# TRENDS AND VOLATILITY IN MACROECONOMIC 

## TIME SERIES DATA



## Lebotsa Daniel Metsileng

Submitted in Partial Fulfillment of the Requirements for the Degree Master of Commerce in Statistics in the Faculty of Commerce and Administration<br>University of North West<br>Mafikeng

March 2001

## Acknowledgements

Very special thanks go to my supervisor, Dr. P.A.E. Serumaga-Zake, for his valuable contribution to the success of this study. Thanks for your patience I really appreciate it. I also thank my parents (Annah and Jacob Metsileng) whom I dedicate this work to, for their moral support. To my family as a whole, I am proud because of your support. I would like to give my special thanks to the research committee of the Faculty of Commerce and Administration for affording me an opportunity. And finally to God be the glory.
L.D.M

March 2001

## CONTENTS

Chapter 1: Literature Review and Importance of Study ..... 1
1.1 Introduction ..... 1
1.2 Survey of the Literature ..... 1
1.3 Importance of Study ..... 4
1.4 Research Methodology ..... 4
1.4.1 Applied Statistical Methods ..... 4
1.4.2 Data and Source of Data ..... 5
1.4.3 Computer Aids ..... 5
1.5 Research Outline ..... 5
1.6 Research Limitation ..... 5
Chapter 2: Review of Some Tests for Stationarity ..... 6
2.1 Introduction ..... 6
2.2 The Augmented Dickey-Fuller (ADF) Unit Root Test ..... 8
2.3 The Phillips-Perron (PP) Unit Root Test ..... 11
2.4 Instrumental Variable (IV) Unit Root Test ..... 12
2.5 The Generalised Least Squares (GLS) Unit Root Test ..... 14
2.6 Multiple Unit Roots Test ..... 15
2.7 Joint Unit Root Test: A Multivariate Setting ..... 16
2.8 Practical Examples ..... 18
2.8.1 Application of Unit Root Tests to Series 1 ..... 18
2.8.2 Application of Unit Root Tests to Series 2 ..... 20
2.9 Conclusion ..... 23
Chapter 3: Trends and Some Decomposition Methods ..... 24
3.1 Introduction ..... 24
3.2 Trends in Macroeconomic Time Series Data ..... 25
3.2.1 Stochastic Trend Models ..... 25
3.2.2 Deterministic Trend Models ..... 27
3.3 Decomposition of Time Series ..... 28
3.3.1 The Beveridge-Nelson Decomposition Method ..... 28
3.3.2 Decomposition of Seasonal Time Series ..... 31
3.3.3 Decomposition of Series into Deterministic and Stochastic Parts ..... 32
3.4 Practical Examples ..... 32
3.4.1 Example 1: Decomposition of Series 3 ..... 33
3.4.2 Example 2: Decomposition of Series 4 ..... 36
3.5 Conclusion ..... 41
Chapter 4: Volatility Modelling in Macroeconomic Time Series ..... 42
4.1 Introduction ..... 42
4.2 Autoregressive Moving Average (ARMA) Processes ..... 43
4.3 Autoregressive Conditional Heteroskedasticity (ARCH) Processes ..... 43
4.3.1 Estimation of the ARCH Process ..... 44
4.3.2 Testing for ARCH ..... 48
4.3.3 Forecasting with an ARCH Process ..... 50
4.4 Extensions of the ARCH Process: A Review ..... 51
4.4.1 The ARCH-in-Mean (ARCH-M) Process ..... 51
4.4.2 The Generalised ARCH (GARCH) Process ..... 54
4.4.3 Integrated GARCH (IGARCH) Process ..... 58
4.4.4 Exponential GARCH (EGARCH) Process ..... 60
4.4.5 Threshold ARCH (TARCH) Process ..... 61
4.4.6 The Component ARCH (CARCH) Process ..... 62
4.4.7 Asymmetric CARCH (ACARCH) Process ..... 63
4.5 Practical Example ..... 63
4.6 Conclusion ..... 71
Chapter 5: Bootstrapping ARCH Processes ..... 72
5.1 Introduction ..... 72
5.2 Basic Bootstrap Concepts ..... 72
5.2.1 Properties of EDF ..... 73
5.2.2 Drawing Bootstrap Samples ..... 74
5.2.3 Parametric Bootstrap ..... 74
5.3 Jacknife-after-Bootstrap and Confidence Intervals ..... 75
5.3.1 Bootstrap and Jacknife Confidence Intervals ..... 75
5.4 Evaluating the Bootstrapped ARCH Process ..... 76
5.5 Conclusion ..... 79
Chapter 6: Conclusions and Recommended Research Area ..... 80
6.1 Summary and Findings ..... 80
6.2 Recommended Research Area ..... 81
Appendices ..... 82
References ..... 96

## CHAPTER 1 <br> INTRODUCTION

### 1.1 Introduction

The aim of this study is basically to discuss the ARCH processes and to make some concepts accessible that are hitherto difficult to assimilate from most texts. It seeks to assess the performance of the ARCH process using bootstrap methods on microeconomic and financial time series data.

The concept of Autoregressive conditional Heteroskedasticity (ARCH) was first introduced by Bachelier (1900) followed by a period of long silence. The concept was however revived by Engle (1982) who formally formulated a model to capture all the earlier stylized facts. In this research, we focus on providing an account of recent theoretical advances in ARCH models and their applications in macroeconomic and financial time series. By ARCH, we mean the phenomenon of conditional Heteroskedasticity in general and all models to capture this phenomenon, and hence does not refer only to Engle's original model.

### 1.2 Survey of the Literature

In recent years, experience in modeling the conditional mean of macroeconomic and financial time series data has emphasized the role of persistence of shocks (volatility), and a large literature has emerged on testing for and estimating unit roots in the Autoregressive representations of univariate processes. Other related topics are the advances in the area of common stochastic trends and co-integration.

In their seminal work, Nelson and Plosser (1982) established that most US macroeconomic series could be characterized as difference-stationary rather than trend-stationary processes, implying that shocks (volatility) were persistent - a characteristic consistent with Real Business Cycle (RBC) models, where fluctuations are driven by technology shocks. The issue of testing for unit roots boils down to checking whether a series in question should be modelled in levels or first differences. A number of tests for a unit root have been proposed, with the most popular being the Dickey-Fuller (DF) test, the Augmented Dickey-Fuller
(ADF) test, and the Phillips-Perron (PP) test. Given the time series $\left\{X_{t}: t=1,2,3, \ldots, N\right\}$, the model favoured by Dickey and Fuller (1982) is

$$
\Delta X_{t}=C+(\beta-1) X_{t-1}+\sum_{j=1}^{p-1} \delta_{j} \Delta X_{t-j}+\varepsilon_{t}, \quad\left\{\varepsilon_{t}\right\} \sim i . i . d\left(0, \sigma_{\varepsilon}^{2}\right)
$$

where C is a constant, and $\Delta X_{t-j}=X_{t-j}-X_{t-j-1}$ for $j=1,2,3, \ldots, p-1 ; p$ is the order of the autoregression model that may be applied in the analysis. If the series contains a time trend, then the preferred model is

$$
\Delta X_{t}=\left(\alpha_{0}+\alpha_{1} t\right)+(\beta-1) X_{t-1}+\sum_{j=1}^{p-1} \delta_{j} \Delta X_{t-j}+\varepsilon_{t}, \quad\left\{\varepsilon_{t}\right\} \sim \text { i.i.d }\left(0, \sigma_{\varepsilon}^{2}\right)
$$

For both the ADF and PP tests, the presence or otherwise of drift and/or trend assumes that the null hypothesis of a unit root, $H_{0}: \beta=1$, is not rejected using Fuller's (1976) $\hat{\tau}_{\tau}$ distribution of the t-statistic of $\hat{\beta}$, then it is necessary to proceed to test the joint hypothesis $H_{0}: \beta=1, \alpha_{1}=0$ using the F-statistic, $\Phi_{3}$ given in Dickey and Fuller (1981). Dolado et al (1990) also points out that if the trend is significant under $H_{0}$, then normality of the tstatistic of $\hat{\beta}$ follows, and hence the standardized normal tables should be used. However, if the trend is not significant, then $H_{0}$ should be tested with $\alpha_{1}=0$ using $\hat{\tau}_{\mu}$ in Fuller (1981). If the constant under $H_{0}$ is significant, then the test for the unit root should be repeated using the standardized normal, otherwise Fuller's $\hat{\tau}$ should be used instead.

As lengthy a talk as it seem to be regarding testing for unit roots, not that much can be said about co-integration. Many ideas in co-integration can be explained by relating them to univariate time series data. It can be shown that co-integration tests are a natural extension of the tests for unit roots in single-equation time series models.

Uncertainty is central in financial markets where markets sometimes appear quite calm and at other times highly volatile. An accurate prediction of volatility is of great importance for at
least two reasons. First, expected future volatility is an important input for all dynamic trading and option-pricing models derived along lines first set out by Black and Scholes (1973). Second, expected future volatility is also an important input for static pricing/hedging, portfolio selection, and margining problems, where volatility is typically used as an explicit measure of risk. A large variety of volatility predictors are available in literature. Apart from directly being based on historical data (time series), one might use observed option prices and an option-pricing model to obtain the volatility forecast that is theoretically implied by market prices. Our study will, however, be limited to predictors based on time series only. Within the class of predictors, one other approach is to model the time series behaviour of volatility explicitly using a two-stage method by first calculating a series of say monthly sample volatilities, and then fitting standard ARIMA models for this time series of monthly estimates - see Poterba and Summers (1986) and French et al (1987). This method assumes implicitly that volatility is constant within a month and becomes variable only for longer periods. Moreover, this method essentially reduces the available amount of information and thus resulting in an increase in the variability of the ARIMA model volatility forecasts. Furthermore, Chou (1988) notes that the parameter estimates are extremely sensitive to the sampling frequency for which the time series of volatility estimates is calculated. This means that the assumption of constant volatility may not be appropriate in some situations. A modification to changing volatility is thus required.

Thanks to Engle (1982), changing volatility can now be modeled using parameterizations of his linear Autoregressive Conditional Heteroskedasticity (ARCH) models. In the ARCH (p) model, the conditional variance is written as a linear function of the first p past squared innovations. Some of the parameterizations of Engle's ARCH models are the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) due to Bollerslev (1986) and the Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) due to Nelson (1991). The issue of proper modeling of the long-run dependencies in the conditional mean of macroeconomic and financial time series led to the formulation of the Integrated Generalized Autoregressive Conditional Heteroskedasticity (IGARCH) by Engle and

Bollerslev in 1986. IGARCH models possess many of the features of the unit root processes for the mean.

There are a number of other models where the conditional variance not only depends on the past variance but also subject to random noise. For instance, Hull and White (1987) used a continuous time stochastic stock return volatility model for the pricing of stock options. Using this model, they were able to explain a significant part of the strike price bias that is typically found in Standards and Poor (S\&P) 500 index option-implied volatilities. In our study, however, our discussions on volatility will be based on Engle's ARCH models and some of its parameterized models.

Bootstrap methods have been a powerful tool when it comes assessing statistics and estimators. In this study, it is employed to the distribution of the replicated variances from an ARCH process.

### 1.3 Importance of the study

Given the importance of predicting volatility in macroeconomic and financial time series, many approaches have been proposed in the literature. Notable among them is the class of autoregressive heteroskedasticity (ARCH) processes originally introduced by Engle (1982). In many macroeconomic and financial time series data, the general assumption of constant variance in the disturbance term is violated. This called for the ARCH concept, where series are modeled taking into consideration the changing variances at different time points. As many methods have been proposed in the literature, it is equally important to assess the performance of these methods. Focusing only on ARCH process, we seek to assess its performance by using the Bootstrap methods, where it is assumed that the original data has no underlying distribution. It is also of interest to illustrate how these processes work.

### 1.4 Research Methodology

### 1.4.1 Applied Statistical Methods

Statistical and econometric methods are used in this study. They include regression, and ARCH modelling.

### 1.4.2 Data and Source of data

All data sets to be used in this study are from the Quarterly Bulletin of Reserved Bank of South Africa and Statistics South Africa.

### 1.4.3 Computer Aids

Data analyses in this study are carried out using SAS and S-plus.

### 1.5 Research Outline

Chapter 1 provides an introduction to the study and briefly provides the structure of the study. In Chapter 2, detailed discussions on the tests for unit roots are given as the concept of ARCH now serves as a standard diagnostic tool in the analysis of macroeconomic time series. The basic ARCH models which capture various stylized facts and their interpretations are described and discussed thoroughly in Chapter 3. In Chapter 4, we show that the basic ARCH models are unable to capture all the observed phenomena, for instance, excess kurtosis and high degree of non-linearity. Generalizing the basic ARCH models to capture these phenomena then becomes a subject matter in this chapter. In Chapter 5, forecasting with ARCH models is considered. We also present bootstrapping ARCH processes. Chapter 6 concludes and gives recommendations. Practical illustrations would be given after every chapter to illustrate the methodology.

### 1.6 Research Limitations

Interpretations of all results will apply only to a panel of series being studied.

## CHAPTER 2

## REVIEW OF SOME TESTS FOR STATIONARITY

### 2.1 Introduction

Most of the time series used in modeling are non-stationary in nature. By nonstationarity, we mean that the mean, variance, and autocovariances may depend on time $t$. A time series $\left\{X_{t}: t=1,2,3, \ldots, N\right\}$ is therefore said to be stationary if its mean, variance, and autocovariance are independent of time. In Box-Jenkins setting, if the mean of the series is less than its corresponding standard deviation, it is representable as

$$
\begin{equation*}
X_{t}=\sum_{j=1}^{p} \phi_{j} X_{t-j}+\varepsilon_{t}+\sum_{k=1}^{q} \theta_{k} \varepsilon_{t-k} \tag{2.1.1}
\end{equation*}
$$

where $\left\{\phi_{j}: j=1,2,3, \ldots, p\right\}$ are the autoregressive parameters of order $p$, and $\left\{\phi_{j}: j=1,2,3, \ldots, q\right\}$ are the moving average parameters of order $q$. If, however, the mean of the series happens to be greater than the standard deviation, an adjustment made to 2.1.1 yields

$$
\begin{gather*}
\text { UWWU } \\
x_{t}=c+\sum_{j=1}^{p} \phi_{j} x_{t-1}+\varepsilon_{t}-\sum_{k=1}^{s} \theta_{k} \varepsilon_{t-k}, \tag{2.1.2}
\end{gather*}
$$

If the series is driven by a polynomial trend, further adjustments to (2.1.2) yields the representation

$$
\begin{equation*}
X_{t}=\sum_{i=0}^{m} \alpha_{i} t^{i}+\sum_{j=1}^{p} \phi_{j} X_{t-j}+\varepsilon_{t}-\sum_{k=1}^{q} \theta_{k} \varepsilon_{t-k}, \tag{2.1.3}
\end{equation*}
$$

In equation (2.1.1) to (2.1.3), $\varepsilon_{t}$ is a white noise process with mean zero and variance $\sigma^{2}$, that is

$$
\begin{equation*}
\varepsilon_{t} \sim \text { i.i.d. } N\left(0, \sigma_{\varepsilon}^{2}\right) \tag{2.1.4}
\end{equation*}
$$

$X_{t}$ is non-stationary in levels, but the differenced series

$$
\begin{equation*}
\Delta X_{t}=X_{t}-X_{t-1} \tag{2.1.5}
\end{equation*}
$$

is stationary, thus $x_{t}$ is said to contain a unit root or simply be a differenced-stationary (DS) series. Consider the case where $\mathrm{p}=1$ and $\mathrm{q}=0$, we obtain the autoregressive AR(1), process

$$
\begin{equation*}
X_{t}=\phi_{1} X_{t-1}+\varepsilon_{t} \tag{2.1.6}
\end{equation*}
$$

If $\left|\Phi_{1}\right|<1,2.1 .6$ is said to be stationary so that

$$
\begin{gather*}
X_{t}=\phi_{1} X_{t-1}+\varepsilon_{t} \\
\left(1-\phi_{1} B\right) X_{t}=\varepsilon_{t} \\
X_{t}=\left(1-\phi_{1} B\right)^{-1} \varepsilon_{t}=\left(1+\phi_{1} B+\phi_{2}^{2} B^{2}+\ldots .\right) \varepsilon_{t} \\
\text { or } \quad X_{t}=\varepsilon_{t}+\phi_{1} \varepsilon_{t-1}+\phi_{2} \varepsilon_{t-2}+\ldots \tag{2.1.7}
\end{gather*}
$$

It follows that

$$
\begin{gather*}
E\left(X_{t}\right)=0  \tag{2.1.8a}\\
\operatorname{var}\left(X_{t}\right)=\frac{\sigma_{\varepsilon}^{2}}{1-\phi_{1}^{2}}  \tag{2.1.8b}\\
\operatorname{cov}\left(X_{t}, X_{t-k}\right)=\frac{\phi_{1}^{2} \sigma^{2}}{1-\phi_{1}^{2}}, \quad k=1,2, \ldots, \tag{2.1.8c}
\end{gather*}
$$

$X_{t}$ is said to have a unit root if $\phi_{1}=1$.

In this case, (2.1.7) becomes

$$
\begin{equation*}
X_{t}=X_{t-1}+\varepsilon_{t}=X_{0}+\varepsilon_{t}+\varepsilon_{t-1}+\varepsilon_{t-2}+\ldots \ldots+\varepsilon_{1}, \tag{2.1.9}
\end{equation*}
$$

assuming that the process starts at $t=0$. For the particular case

$$
\begin{align*}
E\left(X_{t}\right) & =X_{0}  \tag{2.1.10a}\\
\operatorname{var}\left(X_{t}\right) & =t \sigma^{2}  \tag{2.1.10b}\\
\operatorname{cov}\left(X_{t}, X_{t-k}\right) & =\sqrt{\frac{t-\bar{k}}{t}}, \quad k=1,2, \ldots, \tag{2.1.10c}
\end{align*}
$$

Formal tests for non-stationarity have now become a standard starting point in applied time series analysis. Several test statistics have been proposed to test the need for differencing the series before modelling. Notable among these are due to Dickey and Fuller (1979), Phillips and Perron (1988), and Hall (1989).

In this chapter, we review these three unit root test procedures and support these with two real data sets. The remainder of this chapter is structured as follows. In section 2.2, we review the Augmented Dickey-Fuller ADF test. Section 2.3 outlines the Phillips-Perron (PP) test. The Instrumental Variable (IV) test due to Hall is discussed in Section 2.5. Section 2.6 and Section 2.7, respectively, handle the multiple unit root tests and joint unit root test. Section 2.8 provides illustrations of three unit root tests using two time series data sets, Series 1 and Series 2. Series 1 comprises the rand-dollar exchange rate in cents while Series 2 is made up of coin and banknotes in circulation. Section 2.9 summarizes the chapter.

### 2.2 The Augmented Dickey-Fuller (ADF) Test

Consider the AR(1) process with

$$
\begin{equation*}
X_{t}=\phi_{1} X_{t-1}+\varepsilon_{t}, \quad \quad \varepsilon_{t} \sim \text { i.i.d. } \mathrm{N}\left(0, \sigma_{\varepsilon}^{2}\right) \tag{2.2.1}
\end{equation*}
$$

Subtracting $X_{t-1}$ from both side of (2.2.1) yields

$$
\begin{align*}
X_{t}-X_{t-1} & =\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} \\
\text { or } \quad \Delta X_{t} & =\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} . \tag{2.2.2}
\end{align*}
$$

If a constant term is included in the model, we obtain

$$
\begin{equation*}
\Delta X_{t}=c+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} \tag{2.2.3}
\end{equation*}
$$

Similarly, if $X_{t}$ is driven by a linear time trend, then the autoregression we consider is

$$
\begin{equation*}
\Delta X_{t}=\left(\alpha_{0}+\alpha_{1} t\right)+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} \tag{2.2.4}
\end{equation*}
$$

It can be shown that if the $\varepsilon_{t}$ are not i.i.d, then the autoregressions preferred and the Augmented Dickey-Fuller (ADF) autoregressions should be

$$
\begin{align*}
\Delta X_{t} & =\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t}  \tag{2.2.5a}\\
\Delta X_{t} & =c+\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t}  \tag{2.2.5b}\\
\text { or } \quad \Delta X_{t} & =\left(\alpha_{0}+\alpha_{1} t\right)+\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t} . \tag{2.2.5c}
\end{align*}
$$

In what follows, $p$ is selected to ensure that the $\varepsilon_{t}$ are uncorrelated. For the $\operatorname{AR}(1)$ process in 2.2.1, the maximum likelihood estimator of $\phi_{1}$ the least squares estimator

$$
\begin{equation*}
\hat{\phi}_{1}=\frac{\sum_{t=1}^{N} X_{t} X_{t-1}}{\sum_{t=1}^{N} X_{t-1}^{2}} \tag{2.2.6}
\end{equation*}
$$

Substituting $X_{t}=\phi_{1} X_{t-1}+\varepsilon_{t}$ in (2.2.6) yields

$$
\begin{align*}
\hat{\phi}_{1} & =\frac{\sum\left(\phi_{1} X_{t-1}+\varepsilon_{t}\right) X_{t-1}}{\sum X_{t-1}^{2}}=\frac{\phi \sum X_{t-1}^{2}+\sum X_{t-1} \varepsilon_{t}}{\sum X_{t-1}^{2}} \\
\Rightarrow \quad \hat{\phi}_{1} & =\phi_{1}+\frac{\sum X_{t-1} \varepsilon_{t}}{\sum X_{t-1}^{2}} \\
\text { or } \quad \hat{\phi}_{1}-\phi_{1} & =\frac{\sum X_{t-1} \varepsilon_{t}}{\sum X_{t-1}^{2}} . \tag{2.2.7}
\end{align*}
$$

Under the null hypothesis of a unit root, $H_{0}: \phi_{1}=1$, and hence (2.2.7) becomes

$$
\begin{equation*}
\hat{\phi}_{1}-1=\frac{\sum X_{t-1} \varepsilon_{t}}{\sum X_{t-1}^{2}} . \tag{2.2.8}
\end{equation*}
$$

The resultant likelihood ratio test is a function of

$$
\begin{equation*}
\hat{\tau}_{d f}=\frac{\hat{\phi}_{1}-1}{\operatorname{Se}\left(\hat{\phi}_{1}-1\right)} \tag{2.2.9}
\end{equation*}
$$

where $\quad \operatorname{Se}\left(\hat{\phi}_{1}-1\right)=\sqrt{\frac{\sum_{t=2}^{N}\left(X_{t}-\hat{\phi}_{1} X_{t-1}\right)^{2}}{(N-2) \sum X_{t-1}^{2}}}$.

It is obvious that under this null hypothesis, a regression of $\Delta X_{t}$ on $X_{t-1}$ will give a coefficient on $X_{t-1}$ which is an estimate of 0 , since $\phi-1=1-1=0$. However, under the alternative hypothesis $H_{0}:|\phi|<1, \phi-1 \neq 0$ and hence a regression of $\Delta X_{t}$ on $X_{t-1}$ is appropriate. Similarly, if a constant term is included in the unit root autoregression (2.2.1), a regression of $\Delta X_{t}$ on a constant and $X_{t-1}$ is deemed appropriate. Lastly, a linear trend indicated in (2.2.1) suggests regressing $\Delta X_{t}$ on a constant, time and $X_{t-1}$.

When $\phi_{1}=1$, the process generating $x_{t}$ is $\Gamma(1)$. This implies that $X_{t-1}$ will not satisfy the standard assumptions needed for asymptotic analysis. Consequently, Dickey and Fuller (1979) employes Monte Carlo methods to compute the non-standard percentiles for the distributions under the null hypothesis of the unit root. The null hypothesis is rejected if the test statistic is less than the corresponding critical values tabulated by Dickey and Fuller. Otherwise, it is accepted.

If the autoregressive model is of a higher order, the unit root regressions are augmented by lagged differences and $\Delta X_{t-j}$. For example if the sample pacf suggests an $\operatorname{AR}(2)$ process, then the appropriate unit root regression to consider is

$$
\Delta X_{t}=\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{2-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t}
$$

or

$$
\begin{equation*}
\Delta X_{t}=\left(\phi_{1}-1\right) X_{t-1}+\gamma_{1} \Delta X_{t-1}+\varepsilon_{t} \tag{2.2.10}
\end{equation*}
$$

which suggests a regression of $\Delta X_{t}$ on $X_{t-1}$ and $\Delta X_{t-1}$. Where appropriate, a constant term or a linear trend is included in (2.2.10). The inclusion of the terms $\Delta X_{t-j}$ leaves the asymptotic distribution of the parameters of interest unchanged.

### 2.3 The Phillips-Perron (PP) Test

In this section, we review some theoretical background for a unit root test procedure proposed by Phillips and Perron (1988). We shall hereafter refer to this test procedure as the PP test. The unit root test regression is any of the $\operatorname{AR}(1)$ processes

$$
\begin{align*}
& \Delta X_{t}=\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t}  \tag{2.3.1}\\
& \Delta X_{t}=c+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t}  \tag{2.3.2}\\
& \Delta X_{t}=\left(\alpha_{0}+\alpha_{1} t\right)+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} \tag{2.3.3}
\end{align*}
$$

The PP test is non-parametric in nature and has the tendency to correct serial correlation that may be present in the error term, $\varepsilon_{t}$. This test procedure is non-parametric in that the correction in $\varepsilon_{t}$ uses an estimate of the spectrum of $\varepsilon_{t}$ at frequency zero that is robust to heteroskedasticity and autocorrelation of unknown form. The procedure employs the Newey-West (1987) consisting estimate

$$
\begin{align*}
& \quad \xi^{2}=\Gamma_{0}+2 \sum_{k=1}^{\gamma}\left[1-\frac{k}{q+1}\right] \Gamma_{k}  \tag{2.3.4}\\
& \text { where } \quad \Gamma_{k}=\frac{1}{N} \sum_{t=k+1}^{N} \hat{\varepsilon}_{t} \hat{\varepsilon}_{t-k}, \tag{2.3.5}
\end{align*}
$$

and $\Gamma$ is the truncation lag determined by the expression

$$
\begin{equation*}
\Gamma=\text { floor }\left[4\left(\frac{N}{100}\right)^{\frac{2}{9}}\right] \tag{2.3.6}
\end{equation*}
$$

The computed PP test statistic is given by

$$
\begin{equation*}
\hat{\tau}_{p p}=\left(\frac{t_{\phi_{1}-1}}{\xi}\right) \Gamma_{0}^{\frac{1}{2}}-\frac{N}{2}\left(\frac{\xi^{2}-\Gamma_{0}}{\xi \hat{\sigma}}\right) \operatorname{Se}\left(\phi_{1}-1\right) \tag{2.3.7}
\end{equation*}
$$

where $t_{\phi_{1}-1}$ is the $t$-statistic of $\left(\phi_{1}-1\right), \operatorname{Se}\left(\phi_{1}-1\right)$ is the standard error of $\left(\phi_{1}-1\right)$, and $\hat{\sigma}$ is the standard error of the test regression. The asymptotic distributions of the PP test statistics are the same as those of the ADF test statistics. Here again, the null hypothesis of a unit root $\mathrm{H}_{0}: \phi_{1}=1$ is rejected if $\hat{\tau}_{p p}$ is less than the appropriate critical value at some level of significance.

### 2.4. Instrumental Variable (IV) Unit Root Test

In his Monte Carlo study of the empirical powers of some unit root tests, Schwert (1989) observed that the statistics of an earlier version of unit root test proposed by Phillips
(1987a) do not perform well in finite samples in the presence of negative moving average errors. Motivated by the problem, Hall (1989) proposed estimation by instrumental variable (IV) as an alternative to the use of non-parametric corrections. For the AR(1) process

$$
\begin{equation*}
X_{t}=\phi_{1} X_{t-1}+u_{t} \quad \text { where } u_{t}=\varepsilon_{t}+\sum_{k=1}^{q} \varepsilon_{t-k} \tag{2.4.1}
\end{equation*}
$$

It is shown that under the null hypothesis of a unit root $H_{0}: \phi_{1}=1$ the instrumental variable $\hat{\phi}_{1}^{(I V)}$ of $\phi_{1}$ has the standard Dickey-Fuller distribution. For example, let our date generating process (dgp) be

$$
\begin{equation*}
X_{t}=\phi_{1} X_{t-1}+u_{t} \tag{2.4.2}
\end{equation*}
$$

where $u_{t}=\varepsilon_{t}-\theta_{1} \varepsilon_{t-1}$ and $\varepsilon_{t} \sim$ i.i.d. $\mathrm{N}\left(0, \sigma_{\varepsilon}^{2}\right)$. Then the instrumental variable estimator, $\hat{\phi}_{1}^{(I V)}$ of $\phi_{1}$ using $X_{t-2}$ as an instrument for $X_{t-1}$ when $\phi_{1}=1$ is given by

$$
\begin{equation*}
\hat{\phi}_{1}^{I V}=\frac{\sum_{t=1}^{N} X_{t} X_{t-2}}{\sum_{t=1}^{N} X_{t-1} X_{t-2}} \tag{2.4.3}
\end{equation*}
$$

The corresponding test statistic proposed by Hall (1989) is given by

$$
\begin{equation*}
\hat{\tau}_{N}=\left(\hat{\phi}_{1}^{(I V)}-1\right) \sqrt{\frac{\sum_{t=1}^{N} X_{t-1} X_{t-2}}{\hat{\sigma}^{2}}}, \quad \text { where } \quad \hat{\sigma}^{2}=\left[(1+\hat{\theta}) \hat{\sigma}_{s}\right]^{2} \tag{2.4.4}
\end{equation*}
$$

has the ADF $t$-dimensional, and hence the usual ADF critical values are applicable. The null hypothesis of a unit root is rejected if $\hat{\tau}_{I V}$ is less than its corresponding critical value.

### 2.5 The Generalized-Least-Squares (GLS) Unit Root Test

Let's on a series $\left\{X_{t}: t=1,2,3, \ldots, N\right\}$ assuming the representation

$$
\begin{align*}
& X_{t}=\mu+u_{t},  \tag{2.5.1}\\
& u_{t}=\phi_{1} u_{t-1}+\varepsilon_{t} \tag{2.5.2}
\end{align*}
$$

where $\varepsilon_{t} \sim$ i.i.d. $N\left(0, \sigma_{\varepsilon}^{2}\right)$. Concentrating on the $t$-statistic form of the test for (2.5.1), the t -statistic for $\phi_{1}=1$ is obtained by estimating by ordinary least squares (OLS), the autoregression

$$
\begin{equation*}
\Delta X_{t}=\mu+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} . \tag{2.5.3}
\end{equation*}
$$

Then to order $N^{-1}$, this is equivalent to computing the ADF test statistic $\hat{\tau}_{d f}$ from the reparameterized autoregression

$$
\begin{equation*}
\Delta \widetilde{X}_{t}=\mu+\left(\phi_{1}-1\right) \widetilde{X}_{t-1}+\widetilde{\varepsilon}_{t} \tag{2.5.4}
\end{equation*}
$$

where $\tilde{X}_{t}=X_{t}-\mu$, and $\hat{\mu}=\sum_{t=0}^{N} X_{t} /(N+1)$ is the OLS estimator of $\mu$. Next, denote the generalised-least-squares (GLS) test statistic by $\hat{\tau}_{g l s}$. Then $\hat{\tau}_{g l s}$ is obtained simply by calculating the ADF test statistic using the autoregression in (2.5.4), replacing $\widetilde{X}_{t}$ by a de-meaned series using a psuedo-GLS estimator of the mean $\left(\hat{\mu}_{g l s}\right)$, rather than the OLS etimator, $\hat{\mu}$. Based on the testing the hypotheses

$$
\begin{array}{ll} 
& H_{0}: \phi_{1}=1, \\
\text { vs. } & H_{1}:|\phi|<1, \tag{2.5.5b}
\end{array}
$$

the $\hat{\tau}_{g l s}$ statistic is defined as the regression $t$-statistic on the coefficient of $X_{t-1}^{*}$ in the OLS autoregression

$$
\begin{equation*}
\Delta X_{t}^{*}=\left(\phi_{1}^{*}-1\right) X_{t-1}^{*}+\varepsilon_{t}^{*} \tag{2.5.6}
\end{equation*}
$$

$$
\begin{equation*}
\text { where } \quad X_{t}^{*}=X_{t}-\hat{\mu}_{g l s} . \tag{2.5.7}
\end{equation*}
$$

The corresponding test statistic becomes

$$
\begin{equation*}
\hat{\tau}_{g^{l s}}=\frac{\phi_{1}^{*}-1}{\operatorname{Se}\left(\phi_{1}^{*}-1\right)} \tag{2.5.8}
\end{equation*}
$$

where $\operatorname{Se}\left(\phi_{1}^{*}-1\right)$ is the standard deviation of $\left(\phi_{1}^{*}-1\right)$. The same critical values used in the case of the ADF and PP tests apply. $H_{0}$ is rejected if the test statistic is less than the corresponding critical value.

### 2.6 Multiple Unit Roots Test

Much as we have considered testing for the presence of a unit root in a given time series, we must also admit that not all time series processes can well be represented by any of the autoregressions

$$
\begin{align*}
& \Delta X_{t}=\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t}  \tag{2.6.1}\\
& \Delta X_{t}=C+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t}  \tag{2.6.2}\\
& \Delta X_{t}=\sum_{i=0}^{m} \alpha_{i} t^{i}+\left(\phi_{1}-1\right) X_{t-1}+\varepsilon_{t} \tag{2.6.3}
\end{align*}
$$

and their respective higher-order autoregressions

$$
\begin{align*}
& \Delta X_{t}=\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t}  \tag{2.6.4}\\
& \Delta X_{t}=C+\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t}  \tag{2.6.5}\\
& \Delta X_{t}=\sum_{i=1}^{m} \alpha_{i} t^{i}+\left(\phi_{1}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j} \Delta X_{t-j}+\varepsilon_{t} \tag{2.6.6}
\end{align*}
$$

In rare instances, one might suspect more than one unit root. For such cases, Dickey and Pantula (1987) have proposed a simple extension of the ADF methodology capable of handling multiple unit roots. This is essentially nothing but more than performing the ADF tests on successive differences of the series, $X_{t}$. For instance, if two unit roots are suspected, the appropriate autoregression to consider is any of the following:

$$
\begin{align*}
& \Delta^{2} X_{t}=\left(\phi_{1,2}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j}^{*} \Delta^{2} X_{t-j}+\varepsilon_{t}  \tag{2.6.7}\\
& \Delta^{2} X_{t}=C_{1,2}+\left(\phi_{1,2}-1\right) X_{t-1}+\sum_{j=1}^{p-1} \gamma_{j}^{*} \Delta^{2} X_{t-j}+\varepsilon_{t}^{*}  \tag{2.6.8}\\
& \Delta X_{t}=\sum_{i=1}^{m} \alpha_{i}^{*} t^{i}+\left(\phi_{1,2}-1\right) X_{t-1}+\sum_{j=1}^{m} \gamma_{j}^{*} \Delta^{2} X_{t-j}^{2}+\varepsilon_{t}^{2} \tag{2.6.9}
\end{align*}
$$

where $\sum_{i=1}^{m} \alpha_{i}^{*} t^{i}$ is a polynomial time trend of order $m$. Employing the test statistic

$$
\begin{equation*}
\hat{\tau}_{d f}=\frac{\hat{\phi}_{1,2}-1}{\operatorname{Se}\left(\hat{\phi}_{1,2}-1\right)} \tag{2.6.10}
\end{equation*}
$$

and the same critical values used in the case of the ADF and PP tests, the null hypothesis

$$
\begin{equation*}
H_{0}: \phi_{1,2}=1 \tag{2.6.11}
\end{equation*}
$$

is rejected if the test statistic is less than the corresponding critical value.

### 2.7 Joint Unit Root Test: A Multivariate Setting

Here, we outline a simple joint unit root test developed in the multivariate setting and due to Fountis and Dickey (1989). This methodology requires the examination of the eigenvalue and eigenvector. Steps involved are as follows:

Step 1: $\quad$ Fit the linear multivariate time series. That is

$$
\begin{align*}
X_{1, t}= & \phi_{1,1} X_{1, t-1}+\phi_{1,2} X_{1, t-2}+\ldots+\phi_{1, p} X_{1, t-p}+\varepsilon_{1, t} \\
X_{2, t}= & \phi_{2,1} X_{2, t-1}+\phi_{2,2} X_{2, t-2}+\ldots+\phi_{2, p} X_{2, t-p}+\varepsilon_{2, t} \\
& \cdot \\
X_{n, t}= & \phi_{n, 1} X_{n, t-1}+\phi_{n, 2} X_{n, t-2}+\ldots+\phi_{n, p} X_{n, t-p}+\varepsilon_{n, t}  \tag{2.7.1}\\
\Rightarrow \quad \mathbf{X}_{t}= & \Phi_{1} \mathbf{X}_{t-1}+\Phi_{2} \mathbf{X}_{t-2}+\ldots \ldots+\Phi_{p} \mathbf{X}_{t-\mathrm{p}}
\end{align*}
$$

Step 2: $\quad$ Obtain the largest eigenvalue, $\boldsymbol{\lambda}_{\max }$, based on the characteristic equation

$$
\begin{equation*}
\left|\lambda^{p} \mathbf{I}-\Phi_{1} \lambda^{p-1}-\Phi_{2} \lambda^{p-2}-\ldots-\Phi_{p}\right|=\mathbf{0} \tag{2.7.2}
\end{equation*}
$$

where $\mathbf{I}$ is the $p \times p$ identity matrix.

Step 3: $\quad$ Test the following hypotheses

$$
\begin{array}{ll} 
& H_{0}: \mathbf{X}_{t} \text { has a unit root, } \\
\text { vs. } & H_{1}: \mathbf{X}_{t} \text { does not have a unit root, } \tag{2.7.3b}
\end{array}
$$

based on the following test statistic

$$
\begin{equation*}
\hat{\tau}_{m f d}=N\left(\lambda_{\max }-1\right), \tag{2.7.4}
\end{equation*}
$$

where $\lambda_{\text {max }}$ is the largest eigenvalue based on Step 2.
Step 4: $\quad$ For some nominal level, $\alpha$, obtain the critical value from the usual Dickey-Fuller table. $H_{0}$ is rejected if

$$
\begin{equation*}
\left|\hat{\tau}_{m f d}\right|>\text { Critical Value. } \tag{2.7.5}
\end{equation*}
$$

### 2.8 Practical Examples

In this section, we illustrate the concept of the unit root tests using two real data sets. The illustrations will be done using three tests, namely, the ADF, PP, and IV unit root tests. The data consists of monthly exchange rates of the South African Rand to the U.S. dollar, and monthly coin and banknotes in circulation. We shall, hitherto, refer to these datasets as Series 1 and Series 2, respectively. The data are reported in the official Bulletin of the Reserve Bank of South Africa.

Series 1: Rand/Dollar Exchange Rate in cents (Jan 1986 - April 2000),
Series 2: Coin and Banknotes in Circulation (Jan 1990 - June 2000).

### 2.8.1 Application of Unit Root Tests to Series 1

Fig. 2.1 is a graphical representation of Series 1. Fig 2.2 and Fig 2.3 are, respectively, the sample autocorrelation function (ACF's) and the partial autocorrelation functions (PACF's) for Series 1.


Fig. 2.1: Series 1 (January 1986 - May 2000)

| Name of variable | $=$ | $X$ |
| :--- | :--- | ---: |
| Mean of working series | $=574.4195$ |  |
| Standard deviation | $=$ | 206.9606 |
| Number of observations | $=$ | 172 |

## Autocorrelations



Fig. 2.2: Sample ACF's for Series 1 (January 1986 - May 2000)


Fig. 2.3: Sample PACF's for Series 1 (January 1986 - May 2000)

The sample PACF's in Fig 2.3 suggest an AR process of order $p=1$, since only the first PACF falls outside the range

$$
\pm \frac{2}{\sqrt{N}}= \pm \frac{2}{\sqrt{172}}= \pm 0.1525
$$

Hence the unit root autoregression that we consider is

$$
X_{t}=c+\phi_{1} X_{t-1}+\varepsilon_{t}, \quad \varepsilon_{t} \sim \text { i.i.d. } \mathrm{N}\left(0, \sigma_{\varepsilon}^{2}\right) .
$$

From Fig 2.2, the working mean of the series is 574.4195 which its standard deviation is 206.9606. Since the mean is greater than the standard deviation the $\operatorname{AR}(1)$ unit root autoregression that we consider is

$$
X_{t}=c+\phi_{1} X_{t-1}+\varepsilon_{t} \quad \varepsilon_{t} \sim \text { i.i.d. } \mathrm{N}\left(0, \sigma_{\varepsilon}^{2}\right)
$$

For the IV test procedure, we shall use $X_{t-2}$ as an instrument for $X_{t-1}$. Table 2.1 summarizes results from the three tests. In all cases, the null hypothesis of a unit root is accepted since the test statistics are greater than the critical values.

Table 2.1: Unit Root Tests Results

| Test | Statistic | Critical Value |  |  | Comment |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $1 \%$ | $5 \%$ | $10 \%$ |  |
| ADF |  | -3.4673 | -2.8773 | -2.5751 | Accept $\mathrm{H}_{0}$ of Unit root |
| PP | 1.3293 | -3.6473 | -2.8773 | -2.5751 | Accept $\mathrm{H}_{0}$ of Unit root |
| IV | 2.9813 | -3.6473 | -2.8773 | -2.5751 | Accept $\mathrm{H}_{0}$ of Unit root |

### 2.8.2 Application of Unit Root Tests to Series 2

A graph of Series 2 is as shown in Fig. 2.4. Fig 2.5 and Fig 2.6 are, respectively, the sample autocorrelation function (ACF's) and the partial autocorrelation functions (PACF's) for Series 2.


Fig. 2.4: Series 2 (January 1990 - June 2000)

| Name of variable | $=$ |
| :--- | :--- |
| Mean of working series | $=13234.75$ |
| Standard deviation | $=4265.457$ |

Standard deviation $=4265.457$
Number of observations $=126$

## Autocorrelations



[^0]Fig. 2.5: Sample ACF's for Series 2 (January 1990 - June 2000)


Fig. 2.6: Sample PACF's for Series 2 (January 1990 - June 2000)

The sample PACF's in Fig 2.6 suggest an AR process of order $p=1$, since only the first PACF falls outside the range

$$
\pm \frac{2}{\sqrt{N}}= \pm \frac{2}{\sqrt{126}}= \pm 0.1782
$$

Hence the unit root autoregression that we consider is

$$
X_{t}=c+\phi X_{t-1}+\varepsilon_{t}, \quad \varepsilon_{t} \sim \text { i.i.d. } \mathrm{N}\left(0, \sigma_{\varepsilon}^{2}\right)
$$

From Fig 2.5, the working mean of the series is 13234.75 which its standard deviation is 4265.457. Since the mean is greater than the standard deviation the AR(1) unit root
autoregression that we consider is

$$
X_{t}=c+\phi_{1} X_{t-1}+\varepsilon_{t} \quad \varepsilon_{t} \sim \text { i.i.d. } \mathrm{N}\left(0, \sigma_{\varepsilon}^{2}\right)
$$

For the IV test procedure, we shall use $X_{t-2}$ as an instrument for $X_{t-1}$. Table 2.2 reports the results from the three tests. Here again, the null hypothesis of a unit root is accepted in all three cases, since the test statistics are greater than the critical values.

Table 2.2: Unit Root Tests Results

| Test | Statistic | Critical Value |  |  | Comment |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\mathbf{1 \%}$ | $5 \%$ | $\mathbf{1 0 \%}$ |  |
| ADF |  | -3.4835 | -2.8845 | -2.5789 | Accept $\mathrm{H}_{0}$ of Unit root |
| PP | 0.6395 | -3.4835 | -2.8845 | -2.5789 | Accept $\mathrm{H}_{0}$ of Unit root |
| IV | 2.2807 | -3.4835 | -2.8845 | -2.5789 | Accept $\mathrm{H}_{0}$ of Unit root |

### 2.9 Conclusion

In this chapter, we have discussed some unit root tests without considering rigorous derivations. We have considered some applications of three of these unit root tests. In all the three cases, even though the test statistics were different (but do not differ much), conclusions were the same.

## CHAPTER 3

## TRENDS IN MACROECONOMIC TIME SERIES

### 3.1 Introduction

It has become common practice in macroeconomic time series analysis to test for stationarity. The two commonly applied tests are the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests (these methodologies have been discussed thoroughly in Chapter 2). Non-stationarity due to a time-dependent mean and/or variance is another common feature of macroeconomic time series. A trend in a time series may be deterministic or stochastic, or both. While de-trending can remove a deterministic trend, differencing can also remove a stochastic trend. It is, however, inappropriate to de-trend a series having a stochastic trend or difference a series driven by a deterministic trend. In practice, however, the consensus view is that most macroeconomic time series are driven by a stochastic time trend. For instance, the decomposition of real Gross National Product (GNP) series into its permanent and transitory components revealed that innovations in the stochastic trend account for quite a sizeable proportion of the period-to-period movements. (Beveridge and Nelson, 1981). This brings into focus the concept of decomposition of time series into various components. With regards to modelling seasonal data, the traditional approach is to decompose the series into three components: trend, seasonal, and random components.

In Section 3.2, the two types of trends in most macroeconomic time series, namely, deterministic and stochastic trends are discussed. Section 3.3 discusses the BeveridgeNelson decomposition method and the traditional additive-multiplicative decomposition method. In Section 3.4, we focus on giving illustrations on what is entailed in this chapter using real-life data. Section 3.5 concludes.

### 3.2 Trends in Macroeconomic Time Series

A time series, $\left\{X_{t}: t=1,2, \ldots, N\right\}$, can generally be represented as consisting of three parts:

$$
\begin{equation*}
X_{t}=f(t)+\text { Seasonal }+\varepsilon_{t}, \tag{3.2.1}
\end{equation*}
$$

where $f(t)$ is a function of time and $\varepsilon_{t}$, the disturbance term.

### 3.2.1 Deterministic Trends Models

A non-stationary time series that appears to be dependent on the time origin is said to be deterministic. The models for such data therefore need to include functions that depend on the time origin. In macroeconomics, many of such non-stationary time series can be modelled as polynomial, exponential, or sinusoidal functions, dependent on the time origin, to represent the mean of the series. A representation of such a series is

$$
\begin{equation*}
X_{t}=f(t)+\varepsilon_{t} \tag{3.2.2}
\end{equation*}
$$

where $f(t)$ is a function of time and $\varepsilon_{t}$, the disturbance term. For instance, if the series is driven by a linear time trend, then the deterministic trend model becomes

$$
\begin{equation*}
X_{t}=(a+b t)+\varepsilon_{t} \tag{3.2.3}
\end{equation*}
$$

where $a$ and $b$ are constants. The key feature of a time trend is that it has a permanent effect on a time series and hence trending elements will remain in long-term forecasts. In practice, series driven by deterministic trends are usually de-trended in order to induce stationarity. Differencing becomes inappropriate in this case. Consider, for instance, the process:

$$
\begin{equation*}
X_{t}=\alpha_{0}+\alpha_{1} t+\varepsilon_{t} \tag{3.2.4}
\end{equation*}
$$

where $\alpha_{0}$ and $\alpha_{1}$ are constants and $\varepsilon_{t}$, the disturbance term.

Then at time $t-1$, the modified equation is

$$
\begin{equation*}
X_{t-1}=\alpha_{0}+\alpha_{1}(t-1)+\varepsilon_{t-1} . \tag{3.2.5}
\end{equation*}
$$

Subtracting (3.2.5) from (3.2.4) and simplifying further yields

$$
\begin{equation*}
\Delta X_{t}=\alpha_{1}+\Delta \varepsilon_{t} \tag{3.2.6}
\end{equation*}
$$

where $\Delta X_{t}=X_{t}-X_{t-1}$ and $\Delta \varepsilon_{t}=\varepsilon_{t}-\varepsilon_{t-1}$. Equation (3.2.6) is the differenced equation. It is seen that $\Delta X_{t}$ is non-invertible since it cannot be expressed in the form of an autoregressive (AR) process. De-trending can be achieved by regressing $X_{t}$ on the appropriate polynomial time trend. The determination of the appropriate degree of the polynomial can be done by using the Akaike Information Criterion (AIC) or the Schwartz Bayesian Criterion (SBC). On the other hand, if the series is modelled as sinusoidal functions, then the appropriate representation is

$$
\begin{equation*}
X_{t}=\mu+\sum_{k=}^{m}\left\{\alpha_{k} \cdot \cos \left[w_{k}(t-1)\right]+\beta_{k} \cdot \sin \left[w_{k}(t-1)\right]\right\}+\varepsilon_{t} \tag{3.2.7}
\end{equation*}
$$

The fitted model becomes

$$
\begin{equation*}
\hat{X}_{t}=\hat{\mu}+\sum_{k=}^{m}\left\{\hat{\alpha}_{k} \cdot \cos \left[w_{k}(t-1)\right]+\hat{\beta}_{k} \cdot \sin \left[w_{k}(t-1)\right]\right. \tag{3.2.8}
\end{equation*}
$$

where

$$
\begin{gathered}
w_{k}=2 \pi k / N \text { is the } k \text {-th frequency, } \\
\hat{\alpha}_{k}=\frac{2}{N} \sum_{k=1}^{m} X_{t} \cdot \cos \left[w_{k}(t-1)\right], \quad k=1,2, \ldots, m \\
\hat{\beta}_{k}=\frac{2}{N} \sum_{k=1}^{m} X_{t} \cdot \sin \left[w_{k}(t-1)\right], \\
\hat{\mu}=\bar{X}=\sum_{t=1}^{N} X_{t} / N
\end{gathered}
$$

### 3.2.2 Stochastic Trend Models

Consider the general autoregressive integrated moving average $\operatorname{ARIMA}(p, d, q)$ process

$$
\begin{equation*}
\left(1-\phi_{1} B-\phi_{2} B^{2}-\ldots-\phi_{p} B^{p}\right)(1-B)^{d} X_{t}=\left(1-\theta_{1} B-\theta_{2} B^{2}-\ldots-\theta_{q} B^{q}\right) \varepsilon_{t} \text {, } \tag{3.2.9}
\end{equation*}
$$

where $d>0$ and $B$, the back-shift operator defined by

$$
\begin{equation*}
B^{k} X_{t}=X_{t-k}, \quad k=0,1,2,3, \ldots \tag{3.2.10}
\end{equation*}
$$

Then a stochastic trend in a given time series arises when the roots $z=\left(z_{1}, z_{2}, \ldots, z_{p}\right)$ of the characteristic equation

$$
\begin{equation*}
\left|1-\phi_{1} z-\phi_{2} z^{2}-\ldots-\phi_{p} z^{p}\right|=0 \tag{3.2.11}
\end{equation*}
$$

are real. For such a series, taking a suitable difference of the original series can induce stationarity. That is to say, the series $\left\{X_{t}: t=1,2,3, \ldots, N\right\}$ is non-stationary but its d-th differenced series, $(1-B)^{d} X_{t}$, is stationary for some integer $d \geq 1$. Consider, for instance, the random walk process given by

$$
\begin{align*}
& (1-B) X_{t}=\varepsilon_{t} \\
\Rightarrow \quad & X_{t}=X_{t-1}+\varepsilon_{t}, \quad \text { where } \varepsilon_{t} \sim \text { i.i.d. } N\left(0, \sigma_{\varepsilon}^{2}\right) \text {. } \tag{3.2.12}
\end{align*}
$$

Thus, given the past information $X_{t-1}, X_{t-2}, X_{t-3}, \ldots$, the level of the series at any time $t$ is subject to the stochastic disturbance at time $t-1$. Hence, the process is described as having a stochastic trend. In their findings, Nelson and Plosser (1982) suggest that many macroeconomic time series have a stochastic trend as well as a disturbance term. The question we ask is "having observed a series, but not the individual components, is there a way of decomposing the series into the constituent parts? We address this issue in the next section.

### 3.3 Decomposition of Time Series

Numerous macroeconomic theories emphasise the importance to distinguish between permanent and transitory movements in a macroeconomic time series. For instance, the modern theories of the consumption function that classify a person's income into permanent and transitory components highlight the importance of such decomposition. While it is quite easy handling the case of deterministic trends, it is rather a difficult issue if the trend is stochastic.

### 3.3.1 The Beveridge-Nelson Decomposition Method

Consider the first-differenced $\operatorname{ARIMA}(p, 1, q)$ process having the finite-order moving average representation

$$
\begin{equation*}
X_{t}-X_{t-1}=b_{0}+\varepsilon_{t}+\theta_{1} \varepsilon_{t-1}+\theta_{2} \varepsilon_{t-2}+\ldots \tag{3.3.1}
\end{equation*}
$$

Let

$$
\begin{equation*}
a_{t}=\varepsilon_{t}+\theta_{1} \varepsilon_{t-1}+\theta_{2} \varepsilon_{t-2}+\ldots \tag{3.3.2}
\end{equation*}
$$

Then, updating by $s$ periods, we obtain

$$
\begin{equation*}
X_{t+s}=X_{t}+b_{0} s+\sum_{i=1}^{s} a_{t+i} \tag{3.3.3}
\end{equation*}
$$

We note that

$$
\sum_{i=1}^{s} a_{t+i}=\sum_{i=1}^{s}\left(\varepsilon_{t+i}+\theta_{1} \varepsilon_{t+i-1}+\theta_{2} \varepsilon_{t+i-2}+\ldots\right)
$$

or

$$
\begin{equation*}
\sum_{i=1}^{s} a_{t+i}=\sum_{i=1}^{s} \varepsilon_{t+i}+\theta_{1} \sum_{i=1}^{s} \varepsilon_{t+i-1}+\theta_{2} \sum_{i=1}^{s} \varepsilon_{t+i-2}+\ldots \tag{3.3.4}
\end{equation*}
$$

Now, since $E\left(\varepsilon_{t+i}\right)=0$, it follows immediately then the forecast function becomes

$$
E\left(X_{t+s}\right)=E\left(X_{t}+b_{0} s+\sum_{i=1}^{s} a_{t+i}\right)
$$

which simplifies to give

$$
\begin{equation*}
\Rightarrow \quad E\left(X_{t+s}\right)=X_{t}+b_{0} s+\left(\sum_{i=1}^{s} \theta_{i}\right) \varepsilon_{t}+\left(\sum_{i=2}^{s+1} \theta_{i}\right) \varepsilon_{t-1}+\left(\sum_{i=3}^{s+2} \theta_{i}\right) \mathcal{E}_{t-2}+\ldots \tag{3.3.5}
\end{equation*}
$$

The stochastic trend is obtained by taking the limiting value of the forecast expression

$$
\begin{equation*}
E\left(X_{t+s}-b_{0} s\right)=X_{t}+\left(\sum_{i=1}^{s} \theta_{i}\right) \varepsilon_{t}+\left(\sum_{t=2}^{s+1} \theta_{i}\right) \varepsilon_{t-1}+\left(\sum_{t=3}^{s+2} \theta_{i}\right) \varepsilon_{t-2}+\ldots \tag{3.3.6}
\end{equation*}
$$

as $s \rightarrow \infty$. That is

$$
\begin{equation*}
\text { Stochastic Trend }=\lim _{s \rightarrow \infty}\left[E\left(X_{t+s}-b_{0} s\right)\right]=X_{t}+\left(\sum_{i=1}^{\infty} \theta_{i}\right) \varepsilon_{t}+\left(\sum_{i=2}^{\infty} \theta_{i}\right) \varepsilon_{t-1}+\left(\sum_{i=3}^{\infty} \theta_{i}\right) \varepsilon_{t-2}+\ldots \tag{3.3.7}
\end{equation*}
$$

To operationalized the decomposition, we have employed the fact that

$$
X_{t+s}=X_{t+s}-X_{t+s-1}+X_{t+s-1}-X_{t+s-2}+X_{t+s-2}-\ldots-X_{t+1}+X_{t+1}-X_{t}+X_{t}
$$

so that

$$
\begin{align*}
& X_{t+s}=\left(X_{t+s}-X_{t+s-1}\right)+\left(X_{t+s-1}-X_{t+s-2}\right)+\ldots+\left(X_{t+2}-X_{t+1}\right)+\left(X_{t+1}-X_{t}\right)+X_{t} \\
& \Rightarrow \quad X_{t+s}=\Delta X_{t+s}+\Delta X_{t+s-1}+\ldots+\Delta X_{t+2}+\Delta X_{t+1}+X_{t} \tag{3.3.8}
\end{align*}
$$

Hence

$$
\begin{equation*}
\lim _{s \rightarrow \infty} E\left(X_{t+s}\right)=\lim _{s \rightarrow \infty} E\left(\Delta X_{t+s}+\Delta X_{t+s-1}+\ldots+\Delta X_{t+2}+\Delta X_{t+1}\right)+X_{t+s} \tag{3.3.9}
\end{equation*}
$$

With a given time series data, we can employ the Box-Jenkins method to calculate each value of $E\left(X_{t+s}\right)$ in (3.3.9). All that we do is that for each observation in the data, we find all $s$-step ahead forecast and construct the sum given by (3.3.9).

Now, since the random component is $X_{t}$ minus the sum of the deterministic and stochastic trends, we have

$$
\begin{equation*}
\text { Random Component }=X_{t}-\lim _{s \rightarrow \infty}\left[E\left(X_{t+s}\right)-b_{0} s\right]=-\lim _{s \rightarrow \infty} E\left(\Delta X_{t+s}+\Delta X_{t+s-1}+\ldots+\Delta X_{t+2}+\Delta X_{t+1}\right)-s b_{0} . \tag{3.3.10}
\end{equation*}
$$

Thus, the steps involved in the Beveridge-Nelson decomposition are:

Step 1: Obtain the first-difference series $\Delta X_{t}=X_{t}-X_{t-1}$.

Step 2: Fit the best ARMA process of the first-differenced series, $\Delta X_{t}$.

Step 3: Using the best-fitting ARMA process, for each time point $t=1,2, \ldots, N$, obtain the 1 -step, 2 -step, $\ldots, s$-step ahead forecasts. That is, for each value of $t$ and $s$, obtain $E\left(X_{t+s}\right)$.

Step 4: For each $t$, and by setting $s$ to a certain value (say $s=100$ ), use the forecasts to construct the sums

$$
E\left(\Delta X_{t+s}+\Delta X_{t+s-1}+\Delta X_{t+s-2}+\ldots+\Delta X_{t+2}+\Delta X_{t+1}\right)+X_{t}
$$

Step 5: At every time period $t$, form the random component by subtracting the stochastic portion of the trend from the value of $X_{t}$. This mean that, for each observation, and at every time period $t$, the random component is

$$
-E\left(\Delta X_{t+s}+\Delta X_{t+s-1}+\Delta X_{t+s-2}+\ldots+\Delta X_{t+2}+\Delta X_{t+1}\right)
$$

Enders (1995) recommends a small value for $s$ if the ARMA process estimated in Step 1 has fast decaying autoregressive components, for instance $s=2$ for the $\operatorname{ARIMA}(0,1,2)$ process. For slowly decaying autoregressive components, Enders (1995) recommends a large value for $s$.

### 3.3.2 Decomposition of Seasonal Time Series

Time series data that contain seasonal components are quite common in practice, especially in macroeconomics and natural sciences. In macroeconomic time series, the seasonal pattern is frequently stochastic and changes with time. It is of the view that the stochastic nature of seasonality are due to the fact that the series are influenced by many economic factors that do not repeat themselves exactly every season.

The traditional way of modelling seasonal time series data is to decompose the series into three components, viz., a trend, a seasonal component, and an error term. The two traditional decomposition methods are the additive and multiplicative decomposition methods. We refer to the representation given in (3.2.1)

$$
X_{t}=f(t)+\text { Seasonal }+\varepsilon_{t}
$$

as the additive decomposition method. The multiplicative decomposition method specifies the following representation

$$
\begin{equation*}
X_{t}=f(t) \times \text { Seasonal } \times \varepsilon_{t} . \tag{3.3.11}
\end{equation*}
$$

This is readily transformed into an additive model as

$$
\begin{equation*}
\ln X_{t}=\ln f(t)+\ln (\text { Seasonal })+\ln \varepsilon_{t} . \tag{3.3.12}
\end{equation*}
$$

In other instances, it becomes appropriate to model the seasonal component as a linear combination of trigonometric functions. In this characterization

$$
\begin{equation*}
\text { Seasonal }=\sum_{j=1}^{m}\left[\alpha_{j} \cdot \cos \left(\frac{2 \pi j}{l} t\right)+\beta_{j} \cdot \sin \left(\frac{2 \pi j}{l} t\right)\right], \tag{3.3.13}
\end{equation*}
$$

where $m=l / 2$, and $l$ is the seasonal length.

For such a series, the appropriate representation becomes

$$
\begin{equation*}
X_{t}=f(t)+\sum_{j=1}^{m}\left[\alpha_{j} \cdot \cos \left(\frac{2 \pi j}{s} t\right)+\beta_{j} \cdot \sin \left(\frac{2 \pi j}{s} t\right)\right]+\varepsilon_{t} \tag{3.3.14}
\end{equation*}
$$

In the absence of any clear trend, the preferred representation is

$$
\begin{equation*}
X_{t}=C+\sum_{j=1}^{m}\left[\alpha_{j} \cdot \cos \left(\frac{2 \pi j}{s} t\right)+\beta_{j} \cdot \sin \left(\frac{2 \pi j}{s} t\right)\right]+\varepsilon_{t} \tag{3.3.15}
\end{equation*}
$$

where C is a constant term. This means, in essence, that modelling a non-stationary (seasonal) time series is equivalent to decomposing the series into sums of trigonometric functions.

### 3.3.3 Decomposition of Series into Deterministic and Stochastic Parts

Lastly, we consider some cases where the series may appear to be scattered around a time trend $f(t)$, where $f(t)$ may be linear or quadratic. When the data are independent, then the method usually employed consists of decomposing the series into two parts deterministic and stochastic. The method essentially removes the linear(quadratic) deterministic part that causes the non-stationarity in the series, and modelling the remaining stochastic part using the Box-Jenkins methodology.

### 3.4 Practical Examples

In this section, we illustrate the concepts discussed so far in this chapter with two real-life data sets, Series 3 and Series 4. Series 3 consists of monthly data on Consumer Price Index (CPI) for South Africa from January 1994 to Jan 2000. The base year is 1995=100. Series 4 is made up of monthly rand/pound exchange rate in cents. It stretches from January 1986 to July 2000.

### 3.4.1 Example 1: Decomposition of Series 3

Fig. 3.1 below is a graphical representation of Series 3. The plot of the series shows a strong linear trend, suggesting the presence of a linear deterministic trend.


Fig. 3.1: Series 3 (January 1995 - January 2000)

The data points seem to have long drifts indicating a high positive correlation. This means that the series appear to be dependent. The removal of the linear deterministic trend involves fitting the model

$$
\begin{equation*}
X_{t}=\alpha_{0}+\alpha_{1} t+\varepsilon_{t} \tag{3.4.1}
\end{equation*}
$$

where $\alpha_{0}$ and $\alpha_{1}$ are constants and $\varepsilon_{t} \sim$ i.i.d. $N\left(0, \sigma_{\varepsilon}^{2}\right)$. Table 3.1 reports the results obtained from fitting this model.

Now, since we want a precise fit and want to characterise the noise in the series, we need to include the stochastic part. Analysing the residuals using the Box-Jenkins methodology does this. The residuals $\hat{\varepsilon}_{t}$ from the model (3.4.1), after using the estimated values, are pictorially given in Fig. 3.2. Table 3.2 contains the unit root test results with the estimated residuals as the variable.

Table 3.1: Regression Results for Removing the Deterministic Component from Series 3
Dependent Variable: X
CPI

Ordinary Least Squares Estimates

| SSE | 64.47944 | DFE | 71 |
| :--- | ---: | :--- | ---: |
| MSE | 0.908161 | Root MSE | 0.952975 |
| SBC | 206.6857 | AIC | 202.1048 |
| Reg Rsq | 0.9962 | Total Rsq | 0.9962 |
| Normal Test | 4.1086 | Prob>Chi-Sq | 0.1282 |
| Durbin-Watson | 0.0777 |  |  |


| Variable | DF | B Value | Std Error | t Ratio Approx Prob |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
|  |  |  |  |  |  |
| Intercept | 1 | -214.348300 | 2.4160 | -88.720 | 0.0001 |
| T | 1 | 0.023595 | 0.000174 | 135.692 | 0.0001 |



Fig. 3.2: Series 3 Residuals

Table 3.2: ADF Unit Root Test Residuals


Table 3.3: Results of Modelling Series 3 Residual (Stochastic Part)
ARMA $(p+d, q)$ Tentative Order Selection Tests for the Residual Series (5\% Significance Level)

| ESACF | $p+d$ | $q$ |
| :---: | :---: | :---: |
| 2 | 1 |  |
| 1 | 2 |  |
|  | 4 | 2 |
| 5 | 1 |  |
|  | 0 | 4 |

Maximum Likelihood Estimation
Approx.

| Parameter | Estimate | Std Error | T Ratio | Lag |
| :--- | ---: | ---: | ---: | :---: |
| MA1,1 | -0.33936 | 0.11740 | -2.89 | 1 |
| MA1,2 | -0.32123 | 0.11920 | -2.69 | 2 |
| AR1,1 | 0.93093 | 0.04914 | 18.94 | 1 |

Variance Estimate $=0.05741082$
Std Error Estimate $=0.23960556$
AIC $=4.73613677$
SBC $=11.6075151$

The ADF unit root test results in Table 3.2 reveal that the estimated residual series are stationary and hence no differencing is required. Next, we model the residual series using the Box-Jenkins methodology. Table 3.3 shows the five competing models.

Out of these competing models, the ARMA(1,2) proves to be adequate. Combining the deterministic and stochastic parts, the complete model for Series 3 becomes

$$
\begin{equation*}
X_{t}=-214.3483+0.02360 t+W_{t} \tag{3.4.2}
\end{equation*}
$$

In (3.4.2), $W_{t}$ represents the stationary stochastic part that follows an $\operatorname{AR}(1)$ process

$$
\begin{equation*}
W_{t}=0.9309 W_{t-1}+a_{t}+0.3212 a_{t-1}+0.3394 a_{t-2}, \tag{3.4.3}
\end{equation*}
$$

where $a_{t} \sim$ i.i.d. $N\left(0,0.2396^{2}\right)$

### 3.4.2 Example 2: Decomposition of Series 4

In this sub-section, we illustrate the Beveridge-Nelson decomposition method using Series 4: A plot of Series 4 is shown in Fig. 3.3. The plot of the series clearly shows that the series is non-stationary. Fig. 3.4 is a plot of the first-differenced series, $\Delta X_{t}=X_{t}-X_{t-1}$.


Fig. 3.3: Series 4 (January 1986 - July 2000)


Fig. 3.4: First-Differenced Series 4

Table 3.4:
Dependent Variable $=X \quad$ rand/pound rate

| Ordinary Least |  |  |  | Squares Estimates |
| :--- | ---: | :--- | ---: | ---: |
|  |  |  | 173 |  |
| SSE | 1085325 | DFE | 79.20577 |  |
| MSE | 6273.554 | Root MSE | 79 |  |
| SBC | 2035.164 | AIC | 2028.834 |  |
| Reg Rsq | 0.8648 | Total Rsq | 0.8648 |  |
| Normal Test | 3.2837 | Prob>Chi-Sq | 0.1936 |  |
| Durbin-Watson | 0.0566 |  |  |  |


| Variable | DF | B Value | Std Error | t Ratio Approx Prob |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
|  |  |  |  |  |  |
| Intercept | 1 | -990.445977 | 47.6665 | -20.779 | 0.0001 |
| DATE | 1 | 0.129524 | 0.00389 | 33.263 | 0.0001 |

First, we may like to ascertain the truth or falsity of the economic theory that suggests the rand/pound should have a stochastic trend rather than a deterministic one. We consider the consequences of detrending instead of differencing the series. Table 3.4 reports the results from regressing the series, $X_{t}$, on time $t$.

Table 3.5: ADF Unit Root Test on Series 4 Residuals (Detrended)


The large absolute $t$-ratios indicate that the coefficients are highly significant. The residuals from the regression form the de-trended series of the rand/pound exchange rate. Results from the ADF unit test on the de-trended series (see Table 3.5) shows that even the de-trended series is non-stationary, suggesting that de-trending the series is inappropriate. This means that using the differenced series will vastly be superior to a model of the de-trended rand/pound exchange rate.

Let's shift our attention to the natural logarithm of Series 4. For this transformed series, the competing models and their results are given in Table 3.6. Here, both the AIC and SBC criteria select the ARIMA $(0,1,4)$ process as the best model. Thus, in the BeveridgeNelson decomposition method, Step 1 and Step 2 produce the process

$$
\begin{equation*}
\Delta X_{t}=0.006+\left(1+0.357 B-0.079 B^{2}-0.203 B^{3}-0.275 B^{4}\right) \varepsilon_{t} \tag{3.4.4}
\end{equation*}
$$

Table 3.6: Bevridge-Nelson Decomposition Method (Step 1 and Step 2)

## ARMA $(p+d, q)$ Tentative Order Selection Tests (5\% Significance Level)

| ESACF | $p+d$ | $q$ |
| :---: | :---: | :---: |
|  | 2 | 1 |
| 3 | 1 |  |
|  | 1 | 4 |

Maximum Likelihood Estimation
1.

|  | Approx. |  |  |  |
| :--- | ---: | ---: | ---: | :---: |
| Parameter | Estimate | Std Error | T Ratio | Lag |
| MU | 0.0062211 | 0.0030451 | 2.04 | 0 |
| MA1,1 | -0.42230 | 0.17691 | -2.39 | 1 |
| AR1,1 | -0.03393 | 0.19506 | -0.17 | 1 |
|  |  |  |  |  |
| Constant Estimate $=0.00643217$ |  |  |  |  |
|  |  |  |  |  |
| Variance Estimate $=0.00085597$ |  |  |  |  |
| Std Error Estimate $=0.02925698$ | AIC $=-732.07656$ |  |  |  |
| EBC $=-722.59939$ |  |  |  |  |

2. 

|  | Approx. |  |  |  |
| :--- | ---: | ---: | ---: | :---: |
| Parameter | Estimate | Std Error | T Ratio | Lag |
| MU | 0.0061973 | 0.0031566 | 1.96 | 0 |
| MA1,1 | -0.79758 | 0.17929 | -4.45 | 1 |
| AR1,1 | -0.41584 | 0.20075 | -2.07 | 1 |
| AR1,2 | 0.15451 | 0.11988 | 1.29 | 2 |
|  |  |  |  |  |
| Constant Estimate $=0.0078168$ |  |  |  |  |
| Variance Estimate $=0.00085815$ | AIC $=-730.61039$ |  |  |  |
| Std Error Estimate $=0.02929427$ | SBC $=-717.97417$ |  |  |  |

3. 

|  | Approx. |  |  |  |
| :--- | ---: | ---: | ---: | :---: |
| Parameter | Estimate | Std Error | T Ratio | Lag |
| MU | 0.0065500 | 0.0017367 | 3.77 | 0 |
| MA1,1 | -0.35668 | 0.07466 | -4.78 | 1 |
| MA1,2 | 0.07877 | 0.07927 | 0.99 | 2 |
| MA1,3 | 0.20300 | 0.07749 | 2.62 | 3 |
| MA1,4 | 0.27463 | 0.07450 | 3.69 | 4 |


| Constant Estimate $=0.00654997$ |  |
| :--- | :--- |
| Variance Estimate $=0.00080349$ | AIC $=-740.81456$ |
| Std Error Estimate $=0.02834585$ | SBC $=-725.01928$ |

In this case, we have $a_{0}=0.006$, and
1-step ahead forecast: $\quad E\left(\Delta X_{t+1}\right)=0.006+\left(0.357-0.079 B-0.203 B^{2}-0.275 B^{3}\right) \varepsilon_{t}$,
2-step ahead forecast: $\quad E\left(\Delta X_{t+2}\right)=0.006+\left(-0.079-0.203 B-0.275 B^{2}\right) \varepsilon_{t}$,
3-step ahead forecast: $\quad E\left(\Delta X_{t+3}\right)=0.006+(-0.203-0.275 B) \varepsilon_{t}$,
4-step ahead forecast: $\quad E\left(\Delta X_{t+4}\right)=0.006-0.275 \varepsilon_{t}$,
and all other $s$-step ahead forecasts are 0.006 .

If we set $s=50$, then at each time point $t$, the summation

$$
\begin{equation*}
E\left(\Delta X_{t+50}+\Delta X_{t+49}+\Delta X_{t+48}+\ldots+\Delta X_{t+2}+\Delta X_{t+1}\right)+X_{t} \tag{3.4.5}
\end{equation*}
$$

is given by

$$
\begin{equation*}
50(0.006)+\left(0.357-0.079 B-0.203 B^{2}-0.275 B^{3}\right) \varepsilon_{t} \tag{3.4.6}
\end{equation*}
$$

This means that for the first usable observation in the sample, the stochastic portion and the temporary portion of the trend are, respectively, given by

$$
\begin{equation*}
X_{t}+50(0.006)+\left(0.357-0.079 B-0.203 B^{2}-0.275 B^{3}\right) \varepsilon_{t} \tag{3.4.7}
\end{equation*}
$$

and

$$
\begin{equation*}
-\left(0.357-0.079 B-0.203 B^{2}-0.275 B^{3}\right) \varepsilon_{t} \tag{3.4.8}
\end{equation*}
$$



Fig. 3.5: Decomposition of the rand/pound exchange rate.

Using the estimated residuals in Appendix B, (3.4.7), and (3.4.8), the permanent and temporary series are included in the results in Appendix B. Fig. 3.5 portray the temporary and permanent portions of the series. Fig. 3.5 clearly shows that the trend dominates the movements in the temporary portion of the series. Hence, almost all changes in the rand are temporary.

### 3.5 Conclusion

In practice, most macroeconomic time series are non-stationary. It has become so clear the importance to decompose a macroeconomic time series into its permanent and temporary components. This chapter has discussed the types of trends usually found in macroeconomic time series data. It has also dealt thoroughly with some of the decomposition methods and other important representations of such series. In our empirical example of the exchange of the South African rand to the British pound, we established that movements in the rand are essentially temporary.

## CHAPTER 4

## VOLATILITY MODELLING IN MACROECONOMIC TIME SERIES

### 4.1 Introduction

In the analysis of a time series data, the variance of the disturbance term is more often than not assumed to be constant. However, researchers engaged in forecasting macroeconomic and financial time series have revealed that their ability to forecast such variables differs considerably from one time period to another. While the forecast errors are relatively small for some time periods, other time periods have relatively large forecast errors. Such a phenomenon is due to volatility in such macroeconomic variables. The variability in the variance change in the disturbances from period to period may suggest that the variance of the forecast errors is not constant but rather varies from period to period - some kind of autocorrelation in the variance of the forecast errors. In essence, we can argue that the behaviour of the forecast errors is dependent on the behaviour of the disturbances in the series. We can therefore make a case for autocorrelation in the variance of the disturbance term. To handle such a problem, Engle (1982) developed the Autoregressive Conditional Heteroskedasticity (ARCH) process.

In recent years, volatility modelling has become a very active area of research. In this chapter, we present most of the available methods for modelling volatility. Section 4.2 briefly discusses the Autoregressive Moving Average (ARMA) processes and their limitations for modelling volatility. Section 4.3 considers the ARCH process, while Section 4.4 discusses some of its extensions. The extensions to be discussed include ARCH-in-Mean (ARCH-Mean), the Generalised ARCH (GARCH) process, the Integrated GARCH (IGARCH) process, and Exponential GARCH (EGARCH) process. Section 4.4 considers other formulations of the ARCH process, namely, Threshold ARCH (TARCH) process, Component ARCH (CARCH) process, and Asymmetric CARCH (ACARCH) process. A practical illustration is given in Section 4.5. We use monthly data on percentage dividend yield on financial stock traded on the Johannesburg Stock Exchange (JSE). Section 4.6 concludes.

### 4.2 The Autoregressive Moving Average (ARMA) Process

The general representation of the autoregressive moving average process, $\operatorname{ARMA}(p, q)$, is given by

$$
\begin{equation*}
X_{t}=C+\sum_{i=1}^{p} \phi_{i} X_{t-i}+\varepsilon_{t}-\sum_{j=1}^{q} \theta_{j} \varepsilon_{t-j}, \quad \varepsilon_{t} \sim \text { i.i.d. } N\left(0, \sigma_{\varepsilon}^{2}\right) . \tag{4.2.1}
\end{equation*}
$$

In his paper, Tong (1990) outlined several advantages and limitations of the ARMA $(p, q)$ process. For instance, the availability of several statistical packages in this modern era has made the modelling of time series data with ARMA structure very simple. Again, in the literature, the ARMA process is frequently used to model macroeconomic and financial time series data. Poterba and Summers (1986) employed the AR(1) process to model volatility of Standard and Poor's S\&P 500 composite index.

A few mentions have been made with regards to the advantage of the ARMA process in modelling macroeconomic and financial time series. However, one important shortcoming of the ARMA process is the homoskedasticity assumption. This assumption does not allow changes in volatility in macroeconomic and financial time series data to be captured.

### 4.3 Autoregressive Conditional Heteroskedasticity (ARCH) Processes

The underlying property of the Autoregressive Conditional Heteroskedasticity (ARCH) process is its ability to capture the tendency for volatility in macroeconomic and financial time series. In a dynamic linear regression model, the series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ takes the form

$$
\begin{equation*}
X_{t}=Y_{t}^{\prime} \beta+\varepsilon_{t}, \tag{4.3.1}
\end{equation*}
$$

where $\varepsilon_{t}=\sigma_{t} w_{t}, \quad w_{t} \sim$ i.i.d. $(0,1) . Y_{t}^{\prime}$ is an $m \times 1$ vector of independent variables, which may be lagged values of the dependent variable, $X_{t}$, and $\beta$ is an $m \times 1$ vector of regression parameters.

In the basic ARCH process, the square of the disturbance term, $\varepsilon_{t}$, is described as itself following an $\mathrm{AR}(q)$ process:

$$
\begin{gather*}
\varepsilon_{t}^{2}=\lambda_{0}+\sum_{h=1}^{q} \lambda_{h} \varepsilon_{t-h}^{2}+v_{t}, \\
\text { or } \varepsilon_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{r} \varepsilon_{t-q}^{2}+v_{t},
\end{gather*}
$$

where $\nu_{t} \sim$ i.i.d. $\left(0, \delta^{2}\right)$. The conditions $\lambda_{0}>0$ and $\lambda_{i} \geq 0$ for $i=1,2, \ldots, q$ ensure that the conditional variance is always positive. In (4.3.2), the distribution of $\varepsilon_{t}$ conditional $\xi_{t-1}$ is

$$
\begin{equation*}
\varepsilon_{t} \mid \xi_{t-1} \sim N\left(0, \sigma_{t}^{2}\right) \tag{4.3.3}
\end{equation*}
$$

$$
\text { where } \quad \xi_{t-1}=X_{t-1}, Y_{t-1}, X_{t-2}, Y_{t-2}, \ldots
$$

### 4.3.1 Estimation of the ARCH Processes

In a more convenient way, the ARCH process is represented as

$$
\begin{gather*}
\sigma_{t}^{2}=\lambda_{0}+\sum_{k=1}^{q} \lambda_{k} \varepsilon_{t-k}^{2} \\
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{q} \varepsilon_{t-q}^{2}, \tag{4.3.5}
\end{gather*}
$$

or
where

$$
\begin{equation*}
\varepsilon_{t}=\sigma_{t} w_{t}, \quad w_{t} \sim \text { i.i.d. }(0,1) \tag{4.3.6}
\end{equation*}
$$

If $\sigma_{t}^{2}$ evolves according to (4.3.5), then

$$
\begin{equation*}
E\left(\varepsilon_{t}^{2} \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\right)=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{q} \varepsilon_{t-q}^{2} \tag{4.3.7}
\end{equation*}
$$

and hence

$$
\begin{equation*}
\varepsilon_{t} \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots \sim N\left(0, \sigma_{\varepsilon}^{2}\right) \tag{4.3.8}
\end{equation*}
$$

Now, squaring (4.3.6) yields

$$
\begin{equation*}
\varepsilon_{t}^{2}=\sigma_{t}^{2} w_{t}^{2} . \tag{4.3.9}
\end{equation*}
$$

Then, by substituting (4.3.9) and (4.3.5) in (4.3.2) and simplifying yields

$$
\begin{array}{ll} 
& \sigma_{t}^{2} \cdot w_{t}^{2}=\sigma_{t}^{2}+v_{t} \\
\text { or } & v_{t}=\sigma_{t}^{2}\left(w_{t}^{2}-1\right), \\
\Rightarrow & v_{t}^{2}=\left(\sigma_{t}^{2}\right)\left(w_{t}^{2}-1\right)^{2} . \tag{4.3.11}
\end{array}
$$

The expectation of (4.3.11) is


$$
\begin{equation*}
E\left(v_{t}^{2}\right)=E\left(\sigma_{t}^{2}\right) \times E\left[\left(w_{t}^{2}-1\right)^{2}\right] . \tag{4.3.12}
\end{equation*}
$$

Equation (4.3.12) implies that the second moment (or the variance) of $v_{t}$ does not exist for all stationary ARCH processes. For the simple case where the series $X_{t}$ assumes the AR(1) representation

$$
\begin{equation*}
X_{t}=C+\phi_{1} X_{t-1}+\varepsilon_{t}, \quad \varepsilon_{t} \sim \text { i.i.d. } N\left(0, \sigma_{\varepsilon}^{2}\right) . \tag{4.3.13}
\end{equation*}
$$

Then, we have

$$
\begin{align*}
& \varepsilon_{t} \tag{4.3.14}
\end{align*}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+v_{t}, ~ 子 ~ a n d ~ \quad \sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2} .
$$

Squaring both sides of (4.3.15) yields

$$
\begin{equation*}
\left[\sigma_{t}^{2}\right]^{2}=\left[\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}\right]^{2}=\lambda_{0}^{2}+2 \lambda_{0} \lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{1}^{2} \varepsilon_{t-1}^{4} \tag{4.3.16}
\end{equation*}
$$

Hence

$$
\begin{equation*}
E\left[\left(\sigma_{t}^{2}\right)^{2}\right]=\lambda_{0}^{2}+2 \lambda_{0} \lambda_{1} E\left[\varepsilon_{t-1}^{2}\right]+\lambda_{1}^{2} E\left[\varepsilon_{t-1}^{4}\right] \tag{4.3.17}
\end{equation*}
$$

Now,

$$
\begin{align*}
& \operatorname{var}\left[\varepsilon_{t-1}^{2}\right]=E\left[\varepsilon_{i-1}^{4}\right]-\left[E\left[\varepsilon_{t-1}^{2}\right]^{2}\right] \\
\Rightarrow & E\left[\varepsilon_{t-1}^{4}\right]=\left[E\left[\varepsilon_{t-1}^{2}\right]^{2}\right]+\operatorname{var}\left[\varepsilon_{t-1}^{2}\right] . \tag{4.3.18}
\end{align*}
$$

Thus, (4.3.17) becomes

$$
\begin{equation*}
E\left[\left(\sigma_{t}^{2}\right)^{2}\right]=\lambda_{0}^{2}+2 \lambda_{0} \lambda_{1} E\left[\varepsilon_{t-1}^{2}\right]+\lambda_{1}^{2}\left\{\operatorname{var}\left[\varepsilon_{t-1}^{2}\right]+\left[E\left[\varepsilon_{t-1}^{2}\right]\right]^{2}\right. \tag{4.3.19}
\end{equation*}
$$

By (4.3.14), since $E\left[\varepsilon_{t}\right]=E\left[\varepsilon_{t-1}\right]$, we have

$$
\begin{array}{r}
E\left[\varepsilon_{t}^{2}\right]=\lambda_{0}+\lambda_{1} E\left[\varepsilon_{t-1}^{2}\right]+E\left[\nu_{t}\right] \\
E\left[\varepsilon_{t}^{2}\right]=\lambda_{0}+\lambda_{1} E\left[\varepsilon_{t}^{2}\right]+0 \quad \text { or } \quad E\left(\varepsilon_{t}^{2}\right)=\frac{\lambda_{0}}{1-\lambda_{1}} . \tag{4.3.20}
\end{array}
$$

Similarly, we have

$$
\operatorname{var}\left[\varepsilon_{t}^{2}\right]=0+\lambda_{1}^{2} \operatorname{var}\left[\varepsilon_{t-1}^{2}\right]+\operatorname{var}\left[v_{t}\right]
$$

$$
\begin{equation*}
\text { or } \quad\left(1-\lambda_{1}^{2}\right) \cdot \operatorname{var}\left[\varepsilon_{t}^{2}\right]=\operatorname{var}\left[v_{t}\right] \tag{4.3.21}
\end{equation*}
$$

since $\operatorname{var}\left(\varepsilon_{t}^{2}\right)=\operatorname{var}\left(\varepsilon_{t-1}^{2}\right)$.

Since $v_{t} \sim$ i.i.d. $\left(0, \delta^{2}\right),(4.3 .21)$ simplifies to give

$$
\begin{equation*}
\operatorname{var}\left[\varepsilon_{t}^{2}\right]=\frac{\operatorname{var}\left[v_{t}\right]}{1-\lambda_{1}^{2}}=\frac{\delta^{2}}{1-\lambda_{1}^{2}} \tag{4.3.22}
\end{equation*}
$$

Substituting (4.3.20) and (4.3.22) in (4.3.19) and simplifying further yields

$$
\begin{gather*}
E\left[\left(\sigma_{t}^{2}\right)^{2}\right]=\lambda_{0}^{2}+2 \lambda_{0} \lambda_{1}\left[\frac{\lambda_{0}}{1-\lambda_{1}}\right]+\lambda_{1}^{2}\left\{\frac{\delta^{2}}{1-\lambda_{1}^{2}}+\left(\frac{\lambda_{0}}{1-\lambda_{1}}\right)^{2}\right\} \\
\Rightarrow \quad E\left[\left(\sigma_{t}^{2}\right)^{2}\right]=\frac{\lambda_{1}^{2} \delta^{2}}{1-\lambda_{1}^{2}}+\frac{\lambda_{0}^{2}}{\left(1-\lambda_{1}\right)^{2}} \tag{4.3.23}
\end{gather*}
$$

Also, by (4.3.12) we have

$$
\begin{equation*}
\delta^{2}=\left\{\frac{\lambda_{1}^{2} \delta^{2}}{1-\lambda_{1}^{2}}+\frac{\lambda_{0}^{2}}{\left(1-\lambda_{1}\right)^{2}}\right\} \cdot E\left(w_{t}-1\right)^{2} . \tag{4.3.24}
\end{equation*}
$$

Now, since $w_{t} \sim$ i.i.d $(0,1)$, implies

$$
\begin{equation*}
E\left[\left(w_{t}-1\right)^{2}\right]=E\left(w_{t}^{2}-2 w_{t}+1\right)=E\left(w_{t}^{2}\right)-2 \cdot E\left(w_{t}\right)+1=1-2(0)+1=2 . \tag{4.3.25}
\end{equation*}
$$

Hence, (4.3.24) becomes

$$
\begin{equation*}
\delta^{2}=2\left\{\frac{\lambda_{1}^{2} \delta^{2}}{1-\lambda_{1}^{2}}+\frac{\lambda_{0}^{2}}{\left(1-\lambda_{1}\right)^{2}}\right\}=\frac{2 \lambda_{0}^{2}\left(1-\lambda_{1}^{2}\right)}{\left(1-3 \lambda_{1}^{2}\right)\left(1-\lambda_{1}\right)^{2}} \tag{4.3.26}
\end{equation*}
$$

Equation (4.3.26) shows that if $3 \lambda_{1}^{2}<1$, then the $4^{\text {th }}$ moment of $\varepsilon_{t}$ (or the kurtosis) is greater than 3 for positive $\lambda_{1}$, and so the ARCH process yields observations with heavier tails than those of a normal distribution. If $\lambda_{1}<1, \varepsilon_{t}$ follows a white noise process while $\varepsilon_{t}^{2}$ follows an $\operatorname{AR}(q)$ process, yielding volatility clustering (Shepard, 1996).

### 4.3.2 Testing for ARCH

We have stated that the series $X_{t}$ follows an ARCH(q) process if it satisfies the mean equation specification:

$$
\begin{equation*}
X_{t}=Y_{t}^{\prime} \beta+\varepsilon_{t} \tag{4.3.1}
\end{equation*}
$$

where $\varepsilon_{t}=\sigma_{t} w_{t}, w_{t} \sim$ i.i.d. $(0,1) . Y_{t}^{\prime}$ is an $m \times 1$ vector of independent variables, which may be lagged values of the dependent variable, $X_{t}$, and $\beta$ is an $m \times 1$ vector of regression parameters. Then

$$
\begin{equation*}
X_{t} \sim N\left(Y_{t}^{\prime} \beta, \sigma_{t}^{2}\right) \tag{4.3.27}
\end{equation*}
$$

If $\Re_{t}$ is a vector of observations obtained through date $t$, then the conditional distribution of $X_{t}$ is normal with mean $Y_{t}^{\prime} \beta$ and variance $\sigma_{t}^{2}$ (i.e. by (4.3.27)):

$$
\begin{align*}
& f\left(X_{t} \mid Y, \Re_{t}\right)=\frac{1}{\sigma_{t} \sqrt{2 \pi}} \exp \left[-\frac{\varepsilon_{t}^{2}}{2 \sigma_{t}^{2}}\right], \\
& \text { or } \quad f\left(X_{t} \mid Y, \Re_{t}\right)=\frac{1}{\sigma_{t} \sqrt{2 \pi}} \exp \left[-\frac{\left(X_{t}-Y_{t}^{\prime} \beta\right)^{2}}{2 \sigma_{t}^{2}}\right], \tag{4.3.28}
\end{align*}
$$

since $\varepsilon_{t}=X_{t}-Y_{t}^{\prime} \beta$. Denoting the parameters which index the model by $\Theta$, the conditional likelihood and the log conditional likelihood are, respectively, given by

$$
\begin{equation*}
L=\sum_{t=1}^{N} f\left(X_{t} \mid Y_{t}, \mathfrak{R}_{t} ; \Theta\right)=\left(\frac{1}{\sigma_{t} \sqrt{2 \pi}}\right)^{N} \cdot \exp \left[\sum_{t=1}^{N}-\frac{\left(X_{t}-Y_{t}^{\prime} \beta\right)^{2}}{2 \sigma_{t}^{2}}\right] \tag{4.3.29}
\end{equation*}
$$

$$
\begin{equation*}
\ln L=-\frac{N}{2} \ln (2 \pi)-\frac{1}{2} \sum_{t=1}^{N} \ln \left(\sigma_{t}^{2}\right)-\frac{1}{2 \sigma_{t}^{2}} \sum_{t=1}^{N}\left(X_{t}-Y_{t}^{\prime} \beta\right)^{2} . \tag{4.3.30}
\end{equation*}
$$

The $\log$ likelihood function (4.3.30) can then be maximised with respect to the unknown parameters $\lambda=\left(\lambda_{0}, \lambda_{1}, \lambda_{2}, \ldots, \lambda_{r}\right)^{\prime}$ and $\beta$. Consider the simplest ARCH (1) process

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+v_{t}, \quad \text { where } \quad \varepsilon_{t}=\sigma_{t} w_{t}, \quad w_{t} \sim \text { i.i.d. }(0,1) \tag{4.3.31}
\end{equation*}
$$

The log conditional likelihood is

$$
\begin{equation*}
\ln L=\ln f\left(X_{t} \mid X_{0} ; \Theta\right)=-\frac{1}{2} \sum_{t=1}^{N} \ln \left(\sigma_{t}^{2}\right)-\frac{1}{2 \sigma_{t}^{2}} \sum_{t=1}^{N} X_{t}^{2} \tag{4.3.32}
\end{equation*}
$$

where $\Theta=\left(\lambda_{0}, \lambda_{1}\right)^{\prime}$. The null hypothesis that there is no volatility clustering in the series,

$$
\begin{equation*}
H_{0}: \lambda_{1}=1 . \tag{4.3.33}
\end{equation*}
$$

turns out to be the usual analogue of the Box-Pierce Portmanteau test for the $\operatorname{AR}(1)$ process or the MA(1) process, but in squares. With no specific alternative to the test, Engle (1982) recommends a Lagrangian Multiplier (LM) test of the alternative hypothesis of $\operatorname{ARCH}(q)$ disturbances since such a test can be computed from running the auxiliary regression

$$
\begin{equation*}
\hat{\varepsilon}_{t}^{2}=\hat{\lambda}_{0}+\hat{\lambda}_{1} \hat{\varepsilon}_{t-1}^{2}+\hat{\lambda}_{2} \hat{\varepsilon}_{t-2}^{2}+\ldots+\hat{\lambda}_{q} \hat{\varepsilon}_{t-q}^{2} \tag{4.3.34}
\end{equation*}
$$

Under the null hypothesis of no volatility


$$
\begin{equation*}
H_{0}: \lambda_{1}=\lambda_{2}=\ldots=\lambda_{q}=0 \tag{4.3.35}
\end{equation*}
$$

The appropriate test statistic,

$$
\begin{equation*}
T S=N R^{2} \tag{4.3.36}
\end{equation*}
$$

where $R^{2}$ is the coefficient of determination from the auxiliary regression (4.3.34), is tested as $\chi^{2}(q)$.

The hypothesis of no serial correlation (no volatility) is rejected if test statistic is greater than the corresponding chi-square value. Alternatively, we reject the null if the probability of obtaining such a chi-square value is much less than a certain nominal value, say 0.05 .

### 4.3.3 Forecasting with an ARCH Process

In time series analysis, one important aim is to be able to model the series and also to be able to forecast. The relation (4.3.2)

$$
\varepsilon_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{q} \varepsilon_{t-q}^{2}+v_{t}
$$

where $v_{t}$ ~i.i.d. $\left(0, \delta^{2}\right)$ implies that $\varepsilon_{t}^{2}$ follows an $\operatorname{AR}(q)$ process. Thus, the unconditional variance of $\varepsilon_{t}$ is

$$
\begin{gather*}
\operatorname{var}\left(\varepsilon_{t}\right)=E\left(\varepsilon_{t}^{2}\right)=E\left(\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{q} \varepsilon_{t-q}^{2}+v_{t}\right) \\
=\lambda_{0}+\lambda_{1} E\left(\varepsilon_{t-1}^{2}\right)+\lambda_{2} E\left(\varepsilon_{t-2}^{2}\right)+\ldots+\lambda_{q} E\left(\varepsilon_{t-q}^{2}\right)+E\left(v_{t}\right) \\
\Rightarrow \quad E\left(\varepsilon_{t}^{2}\right)=\lambda_{0}+\lambda_{1} E\left(\varepsilon_{t}^{2}\right)+\lambda_{2} E\left(\varepsilon_{t}^{2}\right)+\ldots+\lambda_{q} E\left(\varepsilon_{t}^{2}\right)+0, \tag{4.3.37}
\end{gather*}
$$

since $E\left(\varepsilon_{t}^{2}\right)=E\left(\varepsilon_{t-1}^{2}\right)=E\left(\varepsilon_{t-2}^{2}\right)=\ldots=E\left(\varepsilon_{t-q}^{2}\right)$. Simplify (4.3.37) further yields

$$
\begin{equation*}
\operatorname{var}\left(\varepsilon_{t}\right)=E\left(\varepsilon_{t}^{2}\right)=\frac{\lambda_{0}}{1-\lambda_{1}-\lambda_{2}-\ldots-\lambda_{q}} \tag{4.3.38}
\end{equation*}
$$

or

$$
\begin{equation*}
\sigma^{2}=\frac{\lambda_{0}}{1-\lambda_{1}-\lambda_{2}-\ldots-\lambda_{q}} \tag{4.3.39}
\end{equation*}
$$

The $s$-period-ahead linear forecast is

$$
\begin{equation*}
\hat{\varepsilon}_{t+s \mid t}^{2}=\hat{E}\left(\varepsilon_{t+s}^{2} \mid \varepsilon_{t}^{2}, \varepsilon_{t-1}^{2}, \ldots .\right) \tag{4.3.40}
\end{equation*}
$$

From (4.3.39), we have

$$
\begin{equation*}
\lambda_{0}=\sigma^{2}-\lambda_{1} \sigma^{2}-\lambda_{2} \sigma^{2}-\ldots-\lambda_{q} \sigma^{2} \tag{4.3.41}
\end{equation*}
$$

Substituting (4.3.41) in (4.3.2) and simplifying the results gives

$$
\left(\varepsilon_{t}^{2}-\sigma^{2}\right)=\lambda_{1}\left(\varepsilon_{t-1}^{2}-\sigma^{2}\right)+\lambda_{2}\left(\varepsilon_{t-2}^{2}-\sigma^{2}\right)+\ldots+\lambda_{q}\left(\varepsilon_{t-q}^{2}-\sigma^{2}\right)+\nu_{t}
$$

and hence

$$
\begin{equation*}
\left(\hat{\varepsilon}_{t}^{2}-\sigma^{2}\right)=\lambda_{1}\left(\hat{\varepsilon}_{t-1}^{2}-\sigma^{2}\right)+\lambda_{2}\left(\hat{\varepsilon}_{t-2}^{2}-\sigma^{2}\right)+\ldots+\lambda_{q}\left(\hat{\varepsilon}_{t-q}^{2}-\sigma^{2}\right) . \tag{4.3.42}
\end{equation*}
$$

The s-period-ahead forecast can be calculated from

$$
\begin{equation*}
\left(\hat{\varepsilon}_{t+k \mid t}^{2}-\sigma^{2}\right)=\lambda_{1}\left(\hat{\varepsilon}_{t+k-y \mid t}^{2}-\sigma^{2}\right)+\lambda_{2}\left(\hat{\varepsilon}_{t+k-2 \mid t}^{2}-\sigma^{2}\right)+\ldots+\lambda_{r}\left(\hat{\varepsilon}_{t+k-q \mid t}^{2}-\sigma^{2}\right) \tag{4.3.43}
\end{equation*}
$$

for $k=1,2, \ldots, s$, with $\hat{\varepsilon}_{u \mid t}^{2}=\varepsilon_{u}^{2}$ for $u \leq t$.

### 4.4 Extensions of the ARCH Process: A Review

The ARCH concept has been extended in several ways since its introduction. The most important of these extensions is the Generalised ARCH (GARCH) process due to Bollerslev (1986). In this section, we briefly discuss some of these extensions.

### 4.4.1 The ARCH-in-Mean (ARCH-M) Process

The ARCH-Mean process due to Engle, Lilien and Robins (1987) is an extension of the basic ARCH concept to allow the mean of a series to depend on its own conditional variance.

The motivation has been derived from the fact the mean and the variance of a return are expected to move in the same direction. The process is therefore suitable to the study of the relationship between risky asset and level of volatility. Denote the mean by $\mu_{t}$, where

$$
\begin{equation*}
\mu_{t}=\beta_{0}+b . f\left(\sigma_{t}^{2}\right) \tag{4.4.1}
\end{equation*}
$$

A time series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ follows an ARCH-in-Mean process if it satisfies the mean equation

$$
\begin{equation*}
X_{t}=Y_{t}^{\prime} \beta+b \cdot f\left(\sigma_{t}^{2}\right)+\varepsilon_{t} \tag{4.4.2}
\end{equation*}
$$

where

$$
\varepsilon_{t} \mid \xi_{t-1} \sim N\left(0, \sigma_{t}^{2}\right)
$$

and

$$
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{q} \varepsilon_{t-q}^{2}
$$

$f\left(\sigma_{t}^{2}\right)$ is a function of $\sigma_{t}^{2}$, with $f\left(\lambda_{0}\right)=0$. In finance, $b . f\left(\sigma_{t}^{2}\right)$ represents the expected rate of return due to an increase in the variance of the return (i.e. the risk premium). For the simple ARCH-M process where $\varepsilon_{t} \sim A R C H(1)$

$$
\begin{equation*}
X_{t}=b \cdot f\left(\sigma_{t}^{2}\right)+\varepsilon_{t} . \tag{4.4.3}
\end{equation*}
$$

Then

$$
\begin{gather*}
f\left(\sigma_{t}^{2}\right)=\lambda_{0}+\lambda_{1} \varepsilon_{t}^{2}  \tag{4.4.4}\\
\mu_{t}=\beta_{0}+b \cdot\left[\lambda_{0}+\lambda_{1} \varepsilon_{t}^{2}\right],  \tag{4.4.5}\\
\text { and } \quad X_{t}=b\left[\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}\right]+\varepsilon_{t} \\
\text { or } \quad X_{t}=b \lambda_{0}+b \lambda_{1} \varepsilon_{t-1}^{2}+\varepsilon_{t}
\end{gather*}
$$

Then using the fact that

$$
\begin{equation*}
E\left(\varepsilon_{t}^{2}\right)=E\left(\varepsilon_{t-1}^{2}\right)=\frac{\lambda_{0}}{1-\lambda_{1}}, \tag{4.3.20}
\end{equation*}
$$

it follows immediately that

$$
\begin{equation*}
E\left(X_{t}\right)=b \cdot \lambda_{0}\left[1+\frac{\lambda_{0}}{1-\lambda_{1}}\right] . \tag{4.4.7}
\end{equation*}
$$

Equation (4.4.7) is viewed as the unconditional expected return of holding a risky asset. In a similar fashion, it can be shown that

$$
\begin{equation*}
\operatorname{var}\left(X_{t}\right)=\frac{\lambda_{0}}{1-\lambda_{1}}+\frac{2 \lambda_{0}^{2}\left(b \cdot \lambda_{1}\right)^{2}}{\left(1-\lambda_{1}\right)^{2}\left(1-3 \lambda_{1}^{2}\right)} \tag{4.4.8}
\end{equation*}
$$

In the absence of a risk premium, b. $f\left(\sigma_{t}^{2}\right)=b . \lambda_{1}=0$, and so (4.4.8) becomes

$$
\begin{equation*}
\operatorname{var}\left(X_{t}\right)=\frac{\lambda_{0}}{1-\lambda_{1}} . \tag{4.4.9}
\end{equation*}
$$

Other statistical properties of the ARCH-M process have been considered in Hong (1991). In most applications, using

$$
\begin{equation*}
f\left(\sigma_{t}^{2}\right)=\ln \left(\sigma_{t}^{2}\right) \tag{4.4.10}
\end{equation*}
$$

has been found to work better in the estimation of time-varying risk premiums (Engle et al, 1987). The use of the ARCH-M process for measuring risk has been criticised in the literature, for instance Backus, Gregory and Zin (1989) and Backus and Gregory (1993). It is argued that there does not necessary exist any relationship between risk premium and conditional variances.

### 4.4.2 The Generalised ARCH (GARCH) Process

A time series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ follows the Generalised $\operatorname{ARCH}$ or $\operatorname{GARCH}(p, q)$ process if it satisfies the mean equation specification

$$
\begin{equation*}
X_{t}=Y_{t}^{\prime} \beta+\varepsilon_{t} \tag{4.3.1}
\end{equation*}
$$

where $\varepsilon_{t}=\sigma_{t} w_{t}, w_{t} \sim$ i.i.d. $(0,1) . \quad Y_{t}^{\prime}$ is an $m \times 1$ vector of independent variables, which may be lagged values of the dependent variable, $X_{t}$, and $\beta$ is an $m \times 1$ vector of regression parameters. The specified conditional variance equation is representable as

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\sum_{i=1}^{q} \lambda_{i} \varepsilon_{t-i}^{2}+\sum_{i=1}^{p} \alpha_{i} \sigma_{t-i}^{2}, \tag{4.4.11}
\end{equation*}
$$

where

$$
\begin{gathered}
\lambda_{0}>0 \\
\lambda_{i} \geq 0 \text { for } i=1,2, \ldots, q, \\
\alpha_{i} \geq 0 \text { for } i=1,2, \ldots, p,
\end{gathered}
$$

and

$$
\varepsilon_{t}=\sigma_{t} w_{t} \text { with } w_{t} \sim \text { i.i.d. }(0,1)
$$

The disturbance term is weakly stationary if

$$
\begin{equation*}
\left(\sum_{i=1}^{q} \lambda_{i}+\sum_{i=1}^{p} \alpha_{i}\right)<1 \tag{4.4.12}
\end{equation*}
$$

Writing (4.4.11) as

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\lambda(B) \varepsilon_{t}^{2}+\alpha(B) \sigma_{t}^{2} \tag{4.4.13}
\end{equation*}
$$

where $\lambda(B)=\lambda_{1} B+\lambda_{2} B^{2}+\ldots+\lambda_{q} B^{q}, \quad \alpha(B)=\alpha_{1} B+\alpha_{2} B^{2}+\ldots+\alpha_{p} B^{p}$ and $B$, the backshift operator, (4.4.13) becomes

$$
\begin{array}{ll} 
& \sigma_{t}^{2}-\alpha(B) \sigma_{t}^{2}=\lambda_{0}+\lambda(B) \varepsilon_{t}^{2} \\
\Rightarrow \quad & {[1-\alpha(B)] \sigma_{t}^{2}=\lambda_{0}+\lambda(B) \varepsilon_{t}^{2}} \\
\text { or } \quad & \sigma_{t}^{2}=\frac{\lambda_{0}}{1-\alpha(B)}+\frac{\lambda(B)}{1-\alpha(B)} \varepsilon_{t}^{2} .
\end{array}
$$

If the roots $z=\left(z_{1}, z_{2}, \ldots, z_{p}\right)$ of $1-\alpha(B)$ lie outside the unit circle, (4.4.14) becomes

$$
\begin{equation*}
\sigma_{t}^{2}=\frac{\lambda_{0}}{1-\alpha(1)}+\frac{\lambda(B)}{1-\alpha(B)} \varepsilon_{t}^{2} \quad \text { or } \quad \sigma_{t}^{2}=\lambda_{0}^{*}+\sum_{i=1}^{\infty} h_{i} \varepsilon_{t-i}^{2} \tag{4.4.15}
\end{equation*}
$$

where $\lambda_{0}^{*}=\frac{\lambda_{0}}{1-\alpha(1)}$ and $h_{i}$ is the coefficient of $B^{i}$ in the expansion of $\frac{\alpha(B)}{1-\alpha(B)}$. Equation (4.4.15) is simply a $\operatorname{GARCH}(p, q)$ process with an infinite order ARCH process.

Nelson and Cao (1992) have shown that even though the conditions under (4.4.1) are sufficient to ensure a strictly positive conditional variance, setting

$$
\begin{equation*}
\lambda_{0}^{*}>0 \text { and } h_{i} \geq 0 \tag{4.4.16}
\end{equation*}
$$

where $i=1,2,3 \ldots, \infty$ will equally ensure a strictly positive conditional variance. Consider, for instance, the $\operatorname{GARCH}(1,2)$ process

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\alpha_{1} \sigma_{t-1}^{2} \tag{4.4.17}
\end{equation*}
$$

Nelson and Cao were able to show that the conditional variance is strictly positive if based on the following conditions:

$$
\begin{equation*}
\lambda_{0}>0, \quad \lambda_{1} \geq 0, \quad \alpha_{1} \geq 0, \text { and } \alpha_{1} \lambda_{1}+\lambda_{2} \geq 0 \tag{4.4.18}
\end{equation*}
$$

As in the case of $\operatorname{ARCH}(1)$ process, in the most commonly used $\operatorname{GARCH}(1,1)$ process,

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\alpha_{1} \sigma_{t-1}^{2}, \tag{4.4.19}
\end{equation*}
$$

Hwang and Satchell have shown in Knight and Satchell (1998) that the logarithmic likelihood function is

$$
\begin{equation*}
\ln L\left(\lambda_{0}, \lambda_{1}, \alpha_{1}\right)=-\frac{1}{2} \ln (2 \pi)-\frac{1}{2} \sum_{t=1}^{N}\left[\ln \left(\sigma_{t}^{2}\right)+\frac{X_{t}^{2}}{\sigma_{t}^{2}}\right] . \tag{4.4.20}
\end{equation*}
$$

Hwang and Satchell further showed that the s-step-ahead forecast from the $\operatorname{GARCH}(1,1)$ process is given by
$E\left(X_{t+s}^{2}\right)=\lambda_{0} \sum_{i=0}^{s-1}\left(\lambda_{1}+\alpha_{1}\right)^{i}+\left(\lambda_{1}+\alpha_{1}\right)^{s-1} \lambda_{1} \sigma_{t}^{2}+\left(\lambda_{1}+\alpha_{1}\right)^{s-1} \alpha_{1} X_{t}^{2} \quad$ for $s>1$
and

$$
\begin{equation*}
E\left(X_{t+s}^{2}\right)=\lambda_{0} \sum_{i=0}^{s-1}\left(\lambda_{1}+\alpha_{1}\right)^{i}+\left(\lambda_{1}+\alpha_{1}\right)^{s-1} \sigma_{t+1}^{2} \quad \text { for } \quad s>2 . \tag{4.4.21b}
\end{equation*}
$$

Thus, for large $s$ and $\lambda_{1}+\alpha_{1}<1$, we have

$$
\begin{equation*}
E\left(X_{t+s}^{2}\right) \cong \lambda_{0} \sum_{i=0}^{s-1}\left(\lambda_{1}+\alpha_{1}\right)^{i}=\frac{\lambda_{0}}{1-\lambda_{1}-\alpha_{1}} \quad \text { as } s \rightarrow \infty \tag{4.4.22}
\end{equation*}
$$

Lastly, from the $\operatorname{GARCH}(1,1)$ process, the condition

$$
\begin{equation*}
3 \lambda_{1}^{2}+2 \lambda_{1} \alpha_{1}+\alpha_{1}^{2}<1 \tag{4.4.23}
\end{equation*}
$$

means the $4^{\text {th }}$ moment (or the kurtosis) of $\varepsilon_{t}$ is greater than that of a normal random variable. Consequently, the GARCH process is capable of producing outliers.

One important feature of $\operatorname{GARCH}(q, p)$ processes is that the conditional variance of the disturbances of the series $X_{t}$ follows an ARMA $(r, q)$ process. That is if we let

$$
\begin{equation*}
\varepsilon_{t}^{2}=\sigma_{t}^{2}+u_{t} \tag{4.4.24}
\end{equation*}
$$

then

$$
\begin{align*}
& \varepsilon_{t}^{2}=\lambda_{0}+\sum_{i=1}^{r}\left(\lambda_{i}+\alpha_{i}\right) \varepsilon_{t-i}^{2}+u_{t}-\sum_{i=1}^{q} \alpha_{i}\left(\varepsilon_{t-i}^{2}-\sigma_{t-i}^{2}\right), \\
& \varepsilon_{t}^{2}=\lambda_{0}+\sum_{i=1}^{r}\left(\lambda_{i}+\alpha_{i}\right) \varepsilon_{t-i}^{2}+u_{t}-\sum_{i=1}^{q} \alpha_{i} u_{t-i}, \tag{4.4.25}
\end{align*}
$$

where $r=\max (q, p), \quad \lambda_{i}=0$ for $i>p, \alpha_{i}=0$ for $i>q$. We see from (4.4.25) that $\varepsilon_{t}^{2}$ has an $\operatorname{ARMA}(r, q)$ representation. Therefore, it is expected that the residuals from the fitted ARMA process follow a white noise process. The autocorrelation function of the squared residuals, $\hat{\varepsilon}_{t}^{2}$, aid in determining the order of the GARCH process. In fact, McLeod and Li (1983) suggest estimating the best-fitting ARIMA model (or regression model) and calculating the sample autocorrelation (acf) of the squared residuals, $\varepsilon_{t}^{2}$ :

$$
\begin{equation*}
\hat{\rho}_{k}(\varepsilon)=\frac{\sum_{i=k+1}^{N}\left(\hat{\varepsilon}_{t}^{2}-\hat{\sigma}^{2}\right)\left(\hat{\varepsilon}_{t-k}^{2}-\hat{\sigma}^{2}\right)}{\sum_{i=1}^{N}\left(\hat{\varepsilon}_{t}^{2}-\hat{\sigma}^{2}\right)}, \quad \text { where } \quad \hat{\sigma}^{2}=\sum_{t=1}^{N} \frac{\hat{\varepsilon}_{t}^{2}}{N} \tag{4.4.26}
\end{equation*}
$$

The Box-Pierce Portmanteau statistic

$$
\begin{equation*}
Q(\varepsilon)=N(N+2) \sum_{k=1}^{m} \frac{\hat{\gamma}_{k}(\varepsilon)}{(N-k)} \tag{4.4.27}
\end{equation*}
$$

which is asymptotically distributed as $\chi^{2}(m)$, where $m$ is the number of autocorrelations used in the test, can then be used to test for groups of significant coefficients.

Rejecting the null hypothesis

$$
\begin{equation*}
H_{0}: \hat{\varepsilon}_{t}^{2} \text { are uncorrelated, } \tag{4.4.28}
\end{equation*}
$$

is equivalent to rejecting null hypothesis of no ARCH or GARCH errors. Equivalently, the LM test proposed by Engle (1982) and discussed in Section 4.3 .3 can be used. Researches have revealed that a process greater than $\operatorname{GARCH}(1,2)$ or $\operatorname{GARCH}(2,1)$ are very uncommon.

### 4.4.3 Integrated GARCH (IGARCH) Process

A time series $X_{t}$ following a standard $\operatorname{GARCH}(1,1)$ process takes the following mean equation specification and conditional variance equation:

$$
X_{t}=Y_{t}^{\prime} \beta+\varepsilon_{t},
$$

and

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\alpha_{1} \sigma_{t-1}^{2} \quad \text { where } \quad \varepsilon_{t}=\sigma_{t} w_{t} \tag{4.4.29}
\end{equation*}
$$

Now,

$$
\varepsilon_{t}^{2}=\sigma_{t}^{2}+\alpha_{1} \varepsilon_{t-1}^{2}+\varepsilon_{t}^{2}-\sigma_{t}^{2}-\alpha_{1} \varepsilon_{t-1}^{2},
$$

or

$$
\varepsilon_{t}^{2}=\left(\sigma_{t}^{2}+\alpha_{1} \varepsilon_{t-1}^{2}\right)+\left(\varepsilon_{t}^{2}-\sigma_{t}^{2}\right)-\alpha_{1} \varepsilon_{t-1}^{2},
$$

or

$$
\begin{equation*}
\varepsilon_{t}^{2}=\left(\sigma_{t}^{2}+\alpha_{1} \varepsilon_{t-1}^{2}\right)+\sigma_{t}^{2}\left(\frac{\varepsilon_{t}^{2}}{\sigma_{t}^{2}}-1\right)-\alpha_{1} \varepsilon_{t-1}^{2} . \tag{4.4.30}
\end{equation*}
$$

Substituting the relation $\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\alpha_{1} \sigma_{t-1}^{2}$ in (4.4.30) yields

$$
\begin{aligned}
& \varepsilon_{t}^{2}=\left(\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\alpha_{1} \sigma_{t-1}^{2}+\alpha_{1} \varepsilon_{t-1}^{2}\right)+\sigma_{t}^{2}\left(\frac{\varepsilon_{t}^{2}}{\sigma_{t}^{2}}-1\right)-\alpha_{1} \varepsilon_{t-1}^{2}, \\
& \varepsilon_{t}^{2}=\lambda_{0}+\left(\lambda_{1}+\alpha_{1}\right) \varepsilon_{t-1}^{2}+\sigma_{t}^{2}\left(\frac{\varepsilon_{t}^{2}}{\sigma_{t}^{2}}-1\right)-\alpha_{1}\left(\varepsilon_{t-1}^{2}-\sigma_{t-1}^{2}\right),
\end{aligned}
$$

or $\quad \varepsilon_{t}^{2}=\lambda_{0}+\left(\lambda_{1}+\alpha_{1}\right) \varepsilon_{t-1}^{2}+\sigma_{t}^{2}\left(\frac{\varepsilon_{t}^{2}}{\sigma_{t}^{2}}-1\right)-\alpha_{1} \cdot \sigma_{t-1}^{2}\left(\frac{\varepsilon_{t-1}^{2}}{\sigma_{t-1}^{2}}-1\right)$.

Using the relation $\varepsilon_{t}=\sigma_{t} w_{t}$, we have

$$
\begin{equation*}
e_{t}=\sigma_{t}^{2}\left(w_{t}-1\right) \quad \text { and } \quad e_{t-1}=\sigma_{-1 t}^{2}\left(w_{t-1}-1\right) \tag{4.4.32}
\end{equation*}
$$

Hence, (4.4.31) becomes

$$
\begin{equation*}
\varepsilon_{t}^{2}=\lambda_{0}+\left(\lambda_{1}+\alpha_{1}\right) \varepsilon_{t-1}^{2}+e_{t}-\alpha_{1} \cdot e_{t-1} . \tag{4.4.33}
\end{equation*}
$$

Equation (4.4.33) implies that the GARCH process can be written as an ARMA process. If $\lambda_{1}+\alpha_{1}<1$, then the original series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ is covariance stationary. If $\lambda_{1}+\alpha_{1}=1$, (4.4.33) becomes

$$
\begin{gather*}
\varepsilon_{t}^{2}=\lambda_{0}+\varepsilon_{-t-1}^{2}+e_{t}-\alpha_{1} \cdot e_{t-1} \\
\Rightarrow \quad \varepsilon_{t}^{2}-\varepsilon_{t-1}^{2}=\lambda_{0}+e_{t}-\alpha_{1} \cdot e_{t-1} \quad \text { or } \quad \varepsilon_{t}^{2}-B \varepsilon_{t}^{2}=\lambda_{0}+e_{t}-\alpha_{1} \cdot e_{t-1}, \tag{4.4.34}
\end{gather*}
$$

where B is the backshift operator. Equation (4.4.34) can compactly be written as

$$
\begin{equation*}
(1-B) \varepsilon_{t}^{2}=\lambda_{0}+e_{t}-\alpha_{1} \cdot e_{t-1} . \tag{4.4.35}
\end{equation*}
$$

Equation (4.4.35) leads to an analogy with an ARIMA( $0,1,1$ ) process with an intercept in terms of defining an autocorrelation function of squared observations. Equation (4.3.35) is called Integrated GARCH or IGARCH since the squared observations are stationary in first differences, but does not follow that $\varepsilon_{t}^{2}$ will behave like an integrated process. For many empirical studies using high-frequency data, $\lambda_{1}+\alpha_{1}$ is estimated to be close to 1 , suggesting that volatility has quite persistent shocks.

That is, the null hypothesis of a unit root in variance

$$
\begin{equation*}
H_{0}: \lambda_{1}+\alpha_{1}=1 \tag{4.4.36}
\end{equation*}
$$

is mostly accept using high-frequency data. For example, French, Schwert and Stambaugh (1987), Chou (1988), Pagan and Schwert (1990) do not reject the null hypothesis of unit root in variance $\left(\lambda_{1}+\alpha_{1}+\lambda_{2}+\alpha_{2}+\ldots+\lambda_{q}+\alpha_{p}=1\right)$ when the IGARCH process was applied to different stock market data.

### 4.4.4 Exponential GARCH (EGARCH) Process

A possible limitation of the GARCH process is that the conditional variance $\sigma_{t}^{2}$ responds to positive and negative residuals $\varepsilon_{t-i}$ in the same manner, i.e. $\sigma_{t}^{2}$ may be symmetric in $\varepsilon_{t-i}$. Nelson (1991) argued that a symmetric conditional variance function may be inappropriate for modelling volatility of returns on stocks since it cannot represent the leverage effect which is negative correlation between volatility and past returns. Nelson (1991) therefore proposed concept of Exponential GARCH or EGARCH. The EGARCH process enables the conditional variance to respond to positive and negative residuals asymmetrically. A time series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ follows an $\operatorname{EGARCH}(p, q)$ process if it satisfies the following specifications:

$$
\begin{align*}
& X_{t}=Y_{t}^{\prime} \beta+\varepsilon_{t} \text { with } \varepsilon_{t}=\sigma_{t} w_{t} \text { where } w_{t} \sim N(0,1) .  \tag{4.4.37a}\\
& \ln \sigma_{t}^{2}=\lambda_{0}+\sum_{i=1}^{q} \lambda_{i} \cdot f\left(w_{t-i}\right)+\sum_{i=1}^{p} \alpha_{i} \cdot \ln \sigma_{t-i}^{2} \tag{4.4.37b}
\end{align*}
$$

where

$$
\begin{equation*}
f\left(w_{t}\right)=c_{1} w_{t}+c_{2}\left[\left|w_{t}\right|-E\left|w_{t}\right|\right], \quad c_{1} \text { and } c_{2} \text { are constants. } \tag{4.4.37c}
\end{equation*}
$$

The function, $(4.4 .37 \mathrm{c})$, is independent of with mean zero and a constant variance. Thus, (4.4.37b) is a linear ARMA process for $\ln \sigma_{t}^{2}$ with innovation $f\left(w_{t}\right)$.

### 4.4.5 Threshold ARCH (TARCH) Process

The application of the EGARCH process to represent asymmetric responses in the conditional variance to positive and negative errors has motivated to the proposal of the Threshold ARCH or the $\operatorname{TARCH}(p, q)$ process. Proposed independently by Zakoian (1991) and Glosten, Jaganathan, and Runkle (1993), the specification for the conditional variance is

$$
\begin{align*}
& \sigma_{t}^{2}=\lambda_{0}+\sum_{i=q}^{q} \lambda_{i} \varepsilon_{t-i}^{2}+c_{1} \varepsilon_{t-1}^{2} d_{t-1}+\sum_{i=1}^{p} \alpha_{i} \sigma_{t-i}^{2},  \tag{4.4.38a}\\
& \text { where } \quad d_{t}= \begin{cases}1 & , \varepsilon_{t}>0 \\
0 & , \varepsilon_{t} \leq 0 .\end{cases} \tag{4.4.38b}
\end{align*}
$$

In this specification, news has differential impacts on the conditional variance, $\sigma_{t}^{2}$. Consider the simple TARCH $(1,1)$ process

$$
\begin{align*}
& \quad \sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+c_{1} \varepsilon_{t-1}^{2} d_{t-1}+\alpha_{1} \sigma_{t-1}^{2},  \tag{4.4.39}\\
& \text { where } \quad d_{t}= \begin{cases}1 & , \varepsilon_{t}>0 \\
0 & , \varepsilon_{t} \leq 0 .\end{cases}
\end{align*}
$$

For good news, $\varepsilon_{i} \leq 0$ and so $d_{t}=0$. Hence, (4.4.39) becomes

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\alpha_{1} \sigma_{t-1}^{2} \tag{4.4.40}
\end{equation*}
$$

Similarly, for bad news, $\varepsilon_{t}>0$ and so $d_{t}=1$. The specification equation (4.4.39) is

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}+\left(\lambda_{1}+c_{1}\right) \varepsilon_{t-1}^{2}+\alpha_{1} \sigma_{t-1}^{2} . \tag{4.4.41}
\end{equation*}
$$

Equation (4.4.40) and (4.4.41) show that the impact of good news is $\lambda_{1}$, while bad news has an impact of $\lambda_{1}+c_{1}$. Leverage effects exist if $c_{1}>0$. News impact is asymmetric if $c_{1} \neq 0$.

### 4.4.6 The Component ARCH (CARCH) Process

In the simple $\operatorname{GARCH}(1,1)$ process

$$
\begin{equation*}
\sigma_{t}^{2}-\mu=\lambda_{0}+\lambda_{1}\left(\varepsilon_{t-1}^{2}-\mu\right)+\alpha_{1}\left(\sigma_{t-1}^{2}-\mu\right) \tag{4.4.42}
\end{equation*}
$$

the mean reversion to $\mu$ is constant at all times. The Component ARCH or the CARCH concept, on the contrary, allows the mean reversion to a varying level, $l_{t}$. In this simple case the specification equations are

$$
\begin{gather*}
\sigma_{t}^{2}-l_{t}=\lambda_{1}\left(\varepsilon_{t-1}^{2}-l_{t-1}\right)+\alpha_{1}\left(\sigma_{t-1}^{2}-l_{t-1}\right)  \tag{4.4.43a}\\
l_{t}=\lambda_{0}+c_{1}\left(l_{t-1}-\lambda_{0}\right)+c_{2}\left(\varepsilon_{t-1}^{2}-\sigma_{t-1}^{2}\right) \tag{4.4.43b}
\end{gather*}
$$

In (4.4.43), $\sigma_{t}$ represents the volatility, while $l_{t}$ represents the time varying long run volatility. Equation (4.4.43a) describes the transitory component, $\sigma_{t}^{2}-l_{t}$, while (4.4.43b) describes the long-run component, $l_{t}$. The constant $c_{1}$ is typically between 0.99 and 1.00. Combining (4.4.43a) and (4.4.43b) gives the restricted $\operatorname{GARCH}(2,2)$ process

$$
\begin{equation*}
\sigma_{t}^{2}=\lambda_{0}^{*}+\lambda_{1}^{*} \varepsilon_{t-1}^{2}+\lambda_{2}^{*} \varepsilon_{t-2}^{2}+\alpha_{1}^{*} \sigma_{t-1}^{2}+\alpha_{2}^{*} \sigma_{t-2}^{2} \tag{4.4.44}
\end{equation*}
$$

where $\lambda_{0}^{*}=\left(1-\lambda_{1}-\alpha_{1}\right)\left(1-c_{1}\right), \quad \lambda_{1}^{*}=\left(\lambda_{1}+\alpha_{1}\right), \quad \alpha_{1}^{*}=-\left[\lambda_{1} c_{1}+c_{2}\left(\lambda_{1}+\alpha_{1}\right)\right]$, $\sigma_{2}^{*}=-\left[\lambda_{1} c_{1}-c_{2}\left(\lambda_{1}+\alpha_{1}\right)\right]$.

### 4.4.7 Asymmetric CARCH (ACARCH) Process

The Asymmetric CARCH or the ACARCH process combines the CARCH process with the asymmetric TARCH process. A time series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ follows an $\operatorname{ACARCH}(1,1)$ process if it satisfies the following specifications:

$$
\begin{align*}
& X_{t}=Y_{t}^{\prime} \beta+\varepsilon_{t}  \tag{4.4.45a}\\
& \sigma_{t}^{2}-l_{t}=\lambda_{1}\left(\varepsilon_{t-1}^{2}-l_{t-1}\right)+c\left(\varepsilon_{t-1}^{2}-l_{t-1}\right) d_{t-1}+\alpha_{1}\left(\sigma_{t-1}^{2}-l_{t-1}\right)+a_{1} z_{1, t},  \tag{4.4.45b}\\
& l_{t}=\lambda_{0}+c_{1}\left(l_{t-1}-\lambda_{0}\right)+c_{2}\left(\varepsilon_{t-1}^{2}-\sigma_{t-1}^{2}\right)+a_{2} z_{2, t} . \tag{4.4.45c}
\end{align*}
$$

where $c, c_{1}, c_{2}, a_{1}$ and $a_{2}$ are all constants. $Y_{t}^{\prime}$ is an $m \times 1$ vector of independent variables, which may be lagged values of the dependent variable, $X_{t}$, and $\beta$ is an $m \times 1$ vector of regression parameters. In (4.4.45b) and (4.4.45c), $z_{1, t}$ and $z_{2, t}$ are exogenous variables, while $d_{t}$ is the dummy variable indicating negatives shocks. Transitory leverage effects in the conditional variance is revealed if $c>0$.

## \LIBRARY!

### 4.5 Practical Example

In this example, we illustrate the ARCH concept by estimating an ARCH process using monthly percentage dividend yield on financial shares traded on Johannesburg Stock Exchange (JSE), South Africa. The series is from January 1991 to October 2000. There are 118 observations. The series is plotted in Fig. 4.1. Plots of the sample autocorrelation (ACF) and partial autocorrelation (PACF) functions are given in Fig. 4.2 and Fig. 4.3, respectively. The plot of the sample PACF's cuts off after lag 1 , suggesting that the series can be represented by an $\mathrm{AR}(1)$ process in a unit root test. Results from the PhillipsPerron unit root test on the series given in Table 4.1 suggest that it is non-stationary. Hence, differencing the series will induce stationarity. Out of the tentative models given in Table 4.2, the AIC and SBC criteria select the ARIMA( $0,1,0$ ) process.


Fig. 4.1: Percentage Dividend Yields on Financial Shares, January 1991 - October 2000)

| Name of variable | $=$ | X. |
| :--- | :--- | ---: |
| Mean of working series | $=$ | 2.566949 |
| Standard deviation | $=$ | 0.528805 |
| Number of observations | $=$ | 118 |

Autocorrelations


Fig. 4.2: Sample ACF's (\% Dividend Yields on Financial Shares, January 1991 - October 2000)

## Partial Autocorrelations



Fig. 4.3: Sample PACF's (\% Dividend Yields on Financial Shares, January 1991 - October 2000)

Table 4.1: Phillips-Perron Unit Root Test on Percentage Dividend Yields on Financial Shares


Table 4.2: Tentative Model Selection and Parameter Estimation


The estimated process is

$$
\begin{equation*}
(1-B) X_{t}=-0.0128+\varepsilon_{t} . \tag{E4.1}
\end{equation*}
$$

From the results in Table 4.3, the probability values at all the lags, namely, 0.499, 0.767, 0.480 , and 0.536 are all greater the 0.05 level of significant. Hence, we cannot reject the null hypothesis that the residuals $\varepsilon_{\imath}$ follow a white noise process. We therefore conclude that an appropriate model has been fitted to the series. The plot of the residuals in Fig. 4.4 can verify this. A plot of original series and predicted values using (E1) in Fig. 4.5 further confirms how appropriate the process fits the series.


Fig. 4.4: Residual Plot (\% Dividend Yields on Financial Shares, January 1991 - October 2000)


Fig. 4.5: Plot of Original and Predicted Series (\% Dividend Yields on Financial Shares)

The plot of the sample ACF's and PACF's of the squared residuals from the fitted ARIMA $(0,1,0)$ preocess are shown in Fig. 4.6 and Fig. 4.7, respectively. Based on the 117 residuals, the only significant autocorreation (except that at lag 18, which is considered impractical) is found at lag 1 . That is, only the autocorrealtion at lag 1 falls outside the interval $(-2 / \sqrt{117}, 2 / \sqrt{117})=(-0.1849,0.1849)$. Hence, our ARCH test should be based on the $\mathrm{ARCH}(1)$ process. However, the fact two auotcorrelations fall outside this range suggests the presence of nonlinear dependence. An ARCH(1) process fit to the original series yields the results in Table 4.3.

| Name of variable | $=($ resid $) * * 2$ |
| :--- | ---: |
| Mean of working series | $=0.027702$ |
| Standard deviation | $=$ |
| Number of observations | 0.085352 |
|  | 117 |

Autocorrelations


Fig. 4.6: Sample ACF's of Squared Residuals (January 1991 - October 2000)


Fig. 4.7: Sample PACF's of Squared Residuals (January 1991 - October 2000)

Table 4.3: ARCH Results
Dependent Variable $=X$
\% dividend yield on financial shares

Ordinary Least Squares Estimates

| SSE | 32.9969 | DFE | 117 |
| :--- | ---: | :--- | ---: |
| MSE | 0.282025 | Root MSE | 0.53106 |
| SBC | 189.2762 | AIC | 186.5055 |
| Reg Rsq | 0.0000 | Total Rsq | 0.0000 |
| Durbin-Watson | 0.0988 |  |  |


| Variable | DF | B Value | Std Error | t Ratio Approx Prob |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
| Intercept | 1 | 2.566949 | 0.0489 | 52.507 | 0.0001 |

Stationary GARCH Estimates

| SSE | 44.78633 | OBS | 118 |
| :--- | ---: | :--- | ---: |
| MSE | 0.379545 | UVAR | 3.548903 |
| Log L | -32.621 | Total Rsq | -0.3573 |
| SBC | 79.55415 | AIC | 71.2421 |
| Normality Test | 11.9314 | Prob>Chi-Sq | 0.0026 |


| Variable | DF | B Value | Std Error | t Ratio Approx Prob |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
|  |  |  |  |  |  |
| Intercept | 1 | 2.250863 | 0.01560 | 144.095 | 0.0001 |
| ARCH0 | 1 | 0.009905 | 0.00302 | 3.277 | 0.0011 |
| ARCH1 | 1 | 0.997209 | 0.15600 | 6.393 | 0.0001 |



Fig. 4.8: Estimated ARCH volatility for the Percentage Dividend Yield on Financial Shares Data

The fitted ARCH(1) process is

$$
\begin{equation*}
\sigma_{t}^{2}=0.0099+0.9972 \varepsilon_{t-1}^{2} . \tag{E4.2}
\end{equation*}
$$

The ARCH parameter $\lambda_{1}=0.9972$ is highly significant. Since $3 \lambda_{1}=2.9916$ is slightly less than 3 implies that the kurtosis is slightly greater than 3 . This means that the unconditional distribution of the series has slightly heavier tails than the normal distribution. Since non-linear dependence and heavy-tailed unconditional distribution are characteristic of conditionally heteroskedastic data implies the series is conditionally heteroskedastic. Lastly, the fact that $\lambda_{1}=0.9972$ is close to 1 is an indication that volatility shocks are quite persistent in the series.

Fig. 4.8 gives the path of the percentage dividend yield on financial shares volatility, estimated with an ARCH process.

### 4.6 Conclusion

In this chapter, we have discussed thoroughly the concept of autoregressive conditional heteroskedasticity (ARCH) and some of its extensions discussed in the literature. We have been able to use data on the percentage dividend yield on financial shares traded on the Johannesburg Stock Exchange. Results revealed volatility clustering.

## CHAPTER 5

## BOOTSTRAPPING ARCH PROCESSES

### 5.1 Introduction

In statistical data analysis, the researcher is usually interested not only the point estimation of a parameter, but also the variation of the estimator and the confidence interval of that parameter. It is procedural, traditionally, to employ the central limit theory to obtain some statistical inferences. However, the application of this theory and hence the inferences may be invalid if the sampling distribution of the estimator is not available exactly or approximately. For instance, if the normality assumption does not hold, the application of the central limit theory may not be appropriate to construct the confidence intervals.

Resampling techniques such as the bootstrap and jacknife provide not only the standard errors and confidence intervals of the estimators, but also the distribution of any statistic. The bootstrap technique is a process of repeatedly sampling, with replacement, from data at hand. From the data of size $N, B$ bootstrap samples of size $N$ are drawn to obtain $B$ new estimates. Their distribution then forms a basis for standard errors or confidence intervals (Efron and Tibshirani, 1986). The fundamental assumption of bootstrapping is that the observed data are representative of the original data.

The remainder of the chapter is organised as follows. In Section 5.2, we discuss some basic bootstrap concepts. Section 5.3 gives a brief discussion on Jacknifing. In Section 5.4, we investigate the distribution of replicated variances of an ARCH process using bootstrap methods. Here, we simulate a simple ARCH process using the Monte Carlo simulation method. Section 5.5 concludes.

### 5.2 Basic Bootstrap Concepts

Let $\left\{x_{t}: t=1,2, \ldots, N\right\}$ be realizations of the random variable $\left\{X_{t}: t=1,2, \ldots, N\right\}$, all having the same unknown distribution function $F$.

Then the empirical distribution function (EDF) is the cumulative distribution function

$$
\begin{equation*}
F_{N}(x)=\frac{1}{N} \sum_{t=1}^{N} I\left(X_{t} \leq x\right), \tag{5.1a}
\end{equation*}
$$

where

$$
I\left(X_{t} \leq x\right)=\left\{\begin{array}{l}
1, \text { if } X_{t} \leq x  \tag{5.1b}\\
0, \text { if } X_{t}>x
\end{array}\right.
$$

Equation (5.1) means that $F$ is step function, the height being $1 / N$. By ordering the realizations $X_{t}$ by ascending size, the width of $F$ is equal to the difference between two successive values of the ordered realizations $X_{t}^{*}$. If two or more observations are identical, there may be steps that have height $k(1 / N)$, where $k$ is an integer.

### 5.2.1 Properties of EDF

If $\left\{X_{t}: t=1,2, \ldots, N\right\}$ is a random sample from a population with unknown EDF, $F$, then

$$
\begin{align*}
& P\left[I\left(X_{t} \leq x\right)=1\right]=P\left(X_{t} \leq x\right) \\
\Rightarrow \quad & P\left[I\left(X_{t} \leq x\right)=1\right]=F(x) \tag{5.2.1}
\end{align*}
$$

and hence,

$$
\begin{equation*}
\left.P\left[I\left(X_{t} \leq x\right)\right]=0\right]=1-F(x) \tag{5.2.2}
\end{equation*}
$$

We note that

$$
\begin{equation*}
I\left(X_{1} \leq x\right), \quad I\left(X_{2} \leq x\right), \ldots, \quad I\left(X_{N} \leq x\right) \tag{5.2.3}
\end{equation*}
$$

are independent Bernoulli random variables, hence $N . F_{N}$ is a Binomial random variable.

Thus,

$$
\begin{array}{ll} 
& E\left[N \cdot F_{N}(x)\right]=N \cdot F(x) \\
\text { or } & E\left[\cdot F_{N}(x)\right]=. F(x), \tag{5.2.4}
\end{array}
$$

showing that $F_{N}$ is a point-wise unbiased estimate of $F$. Furthermore,

$$
\begin{align*}
& \operatorname{var}\left[N \cdot F_{N}(x)\right]=N \cdot F(x)[1-F(x)] \\
& \operatorname{var}\left[F_{N}(x)\right]=F(x)[1-F(x)] \tag{5.2.5}
\end{align*}
$$

### 5.2.2 Drawing Bootstrap Samples

Denote the $\mathrm{j}^{\mathrm{th}}$ observation of the $\mathrm{k}^{\text {th }}$ bootstrap sample by $X_{j}^{*}(k)$, where $k=1,2, \ldots, L$. Obtaining $X_{j}^{*}(k)$ involves the following steps:
i. Generate a pseudo-random number from the uniform distribution, $U(0,1)$;
ii. Use these pseudo-random number to generate a random integer $r=1,2, \ldots, N$ with equal probability;
iii. Set $X_{j}^{*}(k)$ to $X_{r}$;
iv. Repeat this process $N$ times to obtain a complete bootstrap sample. Store this bootstrap sample.
v. Repeat the whole process for $k=1,2, \ldots, L$ bootstrap samples to obtain $L$ of such statistics required.
vi. Use these $L$ statistics to obtain whatever features of the distribution of interest.
5.2.3 Parametric Bootstrap

In parametric bootstrap, we assume that the form of the distribution of the random variable, $\left\{X_{1}: t=1,2, \ldots, N\right\}$ in known in priori.

### 5.3 Jacknife-after-Bootstrap and Confidence Intervals

Jacknife-after-bootstrap is a method for obtaining estimate of the variation of such functionals of bootstrap distribution as bias and standard error of a statistic, without necessarily performing second level of bootstrapping, when the need arises. Moreover, it provides information on the influence of each observation on the functionals.

### 5.3.1 Bootstrap and Jacknife Confidence Intervals

Denote a point estimator for $\Theta$ by $\hat{\Theta}$. Let us assume that

$$
\begin{equation*}
(\hat{\Theta}-\Theta) \sim N\left(0, \sigma_{\Theta}^{2}\right) . \tag{5.3.1}
\end{equation*}
$$

Then at some level of significance, $\alpha$, we have

$$
\begin{equation*}
P\left[-Z_{q / 2} \leq \frac{\hat{\Theta}-\Theta}{\sigma_{\Theta}} \leq Z_{q / 2}\right]=1-\alpha . \tag{5.3.2}
\end{equation*}
$$

Hence, the interval

$$
\begin{equation*}
\left(\hat{\Theta}-Z_{q / 2}, \hat{\Theta}+Z_{q / 2}\right) \tag{5.3.3}
\end{equation*}
$$

is a $(1-\alpha) 100 \%$ confidence interval for $\Theta$. In practice, however, the distribution of $(\hat{\Theta}-\Theta)$ is unknown. A fairly good approximation is to let $(\hat{\Theta}-\Theta) \sim N\left(0, \sigma_{\Theta}^{2}\right)$, where $\sigma_{\Theta}^{2}$ is unknown. An estimate of $\sigma_{\Theta}^{2}$ can be obtained using bootstrap or jacknife. In the case of bootstrap estimate, an approximate $(1-\alpha) 100 \%$ bootstrap confidence interval is

$$
\begin{equation*}
\left(\hat{\Theta}-Z_{\alpha / 2} \cdot \hat{\sigma}_{\mathrm{boot}}, \Theta+Z_{q / 2} \cdot \hat{\sigma}_{\mathrm{boot}}\right), \tag{5.3.4}
\end{equation*}
$$

where $\hat{\sigma}_{\text {boot }}$ is the bootstrap estimate of $\sigma_{\Theta}$.

Similarly, a jacknife estimate of $\sigma_{\Theta}^{2}$ will yield

$$
\begin{equation*}
\left(\hat{\Theta}-Z_{\alpha / 2} \cdot \hat{\sigma}_{\text {jack }}, \Theta+Z_{\alpha / 2} \cdot \hat{\sigma}_{\text {jack }}\right), \tag{5.3.5}
\end{equation*}
$$

where $\hat{\sigma}_{\text {jack }}$ is the jacknife estimate of $\sigma_{\ominus}$.

### 5.4 Evaluating the Bootstrapped ARCH Process

In this section, our interest is to investigate the distribution of the replicated variances from the ARCH process. Let the time series $\left\{X_{t}: t=1,2, \ldots, N\right\}$ be generated by

$$
\begin{equation*}
X_{t}=\alpha+\beta . t+\varepsilon_{t}, \tag{5.4.1}
\end{equation*}
$$

where

$$
\begin{align*}
& \varepsilon=\sigma_{t} w_{t}, \quad w_{t} \sim \text { i.i.d.N }(0,1),  \tag{5.4.2}\\
& \sigma_{t}^{2}=\lambda_{0}+\lambda_{1} \varepsilon_{t-1}^{2}+\lambda_{2} \varepsilon_{t-2}^{2}+\ldots+\lambda_{q} \varepsilon_{t-q}^{2} . \tag{5.4.3}
\end{align*}
$$

Then by the step given in Section 5.2.2, the following SAS statements in Prog. 5.1 will be used to obtain the ARCH process. The size of the data is $N=200$. Bootstrapping on the simulated series will be done using the S-plus statistical software. Since the number of replications recommended for accurate estimation of percentiles is 1000 , we will set $B$ to be 1000 .

Prog. 1: Simulating ARCH Process of Size 200
data;
do samp=1 to 200;
sig2 = ranuni(0);
$w=\operatorname{sqrt}(1) *$ rannor(0);
e = sqrt(sig2)*w;
t = _n_;
$x=10+0.5^{*} t+e ;$
put x ;
output;
keep $\times$;
end;
run;
proc print noobs;
run;

In the program, we have $\alpha=10$ and $\beta=0.5$. Using S-plus, the following results in Table 5.1 were obtained. The results in Table 5.1 shows that the $95 \%$ confidence interval for the distribution of the replicated variance has endpoints 0.3596 and 0.5718 . The histogram of the replicated variance along with a smooth density estimate is shown in Fig. 5.1. The solid line is the value of the observed parameter, while the mean of the replicates is indicated by the dotted line. The bootstrap estimate of bias of -0.003346 is the difference between these two values.

Table 5.1: Distribution of the Replicated Variances

| Number of Replications: 1000 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Summary Statistics: |  |  |  |  |  |
| Observed B |  |  |  | $\begin{array}{r} \text { Mean } \\ 0.4386 \end{array}$ | SE 0.05226 |
| Empirical Percentiles: |  |  |  |  |  |
|  | 2.5\% | 5\% | 95\% |  |  |
| var | 0.3437 | 0.3587 | 0.5286 | 0.5 |  |
|  | Confidence | Limits: | $95 \%$ |  |  |
| var | 0.3596 | 0.3714 | 0.5495 | 0.5 |  |



Fig. 5.1: Histogram of Replicated Variances


Fig. 5.2: Normal Quantile-Quantile Plot of the Replicated Variances

Fig. 5.1 shows that the distribution of the replicated variances is slightly skewed, but not normal. The normal quantile-quantile plot in Fig. 5.2 confirms the fact that the distribution of the replicated residuals deviates slightly from the normal distribution.

## Table 5.2: Jacknife-after-Bootstrap

Functional Under Consideration: mean
Functional of Bootstrap Distribution of Parameters:
Func SE.Func
Param 0.43860 .06047

| Observations with Large Influence on Functional: \$Param: |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | Param |  | Param |  | Param |
| 20 | 3.086 | 76 | 2.248 | 188 | 2.575 |
| 25 | 2.266 | 83 | $2.857^{\circ}$ |  |  |
| 38 | 2.867 | 97 | 2.277 |  |  |
| 42 | 3.528 | 120 | 2.442 |  |  |
| 59 | 4.663 | 145 | 2.972 |  |  |
| 73 | 4.101 | 173 | 2.232 |  |  |



Fig. 5.3: Influence plot of Observations on Mean of the Distribution

Next, we consider the results in Table 5.2. Here, the functional of the bootstrap distribution we consider is the standard error of the mean, and the influence measure of each observation of the mean. In all, we have 13 observations having substantial influence on the mean of the distribution of the replicated variances. A graphical representation of the influence of each observation on the mean of the distribution is shown in Fig. 5.3. In particular, observations with absolute influence greater than 2 is considered to be influential.

### 5.5 Conclusion

In this chapter, we have considered bootstrapping the ARCH process with particular reference to the distribution of the replicated variances, and the measure of influence of each observation on the mean of the distribution of the replicated variances. We also obtained a negative bias which according to Christodoulakis and Satchell (in Knight and Satcell, 1998) will result in larger discrepancy between an observed and true forecast error statistic. In conclusion, the bootstrap on the ARCH process may provide worthwhile information about the process. We recommend that it is always better to check on the normality assumption before one assumes it when applying the ARCH processes.

## CHAPTER 6

## CONCLUSIONS AND RECOMMENDED RESEARCH AREA

### 6.1 Summary and Findings

This study has focused on an aspect time series, namely, macroeconomic time series with particular emphasis on modeling trends and volatility. Chapter 1 gave an overview and the structure of the whole dissertation. In Chapter 2, we reviewed some commonly used unit root tests discussed in the literature. A particular attention was paid to the Augmented Dickey-Fuller (ADF) and Phillips-Perron unit root tests. Applications of three of the unit roots using monthly data on the exchange rate of the South African rand to the U.S. dollar and coin and banknotes in circulation gave the same conclusion.

In Chapter 3, our discussions were based on decomposing a macroeconomic time series into its permanent and temporary components. The chapter considered the types of trends usually found in macroeconomic time series data. Some decomposition methods were discussed. The application of decomposition methods using data on the exchange rate of the South African rand to the British pound revealed that movements in the rand/pound exchange rate are essentially temporary. Chapter 4 dealt with the concept of autoregressive conditional heteroskedasticity (ARCH) and most of its extensions discussed in the literature. The extensions include the Generalised ARCH (GARCH) process, ARCH-in-Mean (ARCH-M) Process, Exponential GARCH (EGARCH) process, Integrated GARCH (IGARCH) process, Threshold ARCH (TARCH) process, and Asymmetric Component ARCH (ACARCH) Process. An application of the methodology using monthly data on the percentage dividend yield on financial shares traded on the Johannesburg Stock Exchange was done. Results revealed volatility clustering.

In the final chapter, an empirical study of the distribution of the replicated variances of the bootstrapped ARCH process was conducted. A series of size 200 that follows a simple ARCH process using the SAS statistical software was simulated. The simulated series was bootstrapped 1000 times using the S-plus statistical software. In this study, our interest was to investigate the distribution of the replicated variances. The study showed
that the distribution of the replicated variances is slightly skewed, but not always normal. This is something we would not have established if we had used the traditional approach. The negative biasedness of the distribution of the replicated variances is also an indication of the larger discrepancy that may exist between an observed and true forecast error statistics. The study also revealed 13 observations having substantial influence on the mean of the distribution of the replicated variances.

### 6.2 Recommended Research Area

Further research can be conducted using the same ARCH process but with the Monte Carlo method. It should be of great interest to establish whether the distribution of replicated variances using the Monte Carlo simulation will deviate from the normality assumption.

## APPENDICES

 ..... 핑Rand/Pound Exchange Rates (in cents)
 ..... $\stackrel{0}{0}$
 ..... $\stackrel{0}{4}$
采




Coin and Banknotes in Circulation (in millions of rand)









\section*{| $\circ$ |
| :--- |
| 응 |
| 운 |}



Consumer Price Index（Seasonally adjusted，1995＝100）





득 은 元远 交

ס

$$
\begin{aligned}
& \text { が }
\end{aligned}
$$



毋ை






## Testing for a Unit Root Using the ADF, PP, and IV Methods

```
input x @@;
    date = intnx( 'month', '31dec1989'd, _n_ );
    format date monyy.;
    x1 = lag(x);
    x2 = lag(x1);
    x3 = x1*x2;
cards;
6779
7412
7972
8151}88612 8453 8381 9280 8834 
8893 9104 8645
9336 8976 9126 9522 9851 9535
9939 10022 9561 10037 10331 10016
10337 10244 10038 10429 10550 10482
10516 10721 11858 11688 11627 11359
11997 11270 12135 11759 12458 12237
12005 12241 12473 13099 12351 12989
12821 13064 13725 12977 13846 14331
13546 13992 14594 14183 14818 15111
14326 15010 14762 14815 16101 15938
15552 15943 16392 15764 15979 15535
15641 16793 15983 16804 18017 17308
17660 17939 16784 17754 17978 17114
17666 17356 16882 17912 18409 18505
18642 18985 18258 19195 18758 18545
19856 19382 19730 20917 20943 22660
20830 20486 20945 21874 20714 21726
;
proc gplot;
    plot x * date = 1 / haxis= '1jan1990'd to '1jan2001'd by year;
    symbol1 i=join;
run;
proc arima;
    identify var=x stationarity=(adf=0);
    identify var=x stationarity=(pp=0);
run;
proc model;
    parms c1;
    x = c1*x1;
    fit x / 2sls;
    instruments x2;
run;
proc means sum;
run;
```

NOTE: Same program is used for the Rand/Pound Series

## Decomposition of Series into Deterministic and Stochastic Parts

```
data a;
    /* CPI */
    input x ac;
    date =_n_;
cards;
    88.2
    96.8
    103.5 104.0
```



```
120.8
130.2 130.9 131.9 132.7 133.8
140.8
;
proc reg data=a;
    model x = date / p noprint;
    output out=b residual=r;
    run;
proc forecast out=c lead=6;
        id date;
        var x;
        where date >= 1;
run;
proc print data=c;
run;
data d;
        set b;
        _type_='r';
        output;
    run;
    proc arima data=d;
        identify var=r stationarity=(adf=0);
    run;
    proc arima data=d;
        identify var=r esacf;
    run;
    proc arima data=d;
        identify var=r(1);
        estimate q=2 noconstant method=ml;
        estimate p=1 q=2 noconstant method=ml;
        estimate p=2 q=2 noconstant method=ml;
        estimate p=3 q=2 noconstant method=ml;
        estimate p=4 q=1 noconstant method=ml;
    run;
    proc arima data=d;
        identify var=r(1);
        estimate q=2 noconstant method=ml;
        forecast lead=6;
    run;
```

$x 1=\operatorname{dif}(x)$;
date $=$ intnx ( 'month', '31dec1985'd, _n_);
format date monyy.;
$\begin{aligned} x & =\text { 'log rand/pound exrate } \\ x 1 & =\text { 'first difference } x^{\prime} ;\end{aligned}$


 $6.245586 .254396 .270126 .283976 .249206 .172036 .126156 .14695 \quad 6.155886 .108006 .139506 .194126 .199376 .191776 .21655 \quad 6.22135 \quad 6.25275 \quad 6.23578 \quad 6.21162$ $6.221196 .231866 .235516 .244116 .274016 .301706 .315096 .339926 .319606 .320446 .343146 .328946 .318306 .322606 .32726 \quad 6.356326 .361736 .365116 .36966$

 6.954086 .964856 .952026 .94465
proc gplot;
plot $(x \times 1)^{*}$ date $=1 /$ haxis= '1jan1986'd to '1jan2001'd by year; symbol1 $\mathrm{i}=\mathrm{join}$;
proc arima;
identify var=x esacf;
run;
proc arima;
$\quad$ identify var= $x(1)$;
estimate $p=1 \quad q=1$ method=ml;
$\quad$ estimate $p=2 q=1$ method=ml;
estimate $q=4$ method=ml;
run;
proc arima;
identify
identify var=x(1);
estimate $q=4$ metho
estimate $q=4$ method=ml;
forecast lead=10 out=res

* produces forecasts, residuals, limits */
print data=results;
proc
run;
quit;

|  | rate | $\ln$ (rate) |  | std | L95 | U95 |  | e(t) | e(t-1) | e(t-2) | e(t-3) | (t) |  | temp |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Jan-86 | 335.35 | 5.81517 * |  |  |  |  | * | * |  |  |  |  |  |  |
| Feb-86 | 297.62 | 5.69582 | 5.82172 | 0.031692 | 5.7596 | 5.88384 | -0.1259 | -0.1259 * |  |  |  |  |  |  |
| Mar-86 | 297 | 5.69373 | 5.66205 | 0.030023 | 5.60321 | 5.7209 | 0.03168 | 0.03168 | -0.1259 |  |  | 3 |  |  |
| Apr-86 | 306.56 | 5.72541 | 5.7258 | 0.02922 | 5.66853 | 5.78307 | -0.00039 | -0.00039 | 0.03168 | 0.1259 |  |  |  |  |
| May-86 | 333.11 | 5.80847 | 5.76119 | 0.028895 | 5.70456 | 5.81783 | 0.04728 | 0.04728 | -0.00039 | 0.03168 | -0.1259 | 5 | 5.85357 | -0.0451 |
| Jun-86 | 380.95 | 5.94267 | 5.85282 | 0.028672 | 5.79663 | 5.90902 | 0.08985 | 0.08985 | 0.04728 | -0.00039 | 0.03168 | 6 | 5.96238 | -0.01971 |
| Jul-86 | 384.76 | 5.95262 | 5.96987 | 0.028601 | 5.91381 | 6.02593 | -0.01725 | -0.01725 | 0.08985 | 0.04728 | -0.00039 | 7 | 5.92987 | 0.022747 |
| Aug-86 | 385.51 | 5.95457 | 5.93689 | 0.028383 | 5.88126 | 5.99252 | 0.01768 | 0.01768 | -0.01725 | 0.08985 | 0.04728 | 8 | 5.931 | 0.023567 |
| Sep-36 | 340.37 | 5.83003 | 5.93801 | 0.028379 | 5.88238 | 5.99363 | -0.10798 | -0.10798 | 0.01768 | -0.01725 | 0.08985 | 9 | 5.76888 | 0.061153 |
| Oct-86 | 320.62 | 5.77026 | 5.77609 | 0.028372 | 5.72048 | 5.83169 | -0.00583 | -0.00583 | -0.10798 | 0.01768 | -0.01725 | 10 | 5.77786 | -0.0076 |
| Nov-86 | 320.62 | 5.77026 | 5.78417 | 0.028372 | 5.72856 | 5.83978 | -0.01391 | -0.01391 | -0.00583 | -0.10798 | 0.01768 | 11 | 5.78281 | -0.01255 |
| Dec-86 | 319.28 | 5.76607 | 5.78933 | 0.02835 | 5.73376 | 5.84489 | -0.02326 | -0.02326 | -0.01391 | -0.00583 | -0.10798 | 12 | 5.78974 | -0.02367 |
| Jan-87 | 314.76 | 5.75181 | 5.79618 | 0.02835 | 5.74061 | 5.85174 | -0.04437 | -0.04437 | -0.02326 | -0.01391 | -0.00583 | 13 | 5.74223 | 0.009576 |
| Feb-87 | 317.46 | 5.76035 | 5.74878 | 0.028348 | 5.69322 | 5.80434 | 0.01157 | 0.01157 | -0.04437 | -0.02326 | -0.01391 | 14 | 5.77653 | -0.01618 |
| Mar-87 | 329.49 | 5.79755 | 5.78305 | 0.028348 | 5.72749 | 5.83861 | 0.0145 | 0.0145 | 0.01157 | -0.04437 | - 0.02326 | 15 | 5.81722 | -0.01967 |
| Apr-87 | 328.52 | 5.7946 | 5.82375 | 0.028346 | 5.76819 | 5.87931 | -0.02915 | -0.02915 | 0.0145 | 0.01157 | -0.04437 | 16 | 5.7929 | 0.001699 |
| May-87 | 334.78 | 5.81347 | 5.75944 | 0.028346 | 5.74389 | 5.855 | 0.01403 | 0.01403 | -0.02915 | 0.0145 | 0.01157 | 17 | 5.81466 | -0.00119 |
| Jun-87 | 329.27 | 5.79688 | 5.8212 | 0.028346 | . 76564 | 5.87676 | -0.02432 | -0.02432 | 0.01403 | -0.02915 | 0.0145 | 18 | 5.7890 | 0.007861 |
| Jul-87 | 331.57 | 5.80384 | 5.79559 | 0.028346 | 5.74003 | 5.85114 | 0.00825 | 0.00825 | -0.02432 | 0.01403 | -0.02915 | 19 | 5.81387 | -0.01003 |
| Aug-87 | 331.79 | 5.8045 | 5.82041 | 0.028346 | 5.76485 | 5.87596 | -0.01591 | -0.01591 | 0.00825 | -0.02432 | 0.01403 | 20 | 5.79925 | 0.005253 |
| Sep-87 | 336.25 | 5.81785 | 5.80581 | 0.028346 | 5.75025 | 5.86137 | 0.01204 | 0.01204 | -0.01591 | 0.00825 | -0.02432 | 21 | 5.82842 | -0.01057 |
| Oct-87 | 339.9 | 5.82865 | 5.83495 | 0.028346 | 5.77939 | 5.89051 | -0.0063 | -0.0063 | 0.01204 | -0.01591 | 0.00825 | 22 | 5.82641 | 0.002239 |
| Nov-87 | 350.39 | 5.85905 | 5.83297 | 0.028346 | 5.77741 | 5.88852 | 0.02608 | 0.02608 | -0.0063 | 0.01204 | -0.01591 | 23 | 5.87079 | -0.01174 |
| Dec-87 | 356.63 | 5.8767 | 5.87732 | 0.028346 | 5.82177 | 5.93288 | -0.00062 | -0.00062 | 0.02608 | -0.0063 | 0.01204 | 24 | 5.87239 | 0.004314 |
| Jan-88 | 356.13 | 5.8753 | 5.87895 | 0.028346 | 5.82339 | 5.9345 | -0.00365 | -0.00365 | -0.00062 | 0.02608 | -0.0063 | 25 | 5.87048 | 0.004816 |
| Feb-88 | 359.45 | 5.88458 | 5.87703 | 0.028346 | 5.82148 | 5.93259 | 0.00755 | 0.00755 | -0.00365 | -0.00062 | 0.02608 | 26 | 5.88052 | 0.004062 |
| Mar-88 | 389.71 | 5.9654 | 5.88707 | 0.028346 | 5.83152 | 5.94263 | 0.07833 | 0.07833 | 0.00755 | -0.00365 | -0.00062 | 27 | 5.99368 | -0.02828 |
| Apr-88 | 402.25 | 5.99707 | 6.0002 | 0.028346 | 5.94465 | 6.05576 | -0.00313 | -0.00313 | 0.07833 | 0.00755 | -0.00365 | 28 | 5.98924 | 0.007834 |
| May-88 | 413.05 | 6.02357 | 5.9958 | 0.028346 | 5.94024 | 6.05136 | 0.02777 | 0.02777 | -0.00313 | 0.07833 | 0.00755 | 29 | 6.01575 | 0.007816 |
| Jun-88 | 404.53 | 6.00273 | 6.0223 | 0.028346 | 5.96674 | 6.07786 | -0.01957 | -0.01957 | 0.02777 | -0.00313 | 0.07833 | 30 | 5.97264 | 0.030086 |
| Jul-88 | 407.66 | 6.01043 | 5.97924 | 0.028346 | 5.92368 | 6.03479 | 0.03119 | 0.03119 | -0.01957 | 0.02777 | -0.00313 | 31 | 6.01833 | -0.0079 |
| Aug-88 | 416.49 | 6.03186 | 6.02487 | 0.028346 | 5.96931 | 6.08043 | 0.00699 | 0.00699 | 0.03119 | -0.01957 | 0.02777 | 32 | 6.0282 | 0.003633 |


0.003272
-0.00878
0.00122
-0.00239
-0.01031
-0.00794
0.006891
-0.00638
-0.00428
0.003823
0.007734
0.000163
-0.00945
-0.00449
0.001298
$-9.1 E-05$
-0.00098
0.017187
0.02476
-0.00198
-0.03171
-0.01249
0.016449
-0.02264
-0.02746
0.002278
0.018515
-0.01289
-0.00219
-0.00505
0.019115



| 0.003682 | 0.014737 |
| ---: | ---: |
| -0.04035 | 0.003682 |
| -0.00465 | -0.04035 |
| 0.003464 | -0.00465 |
| -0.03637 | 0.003464 |
| -0.00296 | -0.03637 |
| 0.010207 | -0.00296 |
| -0.00519 | 0.010207 |
| -0.01693 | -0.00519 |
| 0.019038 | -0.01693 |
| 0.002597 | 0.019038 |
| -0.01235 | 0.002597 |
| -0.00826 | -0.01235 |
| -0.0073 | -0.00826 |
| 0.010651 | -0.0073 |
| 0.004419 | 0.010651 |
| -0.00223 | 0.004419 |
| 0.010479 | -0.00223 |
| 0.007209 | 0.010479 |
| -0.0423 | 0.007209 |
| -0.06655 | -0.0423 |
| -0.02769 | -0.06655 |
| 0.012275 | -0.02769 |
| -0.02931 | 0.012275 |
| -0.06691 | -0.02931 |
| 0.041394 | -0.06691 |
| 0.025457 | 0.041394 |
| -0.02875 | 0.025457 |
| -0.01186 | -0.02875 |
| 0.036732 | -0.01186 |
| -0.01463 | 0.036732 |



















 든 N | 0 |
| :--- |
| 0 |
| O |
| $\vdots$ | N 춘 $\infty$

$\stackrel{n}{N} \stackrel{n}{5}$
웅










$\begin{array}{lll}\varphi & \varphi & \varphi\end{array}$





Jun-9
Jul-9
Aug-9
Sep-9
Oct-9
Nov-9
Dec-9
Jan-9
Feb-9
Mar-9
Apr-9
May-9
Jun-9
Jul-9
Aug-9
Sep-9
Oct-9
Nov-9
Dec-9
Jan-9
Feb-9
Mar-98
Apr-9
May-98
Jun-9
Jul-9
Aug-9
Sep-98
Oct-9
Nov-98
Dec-9 $6.90098 \quad 0.001964$ .908530 .004779 $\begin{array}{ll}6.88771 & 0.004841\end{array}$ -0.00492
0.011407 -0.00463
-0.00963
-0.00424
-0.00649
0.010628 $-0.00349$ -0.00169
-0.00686
 88SL00'0 $n$
0
0
0
0
0
0
$\qquad$ o



|  |
| :---: |
|  <br>  <br>  io ío ó io í óo í |
|  <br>  <br>  |
| , 1010 |

6.96373
6.95948
6.96307
6.97065
6.94982
6.97263
6.93001
6.93486
6.96386
6.95667
6.98634
6.95585
6.9654
6.97606
6.98857
6.98895
7.02342
7.02537
7.00501

$$
\begin{array}{r}
6.85196 \\
6.85953 \\
6.83871 \\
6.86152 \\
6.8189 \\
6.82374 \\
6.85275 \\
6.84556 \\
6.87522 \\
6.84473 \\
6.85429 \\
6.86495 \\
6.87746 \\
6.87783 \\
6.91231 \\
6.91426 \\
6.89389
\end{array}
$$

## REFERENCES

Backus, D.K., A.W. Gregory, and S.E. Zin. (1989). " Risk Premiums in the term structure: evidence from artificial economics", Journal of Monetary Economics, 24: 371399

Backus, D.K., and A.W. Gregory. (1993). "Theoretical relations between risk and premiums and conditional variances", Journal of Business and Economic Statistics, 11(2): 177-185

Beveridge, S., and C. Nelson. (1981). "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of Business Cycle", Journal of Monetary Economics, 7: 151-174

Chou, R.Y. (1988). Volatility persistence and stock valuations: some empirical evidence using ARCH, Journal of Applied Econometrics, 4(4): 657-684

Dickey, D.A., and S.G. Pantula. (1987). "Determining the Order of Differencing in Autoregressive Processes", Journal of Business and Economic Statistics, 15: 455-461

Dickey, D.A., and W.A. Fuller. (1979). "Distribution of the estimators for autoregressive time series with unit root". Journal of the American Statistical Association, 74(336): 427431

Efron, B., and R. Tibshirani. (1986). "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy", Statistical Science, 1(1): 54-77 Enders, W. (1995). Applied Econometric Time Series. New York: John Wiley \& Sons, Inc

Engle, R.F. (1982). "Autoregressive Conditional Heteroscedasticity with estimates of the variance United Kingdom Inflation, Econometrica, 50: 286-301

Engle, R.F., D.M. Lilien, and R.P. Robins. (1987). "Estimating Time Varying Risk Premia in the term structure: the ARCH-M model, Econometrica, 55: 987-1007

French, K.R., G.W. Schwert, and R.F. Stambaugh. (1987). "Expected stock returns and volatility, Journal of Financial Econometrics, 19: 3-30

Glosten, L.R., R. Jagannathan, and D. Runkle (1993) "On the Relation between the Expected Value and the Volatility of the Normal Excess Return on Stocks," Journal of Finance, 48: 1779-1801

Hall, A. (1989) "Testing for a Unit Root in the Presence of Moving Average Errors". Biometrika, vol. 79: 49-46

Hong, P.Y. (1991). "The autocorrelation structure for the GARCH-M process", Economic Letters, 37: 129-132
McLeod, A.I., and W.K.Li. (1983). "Diagnostic checking ARMA time series models using squared-residuals autocorrelations", Journal of Time Series Analysis, 4: 269-273

Nelson, C.R., and C.I. Plosser. (1982). "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications", Journal of Monetary Economics, 4: 139162

Nelson, D.B. (1991). "Conditional heteroscedasticity in asset returns: a new approach, Econometrica, 59: 347-370

Nelson, D.B., and C.Q. Cao. (1992). "Inequality constraints in the Univariate GARCH model, Journal of Business and Economic Statistics, 10(2): 229-235

Newey, W.K., and K.D. West. (1987). "A Simple Positive Semi-Definite Heteroskedasicity and Autocorrelation-Consistent Covariance Matrix", Econometrica, 55: 703-708

Pagan, A.R., and G.W. Schwert. (1990). Alternative models for conditional stock volatility, Journal of Econometrics, 45: 267-290

Phillips, P.C.B. (1987a). "Time Series Regression with a Unit Root", Econometrica, 55: 277-301

Phillips, P.C.B., and P. Perron. (1988). "Testing for a unit root in time series regression", Biometrika, 75(2): 335-346

Poterba, J. and L. Summers. (1986). The persistenc of volatility and stock market fluctuations, $A E R, 76$ : 1141-1151

Schwert, G.W. (1989). "Tests for Unit Roots: A Monte Carlo Investigation", Journal of Business and Economic Statistics, 7: 147-159

Shepard, N. (1996). Statistical aspects of ARCH and Stochastic volatility, in D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (editors), Time Series Models in Econometrics, Finance and Other Fields. London: Chapman and Hall

Tong, H. (1990). Nonlinear Time Series: A Dynamic System Approach. Oxford University Press: New York

Zakoian, J.M. (1990) "Threshold Heteroskedastic Models," manuscript, CREST, INSEE, Paris


[^0]:    ." marks two standard errors

