,2) model; however,the ARMA(3,2) model provided the lowest BIC. Once again the G(ARCH)(1,1) model parameters changed very little and remained highly significant.

Repo

There is an AO at observation 38. Once this had been removed all the models in section 2.1.2 were refitted. An ARMA(2,3) model, with no constant, was found to be the best. The parameter estimates are all highly significant and are displayed in table 2.9. However, none of the models performed significantly better with respect to explaining autocorrelation in the series.

Parameter	Estimate	Standard Error
ϕ_1	-0.608	0.150
ϕ_2	-0.692	0.158
θ_1	-1.357	0.134
θ_2	-1.271	0.177
θ_3	-0.631	0.092

Table 2.9: ARMA(2,3) model coefficients: Repo with outlier removed.

The G(ARCH) models were refitted and it was found that a G(ARCH)(1,1) model described all of the ARCH effects in the residuals.

2.4 Quasi-periodicity

According to Fuller 1996 [26] a function is deemed periodic if $\exists$H > 0 such that $\forall$ t, t+H $\in$ T

$$f(t+H) = f(t) \tag{2.13}$$

where H is the period of the function. If one thinks of a time series as a function with time as input, then it is clear that a periodic time series repeats every H time units. A strictly periodic sinusoidal time series is defined as:

$$y_t = A\sin(\omega t + \psi) = A\sin(\frac{2\pi}{P}t + \psi) \tag{2.14}$$

where A is the amplitude, ω is the frequency, ψ is the phase and P is the period of the time series.

A quasi-periodic time series is here defined as

$$y_t = A_t\sin(\omega_t t + \psi_t) = A_t\sin(\frac{2\pi}{P_t}t + \psi_t) \tag{2.15}$$

where A_t, ω_t, ψ_t and P_t are slowly time varying, i.e. they change very little during a cycle, P_t, of variation - see [7]. It is well known that certain autoregressive time series also exhibit quasi-periodicities (e.g.[6]). Checking for quasi-periodicity requires that the AR polynomial, with the backshift operator acting as a variable as in equation 2.16, be set equal to zero and the roots calculated.

$$1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p = 0 \tag{2.16}$$

If any of the roots are complex then the the series contains quasi-periodicities.

Once it has been established that the time series is indeed quasi-periodic the next step is to determine the period. According to Fuller [27] the period of an ARMA process containing an AR(2) component can be determined in the following manner:

Calculate

$$\cos\theta = -\phi_1/2\sqrt{\phi_2} \tag{2.17}$$

where ϕ_i is the i_{th} AR component coefficient.

Solve for θ to obtain

$$\theta = \cos^{-1}(-\phi_1/2\sqrt{\phi_2}) \tag{2.18}$$

The period is given by

$$P = 2\pi/\theta \tag{2.19}$$

Only the USDZAR model was tested as it is the only model with the $\Phi(B)$ of at least order 2. The autoregressive polynomial roots equal -0.3305 + 0.2092i and -0.3305 - 0.2092i. The period of the time series is 2.4379428 $\approx$ 2.44 time units, i.e. 2.44 days.

The periodic nature of the time series does not persist indefinitely. If it did the time series would be predictable. Instead the cycles dampen or decrease over a certain time period. The rate at which the cycles dissipate is determined by the damping factor defined as

$$R = \sqrt{\phi_2} \tag{2.20}$$

where ϕ_2 is as before: see Jenkins and Watts 1968 [28].

The damping factor of this particular time series is $0.391152144 \approx 0.39$. The damping factor is less than one so it leads to a rapidly "declining cyclical" effect in the time series - [27]

Although present, quasi-periodic variability is obviously not of practical importance in the USDZAR time series.

Chapter 3

Univariate Analysis - Alternate Approaches

3.1 Loess Estimation of the Local Variance

An interesting result was obtained from smoothing the series using the Loess function developed by Cleveland *et al* 1992 [17]. The function involves the fitting of a local low order polynomial to a portion of the time series data and using the points on the regression line as estimates of the time series mean. It is a smoothing technique with the degree of smoothing depending on the portion of the data used to estimate the regression line.

The Loess function was applied to the squared USDZAR time series, i.e. the volatility of the USDZAR series, with bandwidth equal to three percent of the time series length. The transformation

$$z_t = y_t/\sqrt{f_t} \tag{3.1}$$

where f_t is the Loess estimate of the time series volatility at time t and y_t is the USDZAR series should then give approximately homoscedastic z_t. A sequence plot of z_t is provided in figure 3.1. The result was a series that closely resembles white noise as seen by the ACF plotted in figure 3.2. This implies that autocorrelation in the USDZAR level series is mainly due to the changeable variance.

An ARCH model was therefore fitted directly to the differenced series. It was found that a G(ARCH)(1,1) model was the most suitable once again. All traces of ARCH were

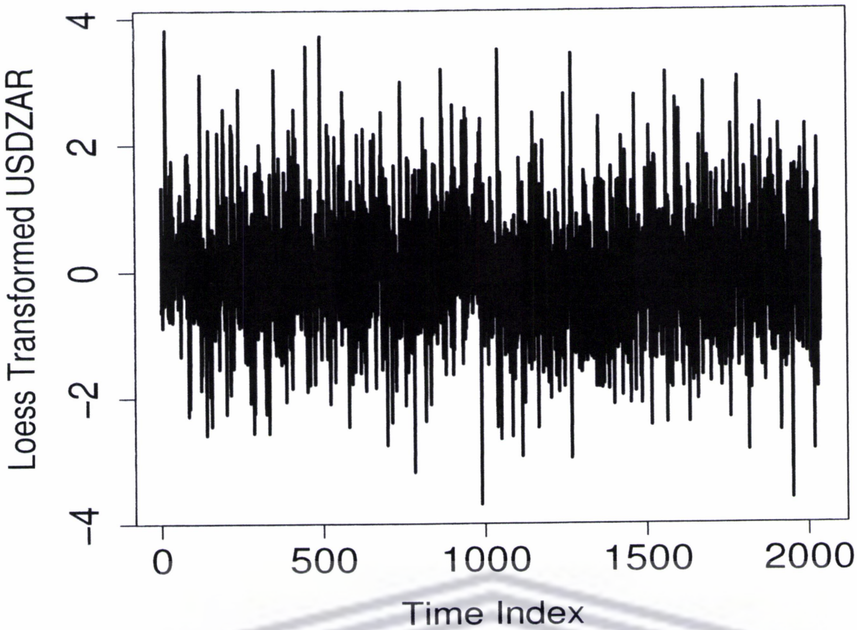

Figure 3.1: Sequence plot of the Loess transformed USDZAR level series

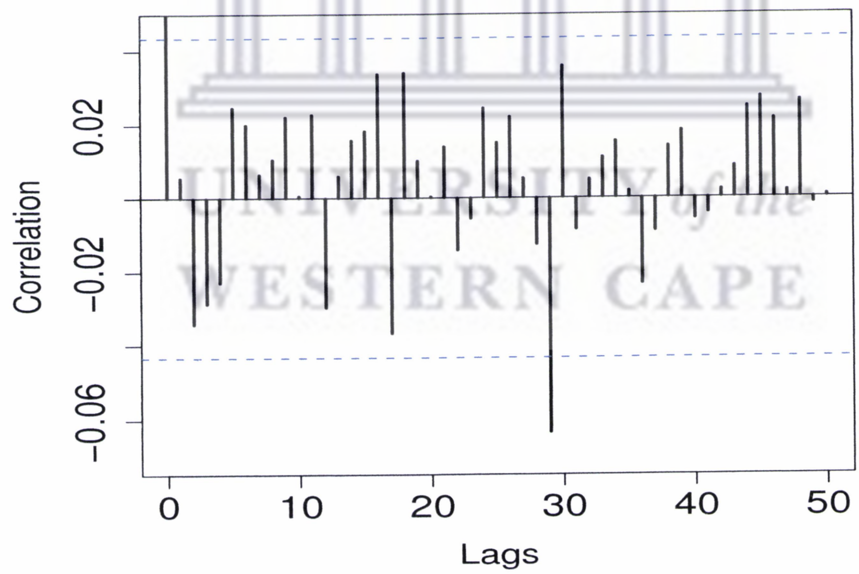

Figure 3.2: ACF of the Loess transformed USDZAR level series

removed from the residuals, as can be seen in figure 3.3, which may be compared to figure 2.28. Each of the parameter estimates was found to be highly significant.

The results of the G(ARCH) model fitted to the differenced level USDZAR series are

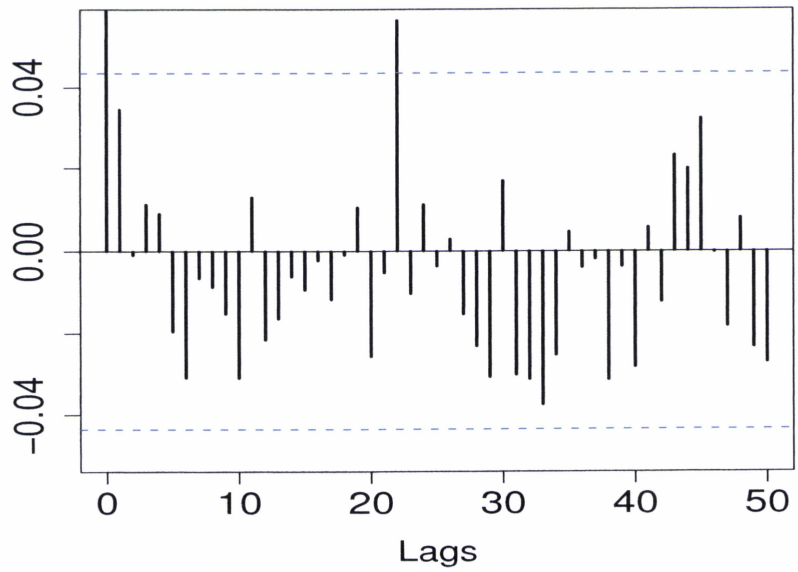

Figure 3.3: USDZAR level series: ARCH in the squared GARCH residuals

summarised in table 3.1. Although it might be tempting to suggest that this model is simpler due to the removal of the ARMA component from the modelling process, the two models are not directly comparable because of the Loess-estimated component.

Parameter	Estimate	Standard Error
α_0	4.85e-07	1.09e-07
α_1	0.12	0.01
β_1	0.89	0.01

Table 3.1: USDZAR level series: GARCH(1,1) summary

This process was repeated for each of the two remaining series. The Repo series yielded no useful results. The volatility in the Alsi series, however, was removed by the loess smoothing technique - see figure 3.4. The ACF contained significant values at lags one and two - see figure 3.5. This tells us that that there is autocorelation present in the series that is not due to the changeable variance. A G(ARCH)(1,1) model was fitted to the differenced series as before and the results are displayed in table 3.2. Once again all parameter estimates are highly significant and all traces of ARCH have been removed from the residuals, as can be seen in figure 3.6, which may be compared to figure 2.26 in section 2.2.2.

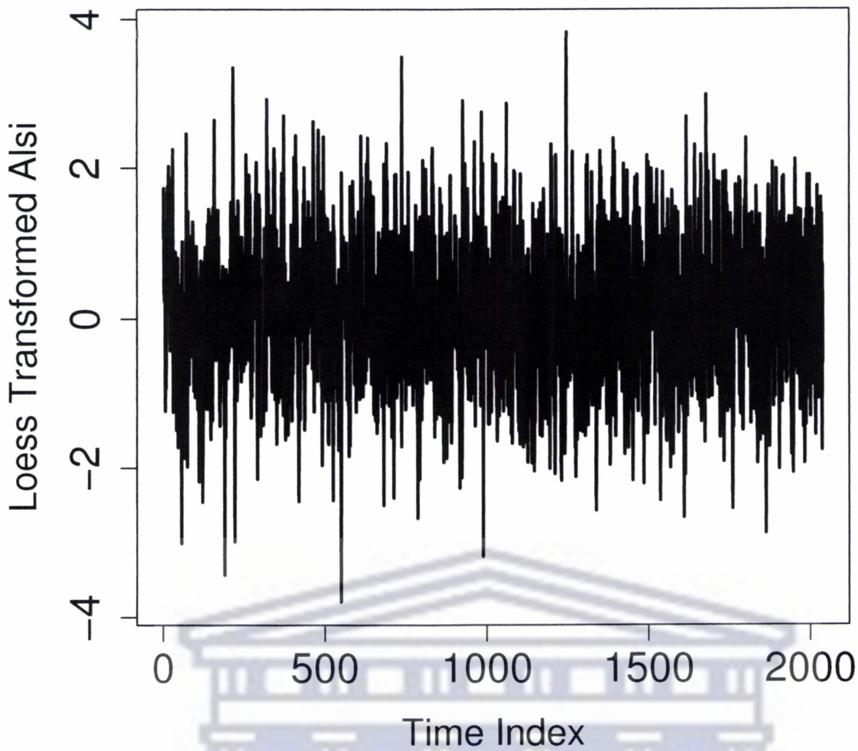

Figure 3.4: Sequence plot of the Loess transformed Alsi level series

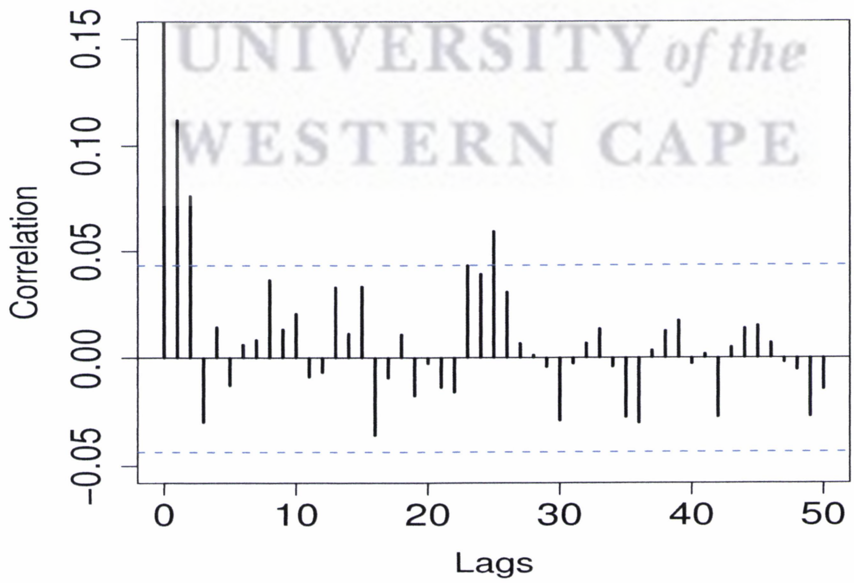

Figure 3.5: ACF of the Loess transformed Alsi level series

Parameter	Estimate	Standard Error
α_0	2.714e-06	7.049e-07
α_1	7.907e-02	7.619e-03
β_1	9.033e-01	9.268e-03

Table 3.2: Alsi level series: GARCH(1,1) summary

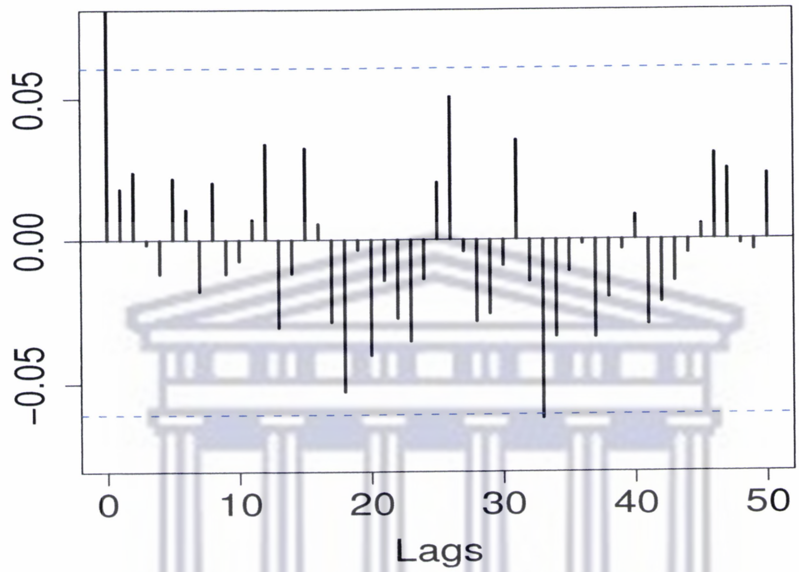

Figure 3.6: Alsi level series: ARCH in the squared GARCH residuals

3.2 Modelling Change in Trend of the USDZAR Series

A possible cause of distortion in the model building results is the change in trend visible in figure 1.2. A cumulative sum (CumSum) plot of the USDZAR series, shown in figure 3.7 , confirms that there is indeed a change in the trend. A parabolic form of the CumSum plot implies that there is a trend present. The CumSum plot for the USDZAR time series consists of two parabolic forms implying the existence of two different trends in the data. The data may need to be differenced twice. Differencing the USDZAR series twice leads to the ACF and PACF shown in figures 3.8 and 3.9. This ACF/PACF combination is characteristic of over-differenced series. This justifies our initial decision to model the singly differenced log-transformed USDZAR series.

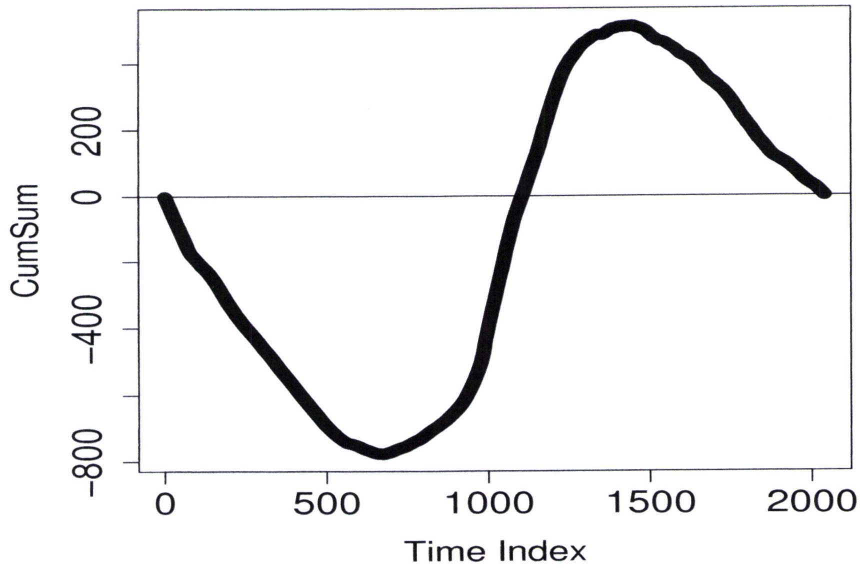

Figure 3.7: CumSum plot of USDZAR

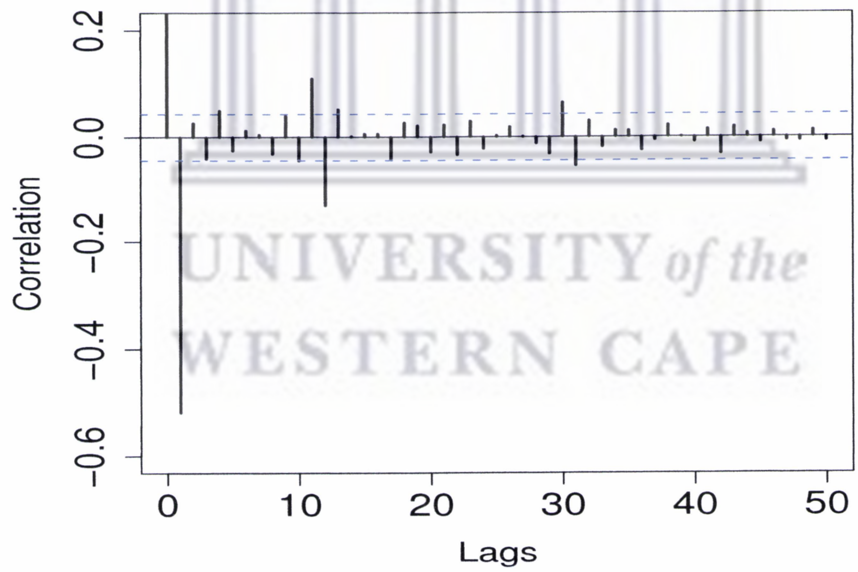

Figure 3.8: ACF of the twice differenced log-transformed USDZAR series

Given that we know that the series contains two distinct trends it might be better to model the USDZAR as two separate series. The most obvious break-point in the series would be at the high point of the initial uptrend, namely the 989^{th} observation. The USDZAR

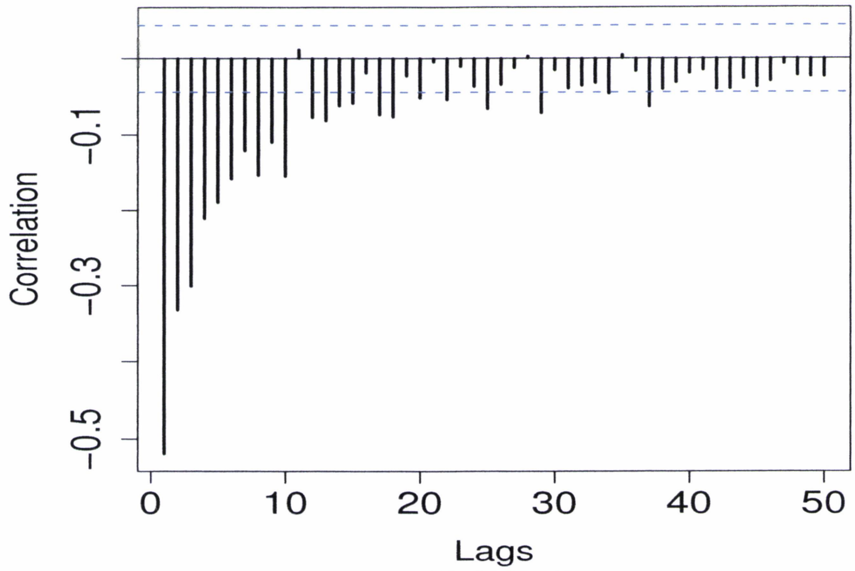

Figure 3.9: PACF of the twice differenced log-transformed USDZAR series

series is therefore split into two parts, the first consisting of the first 989 observations and the second of the observations from 990 to 2040. The two series will be referred to as USDZAR F989 and USDZAR 990L respectively. Both series were differenced before being modelled. This was motivated by the unit root test results displayed in table 3.3.

Time Series	Phillips-Perron Statistic	p-value
USDZAR F989	13.7047	0.99
USDZAR 990L	-6.044	0.7726

Table 3.3: Phillips-Perron unit root test

USDZAR F989

The only model that successfully captured most of the residual autocorrelation in the time series was the ARIMA(3,1,3) model with a constant. The residual ACF and PACF are given in figures 3.10 and 3.11 respectively. Interestingly, all of the parameters are highly significant with the exception of the constant, which is marginally significant (table 3.4). Removal of the constant increases the standard errors for each of the parameter estimates (table 3.4) so it has been included in the model. The QQ plot shown in figure 3.12 reveals that the residuals are once again non-normal.

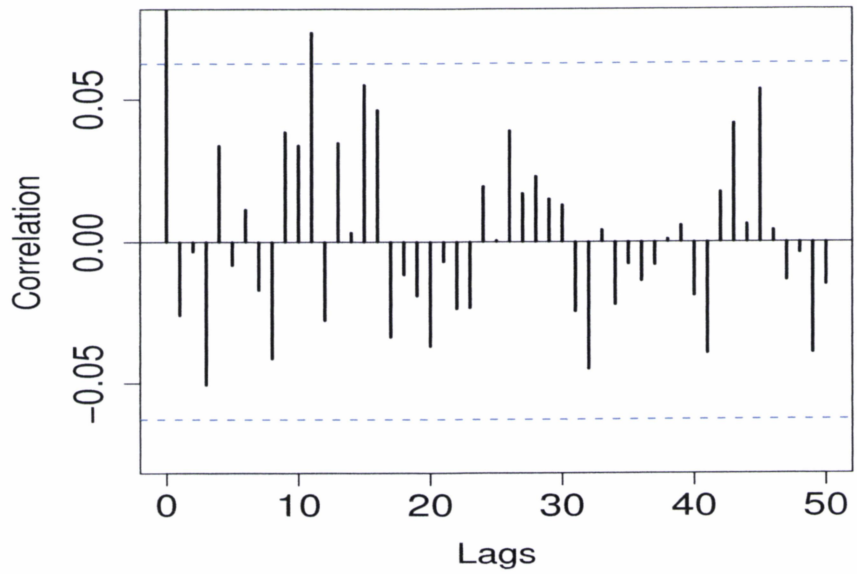

Figure 3.10: ACF of the USDZAR F989 ARIMA model residuals

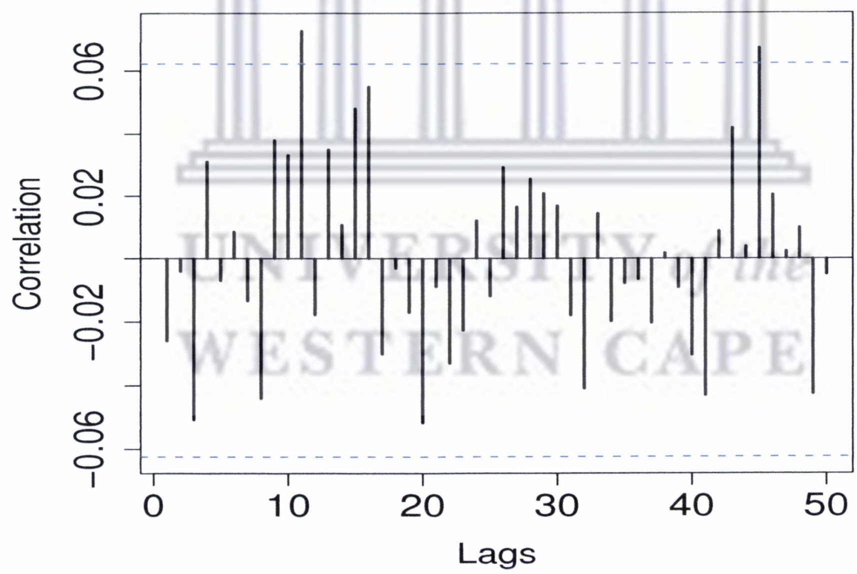

Figure 3.11: PACF of the USDZAR F989 ARIMA model residuals

USDZAR 990L

The most appropriate model for the USDZAR 990L series was found to be an ARIMA(1,1,1) with a constant. The residual autocorrelations obtained sufficiently low Box-Ljung statistics, with corresponding high p-values, for the first 11 lags. The residual ACF and PACF

δ	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	σ^2_ϵ	AIC	BIC
0.002	0.77	-0.50	0.71	0.68	-0.39	0.61	7.3e-05	-6588.348	-6554.078
(0.001)	(0.17)	(0.19)	(0.15)	(0.18)	(0.19)	(0.16)			
	0.59	-0.30	0.70	0.48	-0.21	0.63	7.3e-05	-6587.353	-6557.979
	(0.21)	(0.21)	(0.16)	(0.21)	(0.20)	(0.16)			

Table 3.4: USDZAR F989 model summary

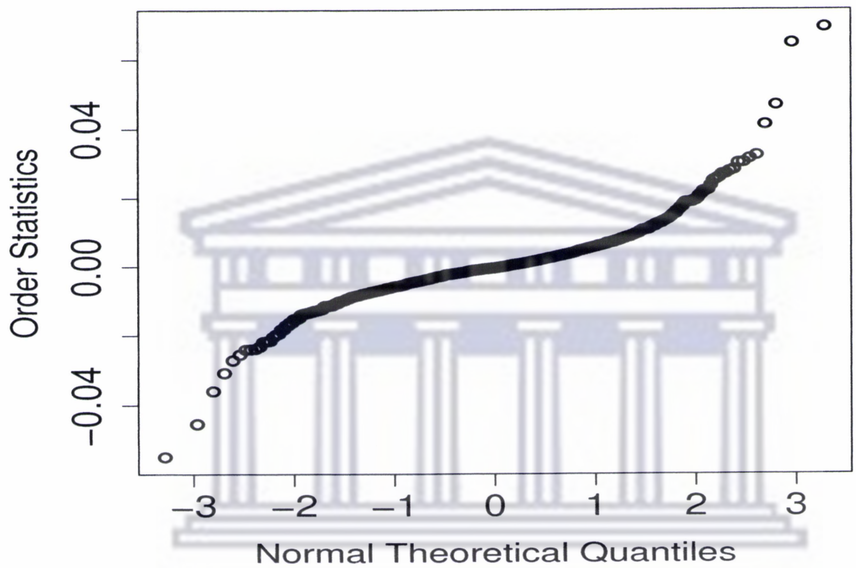

Figure 3.12: QQ plot of the USDZAR F989 ARIMA model residuals

are shown in figures 3.13 and 3.14. The model summary is provided in table 3.5. Each of the parameter estimates is significant at the 6 percent level.

δ	ϕ_1	θ_1	σ^2_ϵ	AIC	BIC
-.001	0.477	.559	1.48e-04	-6272.907	-6258.038
(.000)	(.242)	(.228)			

Table 3.5: USDZAR 990L model summary

The QQ plot in figure 3.15 shows that the model residuals are still non-normal, however, the deviation from normality is less severe than in figure 3.12.

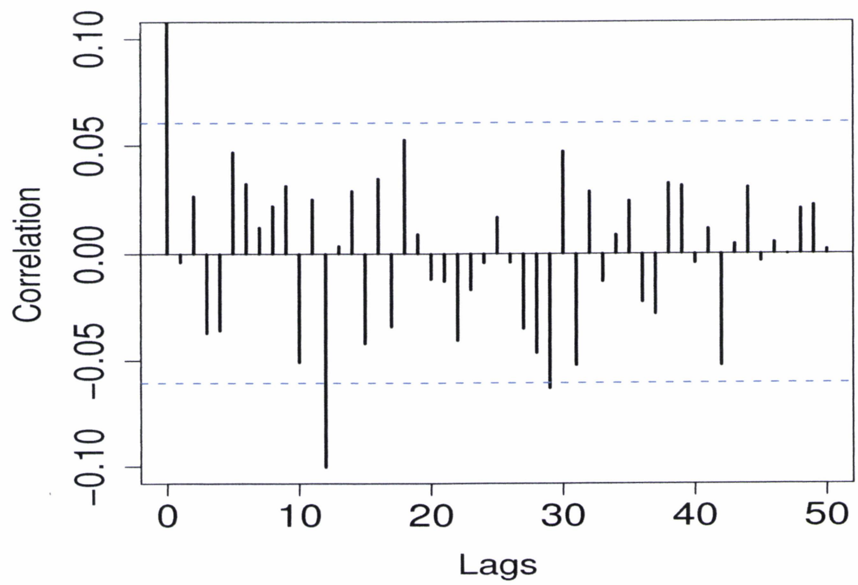

Figure 3.13: ACF of the USDZAR 990L ARIMA model residuals

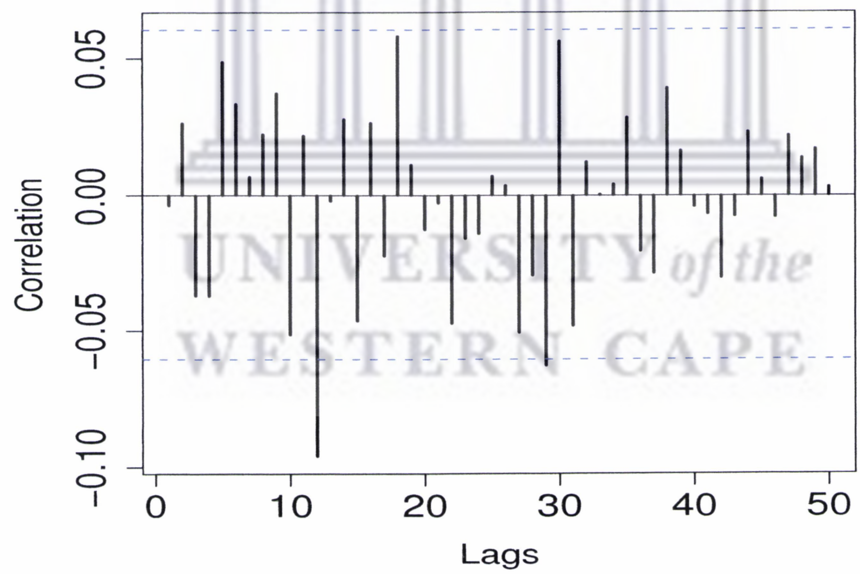

Figure 3.14: PACF of USDZAR 990L ARIMA Residuals

The next objective is to compare the adequacy of the split model against that of the single model fitted in chapter 2.1.2. The models are compared on the basis of information criteria. The AIC and BIC for the split model is defined as

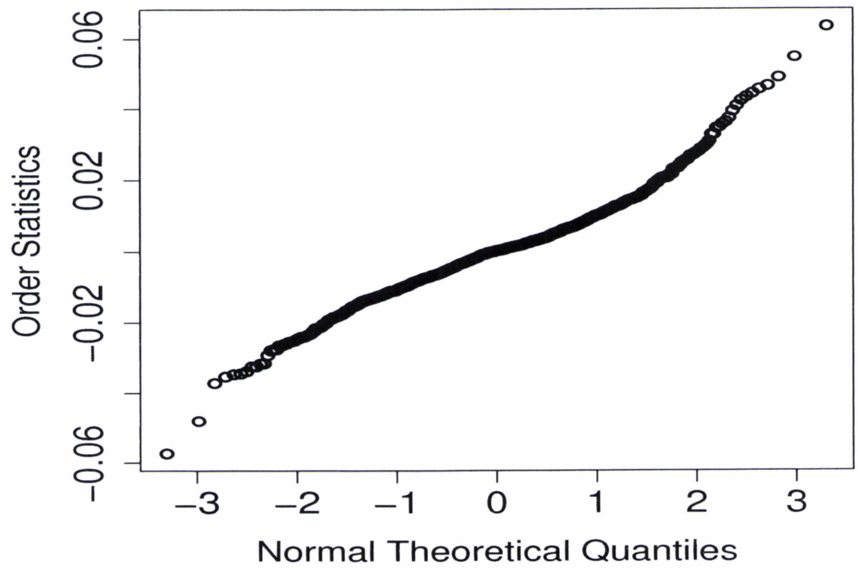

Figure 3.15: QQ plot of the USDZAR 990L ARIMA model residuals

$$AIC_{Total} = -2 * (LL(F989) + LL(990L)) + 2 * (a + b + 1) \quad (3.2)$$

$$BIC_{Total} = -2 * (LL(F989) + LL(990L)) + log(n) * (a + b + 1) \quad (3.3)$$

where LL(x) is the log-likelihood of x, a is the number of parameters in the USDZAR F989 model, b is the number of parameters in the USDZAR 990L model and n is the total number of observations in the combined time series. The total number of model parameters is incremented by one to account for the fact that the split point had to be selected.

The information criteria are displayed in table 3.6. On the basis of this table the conclusion is that the USDZAR series is best modelled as a split series.

Time Series Model	AIC	BIC
Single USDZAR Model	-12587.103	-12559.002
Split USDZAR Model	-12859.256	-12797.439

Table 3.6: Comparison of Information Criteria

3.3 Analysis of Time Intervals Between Repo Changes

As previously explained, the Repo time series is an indexed series where the time between observations is irregular. A possible relationship between the Repo series and the time interval between changes may exist. The series consisting of the times between changes in the Repo series is henceforth referred to as Repo Intervals. The series was, once again, log-transformed before analysis was performed. The sequence plot and histogram of the Repo Intervals series showed that the data are non-normal. The sequence plot in figure 3.16 shows that the bulk of the intervals are very small. Toward the end of the series the periods grow. The histogram in figure 3.17 confirms this. There are very few large values. This is also due to the change in the way the Repo rate is determined. It is modified at Monetary Policy Committee (MPC) meetings only. Often there is no change in the rate for several MPC meetings.

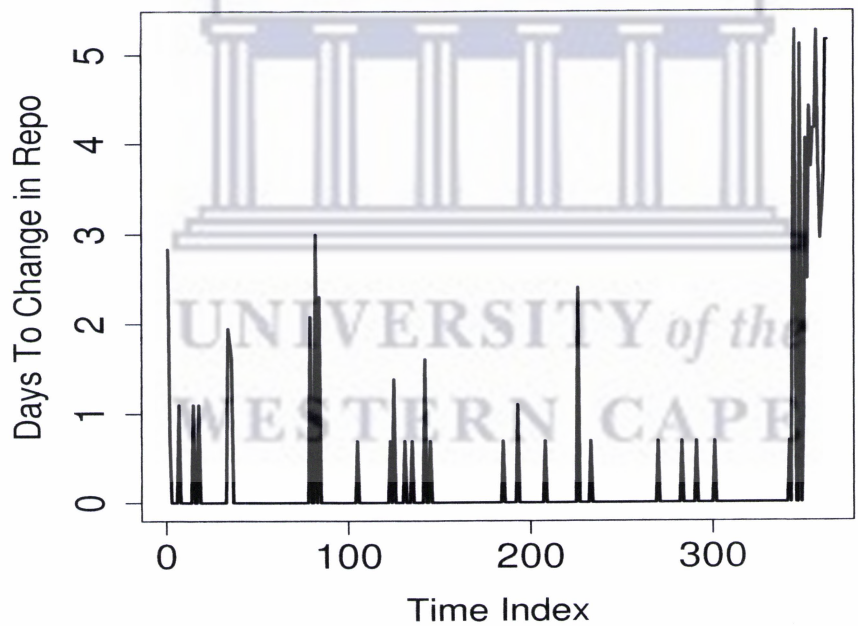

Figure 3.16: Sequence plot of log-transformed intervals, measured in days, between Repo changes

The interval lengths between random events are normally modelled as exponentially distributed. In the case of the log-transformed Repo Intervals series, this distribution does not fit. A contributing factor is the change in the method of determining the Repo rate as discussed above.

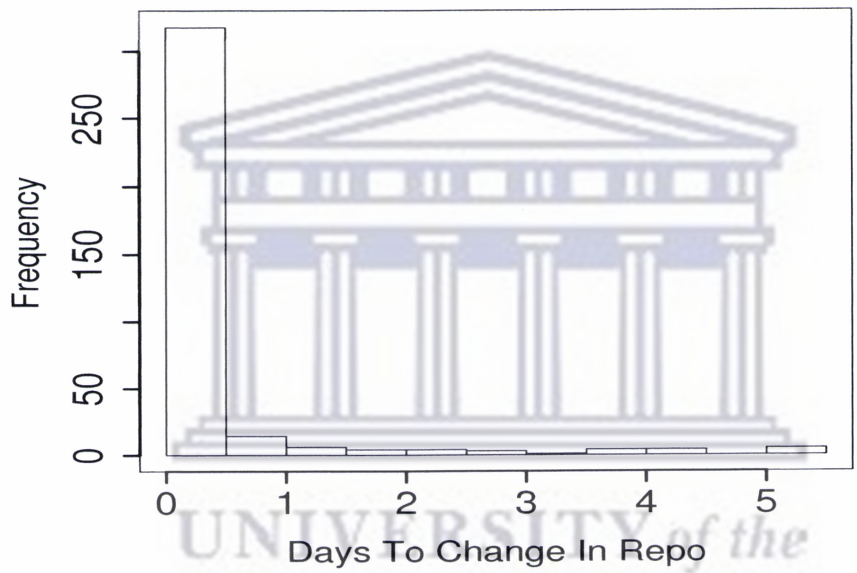

Figure 3.17: Histogram of log-transformed intervals, measured in days, between Repo change

Chapter 4

Multivariate Analysis

4.1 Cointegration Analysis

Cointegration analysis addresses the relationship between two or more inter-related non-stationary time series. This concept is discussed in detail in Hamilton (1994) [7]. The theory states that a vector time series consisting of two or more component time series, all of which contain a unit root, is cointegrated if some linear combination of the component time series does not contain a unit root. This means that the linear combination of the time series has a constant mean and covariance structure. If this is the case, the linear combination can be modelled by applying the same Box-Jenkins methodology outlined earlier.

Let the vector time series $\mathbf{z}_t$ consist of time series x_t and y_t . Each time series is I(1), which means each must be differenced once to remove the unit root and render it I(0) (see [7]).

If x_t and y_t are indeed cointegrated then there exists a vector α such that

$$\alpha' * \mathbf{z_t} \tag{4.1}$$

is I(0).

The Phillips-Ouliaris Cointegration Test is used. A null hypothesis of no cointegration is tested against the alternative hypothesis that the time series are cointegrated. According to Engle and White (1999) [27], the Phillips-Ouliaris test involves the regression of one series against another by ordinary least squares. The residual series obtained through the regression is tested for the presence of a unit root. The Phillips-Ouliaris test is more general than other tests in that it is unaffected by the choice of normalization, i.e. the order in which time series are regressed. Cointegration of the USDZAR and Alsi time series was tested for. Neither series was tested with the Repo series because the observation times of the series are different, due to the definition of the Repo series. The Phillips-Ouliaris test results are shown in table 4.1.

Phillips-Ouliaris Test Statistic	Significance Level
3.0818	0.15

Table 4.1: Cointegration test results

The null hypothesis of no cointegration cannot be rejected at the 15 percent level. This result was not expected as equity traders closely monitor the Rand-Dollar exchange rate for signals on the direction of the movement in the Allshare Index. One would expect that there would be at least some sort of self propagating effect, as traders trying to capitalise on perceived future movements in the Allshare Index would actually cause some of the movement by increasing demand or supply of the stocks in the JSE.

Perhaps a better modelling strategy would involve the USDZAR series as a driver for the Alsi series. There is, of course, no deterministic relationship between the two series. A transfer function model is an appropriate method of checking for a possible directional stochastic relationship between the two series.

4.2 Transfer Function Models

The method of fitting transfer function models involve what is known as pre-whitening (see Koen 1992 [15]). This entails the fitting of a suitable ARIMA model to the input series to obtain a residual series ϵ_t, which is free of autocorrelation, and using the coefficient estimates obtained from the model to filter the output series:

$$\begin{aligned} \Phi(B)i_t &= \Theta(B)\epsilon_t \\ \Rightarrow \epsilon_t &= [\Theta(B)]^{-1}\Phi(B)i_t \end{aligned} \tag{4.2}$$

where i_t is the input series.

The operation $[\Theta(B)]^{-1}\Phi(B)$ is then applied to the output series, o_t, to obtain y_t:

$$y_t = [\Theta(B)]^{-1}\Phi(B)o_t \tag{4.3}$$

The cross correlation function (CCF) of the filtered series, y_t, and the ϵ_t series is generated to test for cross correlation between the two original series. The CCF is used to give indications of possible orders of the transfer function model.

The order of $\Phi(B)$ or the numerator of the transfer model, determines how many lagged values of the input series affect the current value of the output series. The order of $\Theta(B)$ or the denominator, determines the rate of decay in the weights of the lagged values of the input series. The delay τ is the time to the first lagged value of x_t which has a significant effect on y_t. For a more detailed discussion on determining these values consult [14].

As mentioned above (see equation 1.8), the series N_t need not be free of autocorrelation. In this case an ARMA model is fitted to the noise series and the model adequacy is determined as before. Once this ARIMA model has been fitted equation 1.8 can be written as:

$$y_t = C + (\Phi(B)/\Theta(B))x_{t-\tau} + (\psi(B)/\upsilon(B))\nu_t \tag{4.4}$$

where $\upsilon(B)$ is the AR part of the ARMA model for N_t and ψ is the MA part, as explained earlier. Diagnostic checks for the adequacy of transfer function models include the testing of the significance of the estimated parameters, checking that the ν_t series is free of autocorrelation and checking that the CCF of the ν_t and the ϵ_t series obtained earlier, is statistically zero (see [15]). The results of the transfer function modelling follow.

The log-transformed USDZAR series was used as the input series and the log-transformed Alsi series as the output series. The ARIMA(2,1,3) model fitted to the log-transformed USDZAR series (see Section 2.1.2) was used to pre-whiten the log-transformed Alsi series. The CCF of the USDZAR ARIMA model residuals and the pre-whitened Alsi series returned significant values for the cross-correlations at lags 0, 1 and 2 - see figure 4.1. The CCF implies that the current observation of the USDZAR series should be partially explained by the current and two previous observations of the Alsi series. This is of course counter intuitive as our modelling process began with the assumption that the USDZAR series is a driver of the Alsi series. Therefore no practically useful transfer

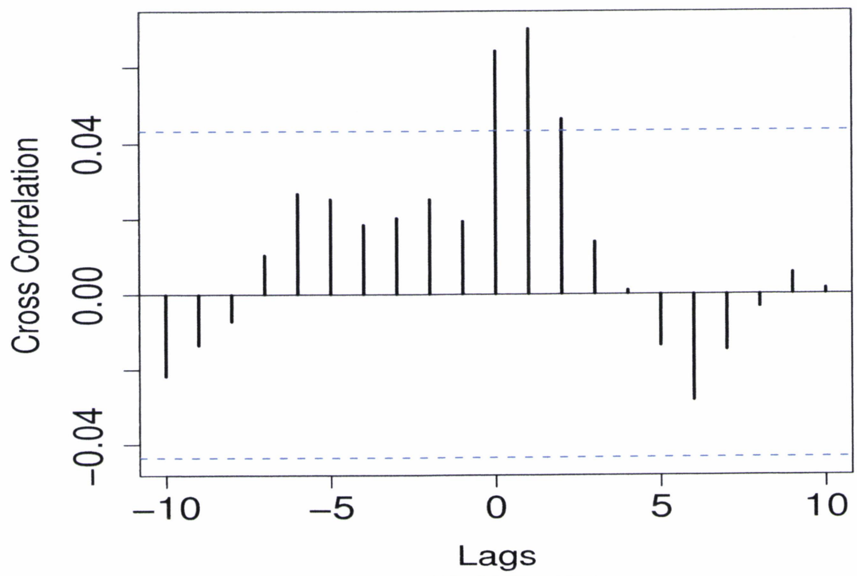

Figure 4.1: CCF of the USDZAR residuals and pre-whitened Alsi series

function model can be fitted. Perhaps the transfer function modeling was unsuccessful because the assumption that the USDZAR series is a driver of the Alsi series, is incorrect.

A different method of creating a transfer function type model involves the use of the residuals of the ARIMA models fitted to the differenced log transformed series. The relationship between these can be studied directly by ordinary least squares linear regression. The Alsi residual series was used as the output (dependent variable) and the USDZAR residual series as the input (independent variable). A CCF was generated for the two sets of residuals. It was found that the only significant cross correlation value occurred at lag zero. A regression line relating the current USDZAR residual to the current value of the Alsi residual was fitted in the following manner:

$$\epsilon_t = \gamma\eta_t + \zeta_t \tag{4.5}$$

where ϵ_t is the log-transformed Alsi ARIMA model residual, γ is a constant, η_t is the USDZAR ARIMA model residual and ζ_t is the regression residual:

$$\begin{aligned} \Phi_A(B)y_t &= \Theta_A(B)\epsilon_t \\ \Phi_U(B)x_t &= \Theta_U(B)\eta_t \end{aligned} \tag{4.6}$$

where subscripts A and U represent the Alsi (y_t) and USDZAR (x_t) series respectively, $\Phi(B)$ is the AR polynomial of the ARIMA model fitted to the particular series and $\Theta(B)$ is the MA polynomial of the same ARIMA model. After rewriting the ARIMA models to obtain expressions for ϵ_t and η_t, equation 4.5 can be expressed as:

$$\begin{aligned} \frac{\Phi_A(B)}{\Theta_A(B)} y_t &= \gamma \frac{\Phi_U(B)}{\Theta_U(B)} x_t + \zeta_t \\ \Rightarrow y_t &= \gamma \frac{\Theta_A(B)\Phi_U(B)}{\Phi_A(B)\Theta_U(B)} x_t + \frac{\Theta_A(B)}{\Phi_A(B)} \zeta_t \\ &= \gamma \Pi(B) x_t + \Lambda(B) \zeta_t \end{aligned} \tag{4.7}$$

where $\Pi(B)$ and $\Lambda(B)$ are infinite order polynomials of the form 1-$\sum_i^n \pi_i(B)^i$ and 1-$\sum_i^n \lambda_i(B)^i$ respectively. The regression analysis provided some interesting results. A highly significant coefficient for the USDZAR residual input was obtained. This coefficient corresponds to γ in equation 4.5. In addition, the residual series generated by the linear regression contained no strongly significant autocorrelation: see figure 4.2.

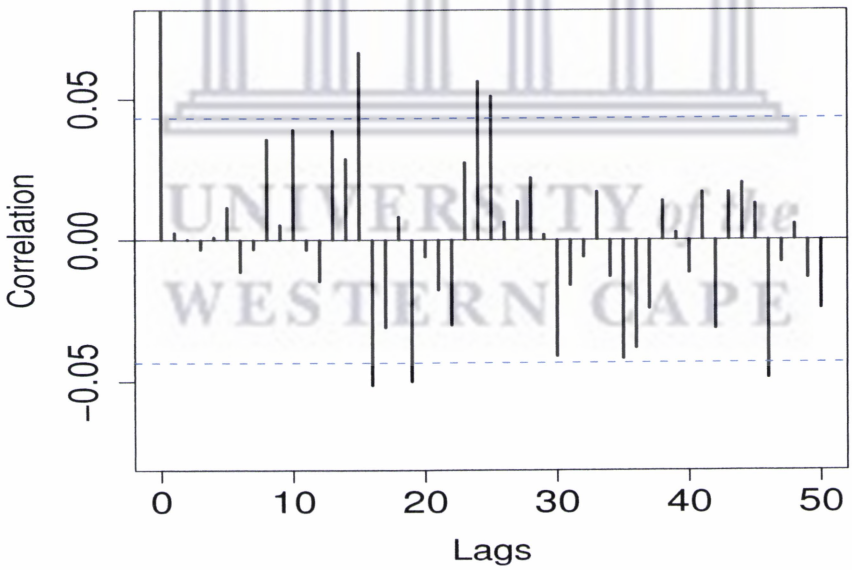

Figure 4.2: ACF of the residuals obtained in regression analysis

It was found that the infinite order polynomial parameter estimates tended to zero rather quickly as shown in figures 4.3 and 4.4.

It seems a transfer function type model, consisting of a few lagged values of the log-transformed USDZAR series and a noise process consisting of a few lagged values of ζ_t,

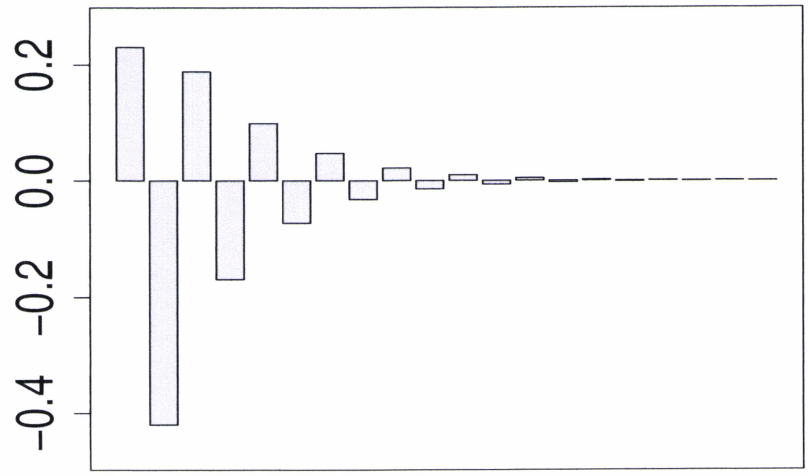

Figure 4.3: Bar-plot of $\Pi(B)$ scaled by γ

Figure 4.4: Bar-plot of $\Lambda(B)$

can be developed for the log-transformed Alsi series. This may be a fruitful avenue for further research, although it appears that calculation of coefficient standard errors will be quite involved.

4.3 VAR Models

Another method of studying the relationship between multiple time series is to fit vector autoregressive (VAR) models. Unlike transfer functions and ARIMA models, VAR models allow possible feedback effects inherent in the time series to be studied (see [14]). Transfer function models require that one of the time series be the output series and the other an input series. The output series depends on lagged values of the input series. In the case of VAR models, each time series being analysed consists of lagged values of itself and lagged values of the other time series. In a sense the time series both act as input series and output series.

Model adequacy can once again be tested by studying the significance of the parameter estimates and checking for autocorrelation in the residuals. Models can be compared on the basis of information criteria as before. In this case, however, the preferred information criteria are the Hannan-Quinn (HQ) criterion and the Schwarz-Bayesian (BIC) criterion (see Lütkepohl 1993 [16]). These criteria are less likely to favour over-parameterised models.
The information criteria are calculated as follows:

$$\begin{aligned} HQ(m) &= ln\,|\Sigma(m)| + (2 * ln(ln(T))/T) * m * k^2 \\ BIC(m) &= ln\,|\Sigma(m)| + (ln(T)/T) * m * k^2 \end{aligned} \tag{4.8}$$

Where m is the maximum lag order, Σ is the estimated residual covariance matrix, k is the number of parameters estimated and T=N-k where N is the number of observations in the time series being modelled.

At first glance it seems that the results of the VAR modelling are promising, but closer inspection tells another story. Both the HQ and BIC criteria suggest a first order VAR model is best. Once fitted it was found that only the diagonal elements of the coefficient matrix Φ_1 in (1.9) are significant, i.e. $\alpha_{1,2}$ and $\alpha_{2,1}$ are effectively zero in

$$\begin{bmatrix} y_{A,t} \\ y_{U,t} \end{bmatrix} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \begin{bmatrix} y_{A,t-1} \\ y_{U,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{A,t} \\ \epsilon_{U,t} \end{bmatrix} \tag{4.9}$$

This means that the two individual series are de-coupled (aside from possible correlation between $\epsilon_{A,t}$ and $\epsilon_{U,t}$). This point might, to some extent, have been anticipated from the results of the cointegration analysis.

In addition the VAR(1) model, fitted with and without a constant had autocorrelated residuals. In particular the PACF returned significant correlation estimates at several low order lags. Increasing the order of the VAR model did not lead to any improvement in the results.

A summary of the VAR(1) model is provided in table 4.2 and the residual correlation properties are shown in figures 4.5 to 4.9. In table 4.2, indices 1 and 2 refer to the USDZAR and Alsi series respectively.

Parameter	Estimate	Standard Error
$\Phi_1(1,1)$	-0.066881	0.022276
$\Phi_1(1,2)$	-0.004662	0.020707
$\Phi_1(2,1)$	-0.02085	0.02380
$\Phi_1(2,2)$	0.13581	0.02212

Table 4.2: A summary of the VAR(1) model fitted to the bivariate time series with USDZAR and Alsi components

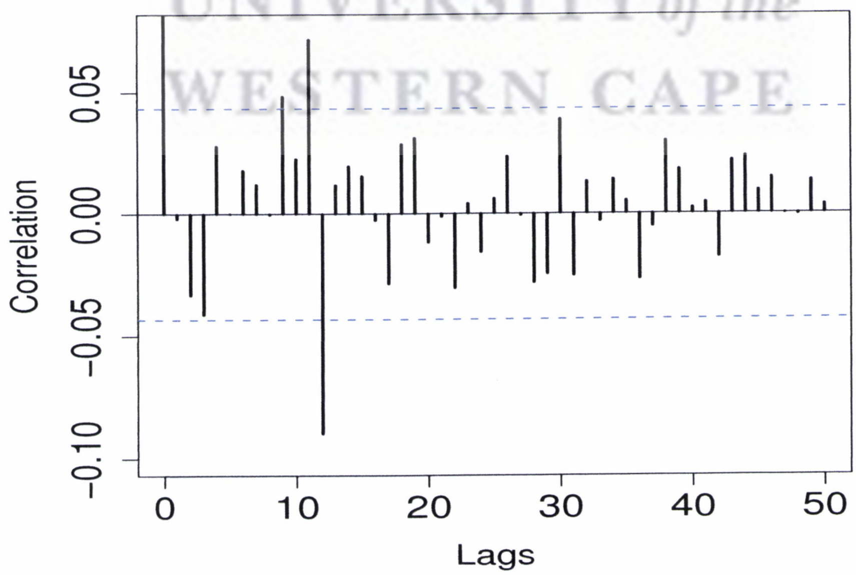

Figure 4.5: ACF of the USDZAR VAR(1) model residuals

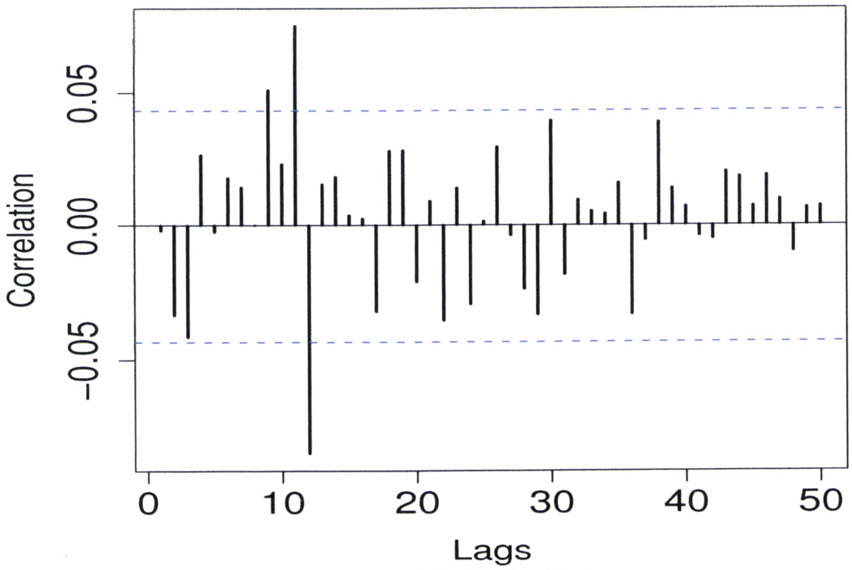

Figure 4.6: PACF of the USDZAR VAR(1) model residuals

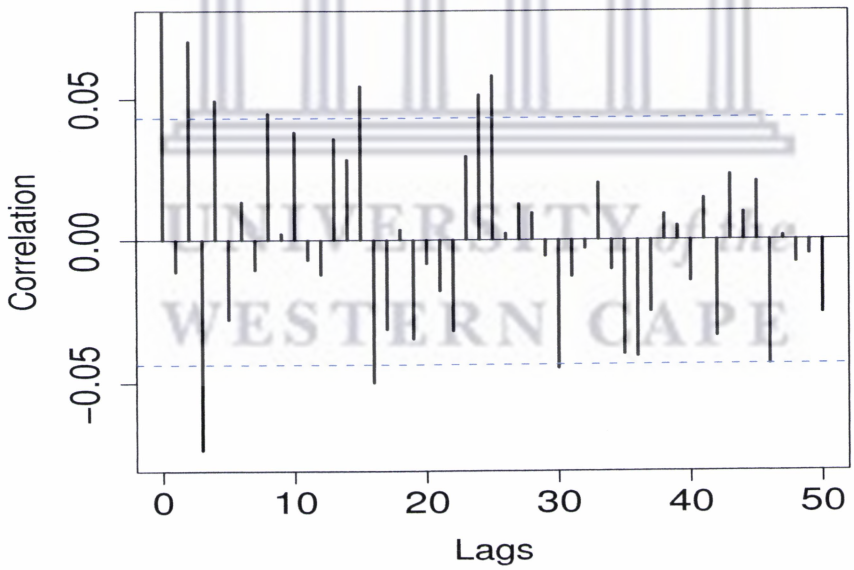

Figure 4.7: ACF of the Alsi VAR(1) model residuals

It seems that no obvious statistical relationship between these two financial time series can be unearthed using VAR modelling procedures, apart from the cross correlation between the innovations $\epsilon_{A,t}$ and $\epsilon_{U,t}$ - see figure 4.9.

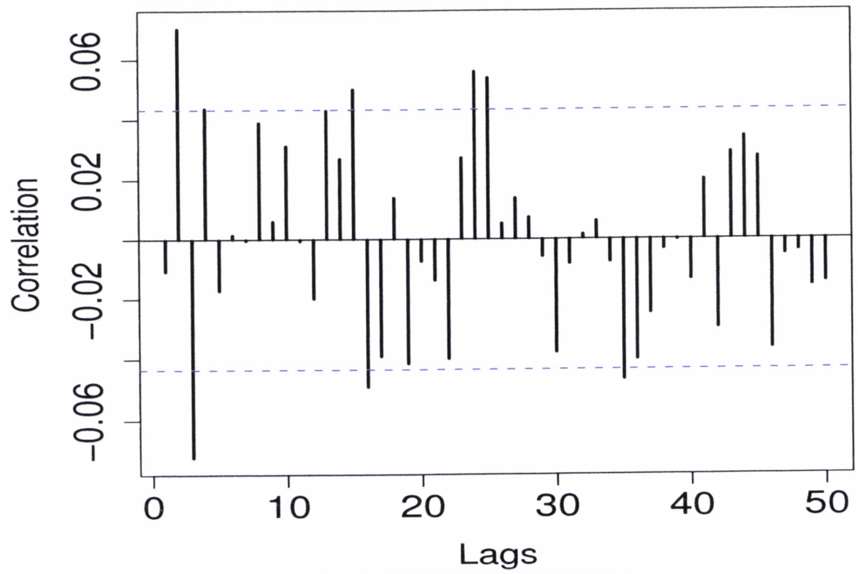

Figure 4.8: PACF of the Alsi VAR(1) model residuals

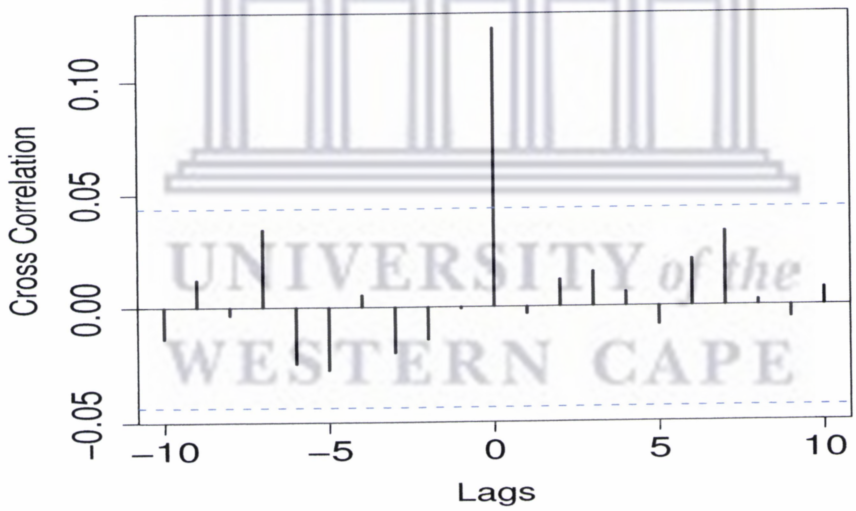

Figure 4.9: CCF of the Alsi and USDZAR VAR(1) model residuals

4.4 Correlation Between Repo Intervals and Repo Series

The correlation between the interval lengths and the Repo series, as well as the differenced Repo series, was tested. In the former case the first observation of the Repo series was deleted. The results can be viewed in table 4.3.

Series	Pearson Correlation	p-value	Kendall's τ	p-value
Repo	-0.3478928	9.131e-12	-0.1500976	0.0003328
Differenced Repo	-0.162311	0.001920	-0.04750724	0.2560

Table 4.3: Correlation between the Repo series and Repo intervals

There is significant negative correlation between the Repo series and the intervals between Repo changes, i.e. the longer the time interval since the previous change the smaller the change in the Repo rate - a rather counter-intuitive result. This result could serve as the basis for a bivariate time series model, but this will not be pursued here.

Chapter 5

Conclusions

This study has been successful in some areas and unsuccessful in others. An adequate ARIMA model was obtained for each of the time series. As was previously mentioned, the Repo time series posed the most problems. It is not clear that ARIMA modelling is the best approach to use to accurately model the Repo rate. An adequate G(ARCH) model was obtained for the residuals of the Alsi and USDZAR models. Once again, the Repo series proved to be tougher. The G(ARCH) model for the Repo was not as successful at capturing the salient information in the residuals.

Cointegration analysis revealed no stationary long-term relationship between the Alsi and USDZAR series. The significant cross-correlation values obtained on the CCF, between the two series, indicate that a relationship does exist, but not in the expected input-output sense. The VAR results show that this relationship cannot be modelled as a low order VAR form. Perhaps vector autoregressive moving average (VARMA) models would be more suited to describing the relationship. These findings could be due to the absence of any relationship between the series or, as mentioned by Pierce 1977 [20], the relationship is present but cannot be uncovered using the current data and methodology.

The relationship between the Repo and the other two series could not be tested using the above methods because of structural differences. As mentioned before, perhaps a multivariate model could be fitted to the Repo and intervals between the Repo changes. According to Kang 1986 [22], indirect modelling of defined variables outperforms direct modelling in terms of forecasting accuracy. He recommends investigating which variables are constituents of the defined variable and deriving forecasts for those. These forecasts are then used to derive forecasts of the defined variable. With this in mind, the Repo series might be more accurately modelled by first modelling the level of inflation, for example.

Frequency domain methods, and in particular the use of bi-spectra, have not been explored in this study. Perhaps this would assist in discovering any relationships between the time series.

The scope of this mini-thesis was restricted to studying a few aspects of the three time series. Obviously, much more could be learnt by including other economically relevant series, such as the Gold price, foreign interest rates, the rate of inflation and the trade balance.

Chapter 6

Appendix 1 - Software packages used in the analyses.

SPSS 14.0

Microsoft Excel

MATLAB R2006a

R 2.3.0

Chapter 7

Appendix 2 - MATLAB program used to test for outliers.

```
function b = outlier(pis,res,alsi,psi)
format bank;                             %Assigning starting values%
alsivar=alsi;
nresvar=res;
resvar=res;
pis_star=0;
n=size(res,1);
I=1;
outlier(1)=5;
io_test=zeros(1,n);
ao_test=zeros(1,n);
    for a=1:n
        squareres(a)=resvar(a)^2;
    end
        variance=(1/n)*(sum(squareres));

while outlier(I)>3
  I=I+1;
  resvar=nresvar;
```

```
for k=1:n   %performing calculation of outlier effect for each time period%


    sum_pis_square=0;
    for j=1:n-k                          %determining the sum of the pi terms%
        pis_square(j)=pis(j)^2;          %up to total-occurence of outlier%
        sum_pis_square=pis_square(j)+sum_pis_square;
    end
        final_sum_pis_square(k)=1+sum_pis_square;


    sum_pis_star=0;
    for t=1:n-k                 %determining pi*(F)et over remaing time periods%
        pis_star(t)=pis(t)*resvar(k+t);
        sum_pis_star=pis_star(t)+sum_pis_star;
    end
        final_sum_pis_star(k)=sum_pis_star;
                                            %determining outlier effect%
wao(k)=(resvar(k)-final_sum_pis_star(k))/(final_sum_pis_square(k));
wio(k)=resvar(k);
                                            %determining lamda's%
lamda_ao(k)=(sqrt(final_sum_pis_square(k))*wao(k))/sqrt(variance);
lamda_io(k)=wio(k)/sqrt(variance);

ao(I)=max(abs(lamda_ao));
io(I) =max(abs(lamda_io));
outlier(I)=max(ao(I),io(I));

      if outlier(I)==abs(lamda_ao(k)) && outlier(I)>3
           ao_test(I)=k;
```

```
        elseif outlier(I)==abs(lamda_io(k)) && outlier(I)>3
          io_test(I)=k;

         else

      end




end                                              %end of k loop%
                               %determining absoulute max lamda and its position%


realout(I)=max(ao_test(I),io_test(I));
if realout(I)==ao_test(I) && realout(I)~=realout(I-1)
    ao_occurs(I)=ao_test(I);
    io_occurs(I)=0;

elseif realout(I)==io_test(I) && realout(I)~=realout(I-1)
    io_occurs(I)=io_test(I);
    ao_occurs(I)=0;

else
    io_occurs(I)=0;
    ao_occurs(I)=0;
end
                               %modifying series and residuals for additive outliers%


if ao_occurs(I)>=1

alsivar(ao_occurs(I))=alsivar(ao_occurs(I))-wao(ao_occurs(I));
```

```
      for q=1:n
         if q<ao_occurs(I)
             nresvar(q)=resvar(q);
         elseif q==ao_occurs(I)
             nresvar(q)=resvar(q)-wao(ao_occurs(I));
         elseif q>ao_occurs(I)
             nresvar(q)=resvar(q)+wao(ao_occurs(I)).*pis(q+1-ao_occurs(I));
         end
         nresvar_squared(q)=nresvar(q)^2;
      end
     new_sum_squared_res=sum(nresvar_squared);
elseif ao_occurs(I)==0
     new_sum_squared_res=variance;

end
                              %modifying series and residuals for innovation outliers%
if io_occurs(I)>=1
     resvar(io_occurs(I))=resvar(io_occurs(I))-wio(io_occurs(I));
      for p=1:n
         if p<io_occurs(I)
             alsivar(p)=alsivar(p);
         elseif p>=io_occurs
             alsivar(p)=alsivar(p)-wio(io_occurs(I))*psi(p+1-io_occurs(I));
         end
         resvar_squared(p)=resvar(p)^2;
      end
     new_sum_squared_res=sum(resvar_squared);
elseif io_occurs(I)==0
     new_sum_squared_res=variance;

end
```

```
if io_occurs(I)==0 && ao_occurs(I)==0;
    outlier(I)=0;
end



new_variance=(1/n)*new_sum_squared_res;



end


io_occurs
ao_occurs
```

Bibliography

[1] Franses, P.H. 1998. Time series models for business and economic forecasting. Cambridge, United Kingdom: Press Syndicate of the University of Cambridge.

[2] Maddala, G.S.and Kim, I. 1998. Unit Roots, Cointegration, and Structural Change. New York: Cambridge University Press.

[3] Wei, W.S. 1989. Time Series Analysis: Univariate and Multivariate Methods. Redwood City, CA: Benjamin Cummings.

[4] Graupe, D. 1989. Time Series Analysis, Identification and Adaptive Filtering, 2nd Edition. Malabar, Florida: Robert E. Krieger Publishing Co.

[5] Enders, W. 1948. Applied Econometric Time Series. Hoboken, New Jersey: John Wiley.

[6] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. c1994. Time Series Analysis, Forecasting and Control. Englewood Cliffs, New Jersey: Prentice Hall.

[7] Hamilton, J.D. 1994. Time Series Analysis. New Jersey: Princeton University Press.

[8] Bodie, Z., Kane, A. and Marcus, A.J. 2005. Investments, Sixth Edition. New York: McGraw-Hill/Irwin.

[9] Nattrass, N. 1997. Macroeconomics: Theory and Policy in South Africa. Claremont, South Africa: David Phillips Publishers.

[10] Lipsey, R.G., Courant, P.N., Purvis, D.D. and Steiner, P.O. 1993. Economics, Tenth Edition. New York: HarperCollins College Publishers.

[11] Nattrass, N., Wakeford, J. and Muradzikwa, S. 2002. Macroeconomics: Theory and Policy in South Africa, Third Revised Edition. Claremont, South Africa: David Phillips Publishers.

[12] Burnham, K.P. and Anderson, D.R. 2002. Model Selection and Multi-Model Inference. New York: Springer.

[13] Lange, O.F. and Grubmüller, H. 2006. Generalized Correlation for Biomolecular Dynamics. Proteins: Structure, Function and Bioinformatics, Volume 62, pages 1053 - 1061.

[14] Pankratz, A. 1991. Forecasting with Dynamic Regression Models. New York: John Wiley and Sons.

[15] Koen, C. 1992. Transfer Function Analysis of Ultraviolet Observations of NGC 5548. Monthly Notices of the Royal Astronomical Society, Volume 262, pages 823 - 830.

[16] Lütkepohl, H. 1993. Introduction to Multiple Time Series. Heidelberg: Springer-Verlag.

[17] Cleveland, W.S. and Grosse, E. and Shyu, W.M. 1992. In: Statistical Models in S [edited by Chambers, J. and Hastie, T.]. Pacific Grove, California: Wadsworth and Brooks/Cole.

[18] Mabert, V.A. and Radcliffe, R.C. 1974. A Forecasting Methodology as Applied to Financial Time Series. The Accounting Review, Volume 49, pages 61 - 75.

[19] Bollerslev, T., Chou, R.Y. and Kroner, K.F. 1992. ARCH modelling in Finance. Journal of Econometrics, Volume 52, pages 5 - 59.

[20] Pierce, D.A. 1977. Relationships and Lack Thereof Between Economic Time Series, with Special Reference to Money and Interest Rates. Journal of the American Statistical Association, Volume 72, pages 11 - 26.

[21] Baillie, T. and Chung, H. 2001. Estimation of GARCH Models from the Autocorrelations of the Squares of a Process. Journal of Time Series Analysis, Volume 22, pages 631 - 650.

[22] Kang, H. 1986. Univariate ARIMA Forecasts of Defined Variables. Journal of Business and Economic Statistics, Volume 4, pages 81 - 86.

[23] Shumway, R.H. 1988. Applied Statistical Time Series Analysis. New Jersey: Prentice Hall.

[24] Harvey, A.C. 1990. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge: Press Syndicate of the University of Cambridge.

[25] Bera, A.K. and Jarque, C.M. 1980. Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. Economics Letters, Volume 6, pages 255 - 259.

[26] Fuller, W.A. 1996. Introduction to Statistical Time Series. New York: John Wiley and Sons.

[27] Engle, R.F. and White, H. 1999. Cointegration, Causality and Forecasting. New York: Oxford University Press Inc.

[28] Jenkins, G.M. and Watts, D.G. 1968. Spectral Analysis and its Applications. California: Holden-Day, Inc.