

**Examining long-run relationships of the BRICS stock market indices to identify opportunities for implementation of statistical arbitrage strategies**

By

Brian Meki

Student Number: 2808333

A full thesis submitted in fulfilment of the requirements  
for the degree of Master of Science in Computational Finance  
in the Department of Statistics, University of the Western Cape



Supervisor

UNIVERSITY of the  
WESTERN CAPE

Professor Danelle Kotze

November, 2012

## Declaration

I declare that **Examining long-run relationships of the BRICS stock market indices to identify opportunities for implementation of statistical arbitrage strategies** is my own work, that it has not been copied from anywhere, and that all sources I have used have been indicated and acknowledged by complete references.

Signed : .....

Date : 31 October, 2012



## **Acknowledgements**

I would be remiss if I did not share in this final word the truth that was involved in the successful completion of this research. The mere fact that I am able to close this chapter of my personal and academic life being of sound mind, body and spirit is testimony, in and of itself, of the kindness and mercy of the divine creator.

Through the highs and the lows I faced in the course of my research, my family kept a constant stream of love, care and undeterred patience flowing my way even though we have been worlds apart. This gave me the zeal to push on, to keep my head up high and to put a smile on my face. For that, I am forever grateful and blessed to have been given unto them as my family.

On the other hand, I most likely would have not been able to survive the world of academia had it not been for the guidance, care and constant encouragement of my supervisor, Professor Danelle Kotze. For that, I am deeply indebted.

Lastly, but not least, I was fortunate to have been the recipient of unwavering, unfaltering and priceless support and companionship given by my friends and the entire Statistics and population studies department staff at the UWC. This, I will remember for all time.

## **Abstract**

### **Purpose:**

This research investigates the existence of long-term equilibrium relationships among the stock market indices of Brazil, Russia, India, China and South Africa (BRICS). It further investigates cointegrated stock pairs for possible implementation of statistical arbitrage trading techniques.

### **Design:**

We utilize standard multivariate time series analysis procedures to inspect unit roots to assess stationarity of the series. Thereafter, cointegration is tested by the Johansen and Juselius (1990) procedure and the variables are interpreted by a Vector Error Correction Model (VECM). Statistical arbitrage is investigated through the pairs trading technique.

### **Findings:**

The five stock indices are found to be cointegrated. Analysis shows that the cointegration rank among the variables is significantly influenced by structural breaks. Two pairs of stock variables are also found to be cointegrated. This guaranteed the mean reversion property necessary for the successful execution of the pairs trading technique. Determining the optimal spread threshold also proved to be highly significant with respect to the success of this trading technique.

### **Value:**

This research seeks to expand on the literature covering long-run co-movements of the volatile emerging market indices. Based on the cointegration relation shared by the BRICS, the research also seeks to encourage risk taking when investing. We achieve this by showing the potential rewards that can be realized through employing appropriate statistical arbitrage trading techniques in these markets.

**Key words:** Cointegration, Rank, Statistical arbitrage, Mean reversion, Pairs trading, Unit roots, Stationarity, Vector Error Correction Model (VECM), Structural breaks, Spread, Threshold.

## List of Acronyms

AIC	Akaike information criterion
ADF	Augmented Dickey and Fuller
AR	Autoregressive (process)
AR(p)	Autoregressive process of order p
ARMA	Autoregressive moving average (process)
ARMA(p, q)	Autoregressive moving average process of order (p, q)
ARCH	Autoregressive conditional heteroscedasticity
AZ	Andrew and Zivot
BRICS	Brazil, Russia, India, China, South Africa
BRIC	Brazil, Russia, India, China
CUSUM	Cumulative sums
DF	Dickey and Fuller
FPE	Final prediction error
GARCH	Generalized autoregressive conditional heteroscedasticity
HQC	Hanna-Quinn (criterion)
JSE	Johannesburg stock exchange
KPSS	Kwiatkowski, Phillips, Schmidt and Shin
LM	Lagrange multiplier
LR	Likelihood ratio
MA	Moving average (process)
MA(q)	Moving average process of order q
MLR	Maximum likelihood ratio
MSCI	Morgan Stanley capital international
OLS	Ordinary least squares
PP	Phillips and Perron
Rec-CUSUM	Cumulative sums of recursive errors
RTSI	Russian trading system index
SBIC	Schwarz-Bayesian information criterion
SSE	Shanghai stock exchange
VAR	Vector autoregressive
VEC	Vector error correction
VECM	Vector error correction model

## List of Figures

Figure 1	Testing sequence for unit roots: With and without structural breaks .....	34
Figure 2a	Pairs trading rules .....	60
Figure 2b	Trading signals .....	61
Figure 3	Time series plots of raw monthly prices .....	66
Figure 4	Time series plots of log-transformed monthly prices .....	68
Figure 5	Time series plots of log-differenced monthly prices .....	70
Figure 6	SENSEX and RTSI monthly stock prices .....	93
Figure 7	RTSI and BOVESPA monthly stock prices .....	94
Figure 8	Price spread between the SENSEX and RTSI stocks .....	96
Figure 9	Price spread between the RTSI and BOVESPA stocks .....	97
Figure 10a	Trading signals at 5% spread (SENSEX-RTSI) .....	99
Figure 10b	Trading signals at 5% spread (RTSI-BOVESPA) .....	100
Figure 10c	Trading signals at 1% spread (SENSEX-RTSI) .....	101
Figure 10d	Trading signals at 1% spread (RTSI-BOVESPA) .....	104
Figure 10e	Trading signals at 10% spread (SENSEX-RTSI) .....	107
Figure 10f	Trading signals at 10% spread (RTSI-BOVESPA) .....	108
Figure 11a	Performance of the SENSEX-RTSI pairs trading at 5% spread .....	109
Figure 11b	Performance of the SENSEX-RTSI pairs trading at 1% spread .....	110
Figure 11c	Performance of the RTSI-BOVESPA pairs trading at 1% spread .....	111

## List of Tables

Table 1a	Descriptive statistics for the raw monthly prices .....	67
Table 1b	Descriptive statistics for the log-transformed monthly prices .....	69
Table 1c	Descriptive statistics for the log-differenced monthly prices .....	71
Table 1d	Correlation matrix of the log-differenced monthly prices .....	72
Table 2	Unit root tests in the absence of structural breaks .....	74
Table 3	Structural break tests for each time series .....	76
Table 4	Unit root tests in the presence of structural breaks .....	77
Table 5	Lag length determination .....	79
Table 6	Rank determination ignoring the shift dummy variable .....	81
Table 7	Rank determination including the shift dummy variable .....	82
Table 8	Testing for cointegrated pairs .....	83
Table 9	Market capitalization data as of 30 December, 2007 .....	84
Table 10a	The cointegrating vector and loading parameters ( $r=1$ ) .....	85
Table 10b	The cointegrating vector and loading parameters ( $r=2$ ) .....	87
Table 11a	Model testing ( $r=1$ ) .....	90
Table 11b	Model testing ( $r=2$ ) .....	91
Table 12	Unit root tests .....	98
Table 13a	Entry position for the SENSEX-RTSI pair trade .....	102
Table 13b	Exit position for the SENSEX-RTSI pair trade .....	103
Table 13c	Profit and loss on the SENSEX-RTSI pair trade .....	103
Table 14a	Entry position for the RTSI-BOVESPA pair trade .....	104
Table 14b	Exit position for the RTSI-BOVESPA pair trade .....	104
Table 14c	Profit and loss on the RTSI-BOVESPA pair trade .....	105

## Table of Contents

Chapter 1	
Introduction .....	1
1.1 Research framework .....	1
1.2 Research question .....	4
1.3 Research value .....	5
1.4 Research outline .....	5
Chapter 2	
Literature Review .....	7
2.1 Historical overview: Prelude to cointegration .....	7
2.2 Stationarity .....	9
2.2.1 Introduction .....	9
2.2.2 Consequences of overlooking stationarity .....	9
2.2.3 Tests for stationarity .....	11
2.2.4 Completeness of unit root tests .....	12
2.3 Cointegration .....	21
2.3.1 Cointegration tests .....	23
2.4 Cointegration based trading strategies .....	26
2.4.1 Index tracking .....	26
2.4.2 Enhanced index tracking .....	26
2.4.3 Statistical arbitrage .....	26
2.5 Heteroscedasticity .....	29
2.5.1 How heteroscedasticity can occur .....	30
2.5.2 Consequences .....	30
2.5.3 Tests .....	30
2.6 Conclusion .....	31
Chapter 3	
Methodology .....	33
3.1 Unit root tests, stationarity and structural change .....	38
3.1.1 Unit root tests assuming no structural change .....	39
3.1.2 Unit root tests accounting for structural change .....	45

3.2	Cointegration.....	49
3.2.1	Granger representation theorem, VAR and VECM.....	50
3.2.2	Johansen and Juselius' cointegration procedure .....	53
3.3	Diagnostic testing.....	55
3.4	Statistical arbitrage.....	57
4	Data description .....	63
5	Data analysis and modelling .....	65
5.1	Graphical analysis.....	65
5.2	Tests for stationarity .....	73
5.3	Tests for structural change .....	75
5.4	Model specification.....	78
5.4.1	Specification of dummy variables .....	78
5.4.2	Lag length determination .....	79
5.4.3	Rank determination .....	80
5.4.4	VECM estimation .....	83
5.4.5	Impulse response functions.....	88
5.5	Model misspecification tests.....	89
5.6	Statistical arbitrage.....	91
6	Conclusion .....	112
	Bibliography .....	116
APPENDICES		
APPENDIX A:	Partial/Autocorrelations .....	134
APPENDIX B:	Differenced stock data .....	139
APPENDIX C:	Impulse response functions.....	144
APPENDIX D:	VAR(2) residual plots .....	149
APPENDIX E:	R code used for the analysis.....	151
APPENDIX F:	Statistical tables .....	164

## CHAPTER 1

### Introduction

Within the last decade, the world recorded one of the most devastating financial crises in recent history. Many emerging markets have been slow to recover from the stock market collapse. As a result, prominent researchers have argued that the emerging markets offer no internationally diversifiable investment opportunities (Balarezo, 2010). In essence, such arguments suggest that both emerging and international markets are highly correlated and pose great investment risk. International diversification is defined as holding a portfolio of assets from both developed and developing markets (Msimanga, 2010).

Although the importance of diversification is quite clear, very little is known about the emerging markets themselves aside from their illiquid nature and exceptionally high volatilities compared to developed markets (Thomas, 2006). Thus in an attempt to expand on the literature covering the developing markets, this research provides a detailed study on the movements of a basket of assets picked only from selected emerging markets. The stock markets from Brazil, Russia, India, China and South Africa (BRICS) provided an adequate platform for our study and the data used was obtained from the countries' respective major stock exchanges. To give some point of perspective, the research assumes the position of an emerging market trader considering holding a portfolio only in the BRICS stock markets.

### 1.1 Research framework

Aggarwal et al. (1999) and Suwardi (2004) attribute the highly volatile characteristic of the emerging markets mainly to taxes and transaction costs, transparency, corruption and liquidity issues. Yet, on the other hand, the very same emerging markets are well known for their fast track reform agendas, investor friendly atmospheres and attractive investment opportunities (Dawson, 2005).

Therefore, the success of any long-term trader in such “delicate” markets hinges on the understanding of the risk-return trade-off of investing across different market conditions presented by the BRICS markets. We explore such understanding by employing Engle and Granger’s (1987) cointegration technique. This approach uses historical data to determine if the different stock data show past co-movement. If the data are found to have historically moved together, then it can be assumed that they will also do so in future. Hence, the interest of this research is in long-run equilibrium relationships. Unlike basing the study on correlation, cointegration overcomes the inadequacies of solely relying on short-term returns to make investment decisions (Granger and Hallman, 1991). The success of the cointegration technique depends on the data being stationary, that is, exhibiting a constant mean and variance over time. We examine these properties by testing for unit roots within the data using the augmented Dickey-Fuller test (Dickey and Fuller, 1981) and the Kwitkowski, Phillips, Schmidt, and Shin (1992) test.

Next, we treat all five variables endogenously using Johansen and Juselius’ (1990) method. In doing so, we simultaneously test if the data are cointegrated and determine how many cointegration relationships exist between them. The method also allows for the cointegrated variables to be defined by a vector error correction model (VECM) with parameters estimated by the method of ordinary least squares (OLS). In this way, the model builds the cointegration relations “into the specification so that it confines the long-term behaviour of the endogenous variables to converge to their cointegrating relationships while allowing for short-term dynamics” (Kotzeva and Pauna, 2006, p. 17).

We hypothesize that the economic meltdown of 2007-2009 may have had significant and similar effects on the BRICS stock markets. Ignoring such effects could lead to biased unit root tests and consequently, all following analyses based on those results could be unreliable. It can be expected that the turmoil may have resulted in major policy changes at both government and economic levels within the BRICS. As a result, the markets would have experienced changes in the long-run equilibrium stock prices. Thus we investigate these changes by testing the data for structural breaks. The tests employed are the supremum F test (Zeileis et al., 2001) and the cumulative sum of recursive errors (Brown et al., 1975).

The Johansen and Juselius (1990) procedure is reapplied to the data with structural breaks taken into account. This is crucial because a structural break has the potential to influence the number of long-run relationships that are determined (Lutkepohl, 2004). The VECM is also re-estimated and further stabilized by capturing the qualitative effects of the break by a shift dummy variable.

Among many other authors, Febrian and Herwany (2007) point out that time series data often show volatility clustering, negative correlation between volatility and stock prices and distributions of their returns are often leptokurtic. It is mainly for these reasons that we check the underlying vector autoregressive (VAR) model for adequacy after estimating the VECM. This ensures that the estimates are efficient and their standard errors are not inflated. We carry out the model checks by applying the Breusch-Godfrey (1980) LM test for serial correlation, Jarque-Bera (1980) test for normality and the Engle (1982) LM test for heteroscedastic errors.

Granger (1981) observes that the existence of a cointegration relation implies that there is at least one causal relation. Thus, after model checking we investigate the causal relationships among the BRICS using impulse response functions (IRF). The IRF measure the magnitude of the effects of external shocks applied to the system. This exercise not only maps out how shock is transmitted throughout the system but also identifies which of the variables are exogenous and which are endogenous.

The VECM also has the capacity to capture the rate at which the system reverts back to its equilibrium after any temporal deviations. This property allows for investigation of asset mispricing, the very essence of the study of statistical arbitrage. Our research identifies that the statistical arbitrage opportunities presented by the BRICS can be attributed to the markets' volatile nature. This is because the flow of information that determines asset prices in the markets is impeded on its way to the traders. As a result, Fama's (1970) efficient market hypothesis (EMH) does not always hold for the BRICS.

According to Lutkepohl (2004), finding that the BRICS are cointegrated implies that we can expect at least two of the five variables to be cointegrated. Using the mean reversion property, we investigate statistical arbitrage by means of pairs trading. The pairs trading

technique offers a trader the opportunity to realize gains by taking advantage of asset mispricing while maintaining a market neutral position before the market corrects itself. Thus profits are gained when the value of the portfolio of the two cointegrated stocks is sufficiently above the equilibrium value. The portfolio is determined to be sufficiently above the equilibrium value if the spread between the two stock prices exceeds two historical standard deviations. We use the two standard deviations as a default trading rule indicating when to enter and when to exit market with each trade. However, we take the liberty of investigating how the simulated trading technique can be affected by considering price spreads of 1% and 10% levels of significance as well. The remainder of the strategy makes use of Takayanagi and Ishikawa's (2012) trading algorithm to track and analyse the portfolio's fluctuations about the boundaries set by the trading rules. Lastly, the cumulative returns from the multiple trades entered over time are discussed with respect to the performance of the portfolios.

## 1.2 Research question

The key issues that drive this research are as follows;

- i) To an emerging market trader, the recent inclusion of South Africa into the BRIC group may stir some investor curiosity into the performance of the emerging markets in the long horizon. Can an emerging markets investor hold a diversified stock portfolio in the BRICS? If not, what is the upside of holding a portfolio of stocks that have been found to historically move together?
- ii) Following Zeileis (2001) and Lutkepohl (2004), can it be said that the 2007-2009 melt down was significantly detectable particularly considering the volatilities characterizing the emerging markets? What bearing, if any, did the recession have on the unit root and cointegration tests applied to the data? Provided the BRICS are cointegrated and structural breaks are significant, how is the estimation of the VECM affected, if at all, by including a structural shift dummy variable in the model?

By attempting to answer these questions and more, we hope to illuminate on whether any relationships found to exist among the BRICS are economically and statistically legitimate or simply spurious.

### **1.3 Research value**

There is a plethora of literature on emerging markets. However, only a few authors have considered the BRICS, more so since South Africa's recent inclusion as the fifth member of the group. Accordingly, this research seeks to add to the literature covering the emerging markets, and consequently, the BRICS economies by drawing insight from such scholars as Dubois and Bacmann (2001), Alagidede and Panagiotidis (2009), Khan (2010) and many others.

Hence, the defining characteristics of this research are its attempt to gradually develop guided, concise and informative steps involved when investigating long-term relationships that involve the challenging aspect of structural shifts. In this way, the research intends to improve on the understanding of the emerging markets, the implementation of the cointegration technique and the use of the mean reversion property in statistical arbitrage.

However, the findings of this research are only specific to the data and sample period defined and investigated in this research.

### **1.4 Research outline**

The rest of the research is structured as follows: academic literature and relevant empirical reviews are summarized in Chapter 2. Throughout this chapter we track how time series analysis has evolved over time leading up to the general acceptance of the cointegration technique by many researchers and scholars as an acclaimed analysis tool. We compare the cointegration technique to other long-run techniques, investigate its relation to the mean reversion property and its role in the implementation of statistical arbitrage trading. In Chapter 3, we outline the methodology of how data is analysed in the remainder of the research. An attempt is made to list, with explained reasons, relevant formulae and techniques

sourced from various texts. A short description of the data is given in Chapter 4. This includes description summaries of the individual timelines of the five BRICS markets with a particular focus on the 2007-2009 time period. The analysis of the data is carried out in Chapter 5. By following the methodology in Chapter 3, we test and investigate the data for unit roots, structural breaks, cointegration and statistical arbitrage. Finally, a conclusion to the investigation is given in Chapter 6. Here we discuss the findings of our research to various aspects found in similar areas of research and texts, but mainly with respect to the research agenda outlined in Chapter 1.



## CHAPTER 2

### Literature review

The intention of this study is to identify and implement possible statistical arbitrage techniques from a long-term meaningful relationship between the Brazilian BOVESPA Stock Index, the Russian Trading System Index, the Indian Bombay SENSEX Stock Exchange, the Chinese Shanghai Stock Exchange and the South African Johannesburg Stock Exchange All Share.

Any long-run meaningful relationship(s) modeled on the variables representing this data have to be statistically and economically valid. We mention this on account of the data being just realizations of some underlying stochastic process (Pfaff, 2008). As will be further discussed later, applying traditional methods to model such kind of data may lead to unreliable model estimates where the variables are actually related or may result in suggested relations where none exist (Yule, 1926). Thus we investigate any possible long-run relationship(s) that the five variables may share using the cointegration technique (Granger and Newbold, 1974; Granger, 1981; Engle and Granger, 1987; Johansen and Juselius, 1990).

### 2.1 Historical progression: Prelude to cointegration

We begin by giving a form of preface to the chapter. This section is mainly based on Hendry (2010). By referencing some earlier work by Jevons (1884), Hendry (2010) immediately reveals that time series analysis is not relatively new, more so, with respect to financial data.

History shows that time series data has mostly been assumed stationary over time. However, economic development has resulted in the growth of sample sizes, and consequently, economic data has become subject to the influence of time. Other factors that have had significant impact on time series data include political changes and technological advancements. As a result, the stationarity assumption has had to be reconsidered when modelling such data (Nelson and Plosser, 1982; Hendry and Juselius, 2001; Rahman, 2005). This has invalidated many traditional modelling techniques which could have failed to

anticipate such varied influence on the data (Nielsen, 2005). For example, continued use of standard techniques which assume unverified stationarity could result in suggested statistically valid relationships between certain time series variables that may in fact be entirely unrelated. This kind of biasedness is pointed out in Yule (1926). The author was a statistician who pioneered what is now known as spurious regression.

These spurious regressions or ‘nonsense relations’ are further researched by Granger and Newbold (1974). The authors prove that when independent non-stationary series are regressed on each other and the resultant error terms exhibit autocorrelation, then the hypothesis of no relation would frequently be rejected. This implies that the favoured alternative would then be a spurious relationship (Rahman, 2005).

Earlier scholars, such as Sargan (1964), demonstrate that dynamic regression systems can intuitively be strengthened by including the economic theory of equilibrium in the investigations of such systems. This is accomplished by using the autoregressive distribution to capture past disequilibria in models using the first differences of the variables investigated. As a result, Sargan’s accomplishments have led to many authors, including Davidson (1978), to popularize the now “error correction models”. To which, thereafter, Granger and Newbold (1974), Granger (1981), Engle and Granger (1987), Johansen and Juselius (1990) have championed the approach of systematically combining certain non-stationary variables to produce meaningful steady-state relationships by using the error correction models.

The analyses carried out in this study are based on the mentioned error correction models. Since the success of the error correction models has been shown to heavily depend on stationary data (Engle and Granger, 1987), we therefore proceed with the chapter by taking a closer look at this property.

## 2.2 Stationarity

### 2.2.1 Introduction

Stationarity became a fundamental part of time series forecasting since the work of Box and Jenkins (1976) in fitting time series data to autoregressive moving average models. A process is said to be stationary when it has a constant mean and variance over time. Additionally, the covariance between any two time periods depends, not on the time at which the covariance is calculated, but on the lag between the two time periods (Greene, 1993).

From the definition, two types of stationary processes are identified; a weakly stationary process and a truly/strongly stationary process. The former, also called covariance stationarity, only requires that the mean and the variance be constant over time. The latter, on the other hand, demands that not only should a process be covariance stationary, but also that all its higher-order moments are constant as well. This makes strongly stationary series less frequently witnessed in reality than weakly stationary series (Chatfield, 1980). Therefore, we only acknowledge the strongly stationary series for their mathematical properties. We do not highlight these mathematical properties in this research. Henceforth, any mention of stationary processes throughout this research will refer to weakly stationary series for reasons already mentioned.

### 2.2.2 Consequences of overlooking stationarity

Stationary processes are generally fairly predictable since their statistical properties can be assumed to be the same in future as they would have been in the past. Therefore, applications of trading strategies that wrongfully assume stationarity may lead to undesirable outcomes. This is akin to investment gambling with maximum improbability (Steenbarger, 2003).

Most financial time series data exhibit non-stationarity. Therefore, it is important to rectify this prior to any analysis to be carried out. This is true particularly if the intent is to test for cointegration. Dickey (2005: p1) remarks that this act facilitates "... estimation of model parameters whose properties are standard..." because any regression model based on non-

stationary variables invalidates any standard assumptions for asymptotic analysis (Beck, 2001). This basically means that attempting to model data in levels by means of such conventional techniques as the ordinary least squares estimation, the maximum likelihood estimation, for example, may give misleading results due to model misspecifications (Greene 1993).

Song and Chang (2003) show some of the effects of overlooking stationarity that have undermined some of the prior studies done on the southern US timber market. The authors show that failure to correct for stationarity in the studies they review resulted in inflated model coefficients and overestimation of the effects of the model variables due to serial correlation. The latter is the nuisance tendency of shock to persist indefinitely in non-stationary data. However, though serial correlation does affect the efficiency of the ordinary least squares estimators, it does not affect their consistency. This is demonstrated by Ramos (1996) and Beck (2008) in the investigation of individual and aggregate data regression analysis and vector autoregressive forecasting, respectively.

Stationarity is further defined in one of two ways, trend stationarity or difference stationarity. One school of thought argues that macroeconomic time series are trend stationary (Lucas, 1973; Sargent, 1978), yet another group of scholars suggests that financial time series exhibit stochastic trends instead (Dickey and Fuller, 1981; Campbell and Shiller, 1987; Lo and MacKinlay, 1988).

The former means that to achieve stationarity, the endogenous variable has to be regressed on the deterministic function (Gujarati, 1995). The latter simply says to achieve stationarity, the data must be differenced instead. This distinction helps to achieve two goals when modelling;

- 1) To give statistical validity to any subsequent tests and analysis by incorporating the correct form stationarity and,
- 2) Sound economic interpretation can be provided (Cochrane, 1991).

However, Granger and Newbold (1974), Song and Chang (2003) and Nelson and Plosser (1982) point out that application of stationarity may, in itself, culminate in new problems which in turn will require some other techniques to be addressed by. We leave and recommend this query for future research.

### 2.2.3 Tests for stationarity

We begin by looking at the less formal tests for stationarity such as the correlogram test. This test is presented as a plot of the sample autocorrelation function against time lags. For a non-stationary process, the autocorrelation function decays at a slow rate towards zero over time. However, when the process is randomly (to be discussed in detail in the Chapter 3) generated the autocorrelation function will be near or equal to zero for any and all time lags. A more comprehensive study of autocorrelations can be found in Box and Jenkins (1976).

Since random walk processes are characterized by unit roots, it has therefore become tradition to test for unit roots and use the results of such tests to infer upon the stationarity of the given series. Classic examples of these tests include the augmented Dickey-Fuller test (Dickey and Fuller, 1981), the Phillips-Perron test (Phillips and Perron, 1988), the Kwiatkowski-Phillips-Schmidt-Shin test (1992), the Ng-Perron test (Ng and Perron, 1995) and the Elliot-Lothman-Stock point-optimal test (Elliott et al., 1992). Cochran (1991) warns that unit root tests are often of low power in finite samples. Despite this, the tests provide a more formal criterion for classing a series into either a trend- or difference stationary model (Diebold and Kilian, 2001). This is crucial because both model types may imply very different outcomes even when applied to the same data.

Generally, unit root tests are applied to investigate the null hypothesis of difference stationarity against the alternative of trend stationarity. However, to capture the more complicated dynamics of most financial time series, Said and Dickey (1984) augmented the Dickey-Fuller (1981) test to accommodate the more general autoregressive moving average form. Under the augmented Dickey-Fuller (1981) test, the error terms are assumed to be homoscedastic and serially uncorrelated by specifying the lag length to that effect. Neglecting this assumption or incorrectly specifying the lag length may lead to the test being biased. Ng-Perron (1995) proposed a remedy that allows the lag length some flexibility so that it can change as sample size changes on condition that the error terms follow the general autoregressive moving average process.

Similarly, the Phillips-Perron (1988) test assumes the series is non-stationary under the null hypothesis. The test's test statistic shares the same asymptotic distributions as the Dickey-

Fuller (1981) test but the test has the advantage of being more robust to error terms exhibiting heteroscedasticity. Also, no prior specification of the lag length is required when applying the test.

An example of a case where the augmented Dickey-Fuller (1981) test is implemented is in West (2008). The author shows that the South African prime and money market rates are random walk processes. In order to achieve stationarity, the data is differenced and done so only once in this case as it was sufficient. The data used covered a 6 year period from January 2000 to January 2006.

Unlike the conventional unit root tests, the method proposed by Kwiatkowski-Phillips-Schmidt-Shin (1992) assumes stationarity of the data in the null hypothesis. This means that when the null hypothesis is rejected, the alternative hypothesis of unit root is concluded. Pfaff (2008) says that this testing procedure fits a conservative testing strategy that a researcher should always follow since the strategy puts the hypothesis of interest as the alternative one.

#### **2.2.4 Completeness of unit root tests**

Some unit root tests fail to distinguish between random walk processes and processes that are stationary but exhibit high persistence in both univariate and multivariate systems. As a result, the performance of such tests is adversely affected. This brings into question the efficiency of these tests and the effects that structural breaks may have on them. These and other such properties are discussed below as a means of addressing the completeness of unit root tests.

### 1. *Efficiency*

Issues that call for concern when considering the efficiency of some of the common unit root tests include adding deterministic terms to the test regression. This makes the tests severely size distorted leading to loss of power (Schwert, 1989; Caner and Kilian, 2001). To overcome such issues, Ng and Perron (2001) and Elliot, Rothenberg and Stock (1996) suggest the modified Phillips-Perron (1988) test and the Dickey-Fuller (1979) generalized least squares test as efficient test alternatives.

The needed lag lengths used in tests can be selected using the conventional Akaike (Akaike, 1970), Hannan-Quinn information criterion (Hannan and Quinn, 1979) and Schwarz-Bayesian Information Criterion (Schwarz, 1978) information criteria. However, the criteria must be modified to properly accommodate integrated data (Ng and Perron, 2001).

### 2. *Structural shifts*

Steenbarger (2003) says a successful trader is one who is mindful of in the market and allows for changes in trading strategies according to the prevailing distribution of price changes. It is, therefore, in the interest of this research to address the events that led to the recession recorded particularly between 2007 and 2009. Reasons for this enquiry are that these events may have been significant enough to affect the stock price distributions of the five market variables investigated here. This in turn may affect the determination of the “cointegratedness” of the five processes and the subsequent conclusions given after (Pfaff, 2008; Perron, 1988).

At this point in the research, we detour briefly to illuminate the subject of structural change.

### 3. *Structural change: Univariate and multivariate systems*

Literature defines structural breaks as changes in time series data following significant local or international occurrences. These include policy changes and strategic government moves such as passing new legislation that might influence economic activity in a particular manner. Whatever the causes maybe, structural shifts may have lasting effects particularly if they

contaminate non-stationary series (Pfaff, 2008). As a result, the testing power of standard unit root tests such as the augmented Dickey-Fuller (Dickey and Fuller, 1979) is weakened. It follows that the chances of committing a type two error are increased, that is, chances of not rejecting the null hypothesis of unit root when in fact it should be rejected (DeJong and Charles, 1991). In addition, overlooking such changes may provide a platform for bias towards selecting a unit root model while the underlying data generating process is (trend) stationary within specified breakpoints (Perron, 1988; Perron and Yabu, 2009).

For example, Nelson and Plosser's (1982) find that 13 out of 14 macroeconomic variables collected from the archives of the United States of America follow random walk processes after using the standard augmented Dickey-Fuller test (Dickey and Fuller, 1979). The data covered the period 1909 to 1970. However, the authors do not account for the 1929 market crash as a significant variable in this analysis.

It is also common that within such specified regimes, the investigated process(es) may alternate between stationary and non-stationary states, which may invalidate the consistency of the standard unit root tests. For example, Burdekin and Siklos (1999) give evidence that the exchange rate regime can explain increases in inflation persistence. Data from the United Kingdom, United States of America, Canada and Sweden is used in the investigation and attention is given to the effects of the end of the gold standard in 1914 and devaluation of the Sterling in 1967. These authors suggest that the end of the gold standard explains the shift in the inflation rate after 1914 better than the devaluation of the Sterling can explain the shift after 1967. It is therefore prudent to monitor not only change in the model coefficients but also the state of the process from regime to regime.

It is important to choose appropriate unit root testing procedures that account for shifts when data is suspected to be contaminated by breaks. For example, the testing procedure in Perron (1988) tracks changes in the drift and trend components while exogenously introducing a single break. However, some authors challenge this approach as data mining because of the exogenous introduction of a break in the test. The argument is that bias in such unit root tests can be minimised if the breaks are endogenously accounted for (Christiano, 1992; Zivot and Andrews, 1992; Banerjee et al., 1992).

One possible drawback that can result from analysing contaminated data is that valuable information can be lost in the process. This is accounted for by the fact that the volatile times characterized by breaks can be potentially informative about stock price movements (Mavrakakis and Alexakis, 2008). Therefore, to these authors, it is intuitive that in order to retain as much information as possible when testing the data for unit roots, the breaks should be accounted for endogenously. That is, without claiming prior knowledge of when the breaks may have occurred.

Perron (1988) re-examines the data in Nelson and Plosser (1982) mentioned before under the presumption that the 1929 stock crash could have permanently altered the behaviour of the data. Though Perron (1988) exogenously introduces this break in the analysis, the author, nevertheless, shows that it is significant enough to account for. Consequently, only 3 microeconomic variables are found to be unit root processes instead. On the other hand, Zivot and Andrews (1992) almost completely negate Perron's (1988) findings in favour of Nelson and Plosser's (1982) results. The reason for this is that Zivot and Andrews (1992) test for unit roots in the same data while endogenously exploring the possibility of a break. They find that only 3 of the variables might be stationary processes with breaks. Though this 'endogenous approach' proves to be an improvement on its predecessors, it however fails to address the possibility of significant multiple breaks within the data.

It is for this reason that Lumsdaine and Papell (2007) extend the Zivot and Andrews (1992) testing method. This extension allows for the testing of the unit roots under the null hypothesis while accommodating two breaks under the alternative hypothesis for both the level and trend components. Since the test does not allow for breaks in the null hypothesis, there is then a tendency to conclude against random walk processes (John, Reetu and Nelson, 2007; Allaro et al., 2010).

Therefore, in order to account for breaks under both the null and alternative hypotheses, John et al. (2007) and Allaro et al. (2010) argue that Lee and Strazicich's (2003) endogenous Lagrange Multiplier test might be a superior choice. This is because this procedure reduces spurious rejection of the null hypothesis by using the Lagrange Multiplier principle in the calculation of the test statistic as Lee and Strazicich (2003) demonstrate.

The reader can find a more detailed and comprehensive summary of other unit root tests and results based on Nelson and Plosser's (1982) data in John et al. (2007) for further understanding of material that has been discussed in this section of the chapter so far.

It is also possible to explicitly test data for unit roots and structural breaks separately. In this case, testing for breaks involves investigating parameter constancy of the regression model of interest. This is the null hypothesis. The alternative hypothesis is that breaks are present in the data. Zeileis (2001) categorises such tests into two groups based on their deviation patterns from the null hypothesis as the  $F$  test framework and the generalized fluctuation test framework (Kuan and Hornik, 1995; Hansen, 1992; Andrews, 1993).

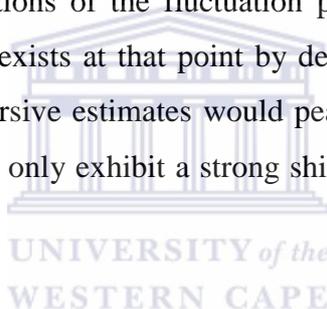
The  $F$  test framework is based on Chow's (1960) work on single shift processes. The unknown break point is investigated by calculating  $F$  statistics for an interval of potential change points, then the supremum of the calculated statistics is compared to some critical value. If the supremum is found to be greater than the critical value, the null hypothesis is rejected and the data is said to be contaminated by a structural break. Thus the test determines that it is significantly better to estimate two sample portions by the method of least squares compared to estimating the entire data sample all at once (Zeileis et al., 2001). Chow's (1960) test is later developed further by Andrews and Ploberger (1994) to include the average and exponential variants of the  $F$  test. These differ from the supremum  $F$  test in that they can only detect the presence of the breaks but not their corresponding break dates.

By dropping the *a priori* assumption that there only exists one significant break in the data, we now turn to multiple break cases. One such example is Yamamoto's (1996) stepwise Chow procedure. Hayashi (2005) demonstrates that the stepwise Chow test has the added benefit of tracking changes in the lagged structures and in determining whether shifts are temporal or continuous. Another approach is to expand the  $F$  test into the double maximum tests (Bai and Perron, 1998, 2003). This allows for the accommodation of an unknown number of breaks in the alternative hypothesis while allowing for the discretion to select a maximum number of breaks to be tested for.

The generalized fluctuations outline provides tests that are either estimate or residual based. This means that in the presence of a structural break, the estimated coefficients of the

regression fitted to the entire data sample should be significantly different from the estimates based on the subsamples of the data provided the subsamples do not contain the breaks (Zeileis, et al., 2001; Zeileis, et al., 2003). However, if the true coefficients remain constant over time then the corresponding estimates would be quite similar. The choice of the subsamples can be achieved either recursively, that is, starting with a selected number of observations and then sequentially, including each successive observation, or by a window of constant size that “rolls” over the entire sample period. The differences between the estimates of the subsamples and the overall data sample result in the computation and mapping of empirical processes.

Thus, under the null hypothesis of no breaks, the plotted empirical process defined within some confidence bounds should not deviate too far from zero. The bounds are constructed from known asymptotic distributions of the fluctuation process. If the fluctuation process crosses the bounds then a break exists at that point by defining an appropriate significance level. The plot based on the recursive estimates would peak around a break point while the moving (rolling) estimates would only exhibit a strong shift about the same point (Zeileis et al., 2001).



The fluctuation processes can also be computed from cumulative and moving sums of either ordinary least squares residuals or recursive residuals. Brown et al. (1975) introduce the use of cumulative sums of recursive residuals such that under the alternative hypothesis, the residuals will have zero mean up until a break point in time. Based on the standard Brownian motion, the test statistic would favour a structural break when the process deviates from its zero mean as the errors become large. Alternatively, Ploberger and Krämer (1992) suggest that the cumulative sums test be carried out by using the residuals estimated by means of the ordinary least squares method. In this case, the fluctuation processes peak around a structural break while a moving sums test based on the same estimates would show a strong shift.

Other tests that are used as aggregates of fluctuation processes in order to investigate their deviations from their respective limiting processes include the Cramér-von Mises statistic (Nyblom, 1989; Hansen, 1992). This test is based on maximum likelihood scores and tests the null of parameter stability against the alternative of random walk. According to Zeileis (2006, p. 5) the Cramér-von Mises statistic can be constructed by “...aggregating first over

the components using the squared Euclidean norm  $L_2$  and then over time by taking the mean". However, if the parameters shift in one direction and then shift back, then aggregating over time would be better served using the range in this case (Zeileis, 2006). The null is then rejected if the maximum range is found to be too large. Lastly, we acknowledge the sequential testing procedure (Perron and Zhou, 2008; Bai and Perron, 1998). The test investigates  $l$  initial breaks against additional breaks in each  $(l + 1)$  data segment. Thus the testing procedure amounts to performing a one break test for each successive segment. The estimated break points are obtained by a global minimization of the sum of squared residuals (Perron, 2006) or by maximizing the Gaussian log-likelihood function.

Besides minimizing the residual sum of squares to estimate the number of breaks, the Bayesian information criterion or the modified Schwarz' criterion (Liu et al., 1997) could also be used in that regard.

The vast range of disciplines that have focused on parameter stability has grown to show how important it is to consider structural shifts in real world time series data. For example, Ramirez-Beltran and Olivares (1999) estimate the expiration date of one pharmaceutical company's product based on tests of parameter stability by comparing subsample regression equations to the population regression. However, the authors consider a limited sample size across a short time line. If a substantially large data set over much longer time span is considered instead, then the time delay phenomenon can become evident as reported in Allaro et al. (2010). In their report, the authors suggest that in the time period from 1974 to 2009 the Ethiopian economy may have experienced structural change years after the corresponding policy changes were made.

We therefore emphasize and reason that our choice to recognize such volatile periods in our sample instead of discarding them is that they could be potentially informative about the stock market behaviour as pointed out by Mavrakis and Alexakis (2008). This view is demonstrated by Zhang et al. (2006). In their research, the authors investigate breaks in a sample of A and B share indices traded on the Chinese stock market covering the time period from 1995 to 2005. Two breaks are found to have had significant effect on their supporting data and these disturbances are attributed to the regulatory shift in 2001 and the Asian

financial crisis. After identifying and dating the breaks, the authors then test the data for cointegration. For this, they use the Gregory and Hansen (1996) test, a procedure that tests for cointegration under the assumption that the data is contaminated by breaks.

The results indicate that the indices are cointegrated. However, this is true only within subsamples defined between and around the two break points. On the other hand, the data is found not to be cointegrated over the entire sample period. However, the authors only test the individual data series using the standard augmented Dickey-Fuller and Phillips-Perron unit root tests.

Another example of data shown to be afflicted by breaks can be found in Kankesu et al. (2007). In this case, the authors use the Zivot and Andrews (1992) unit root test on the Australian export and import data. The estimated break dates show some indication that the shifts in the data may possibly have been the result of the country's trade liberalization policy review. Similarly, Rubio et al. (2008) use Bai and Perron's (1998, 2003) multiple break procedure to assess the sustainability of the US budget deficit over the period from 1947 to 2005. The authors show that cointegration between public revenues and expenditures exists only in some but not all regimes defined around the break dates. As a result, inference on the sustainability of the US budget deficit is explained within each regime based on the coefficients estimated for each cointegrating equation modelling the data in each of the said regimes.

Since this research only considers cointegration over the entire sample period, it is our intention to fit at least one cointegrating equation to the entire data set where cointegration is shown to exist. The effects of any structural breaks found to contaminate the data sample will be captured by shift dummy variables. These variables will provide a qualitative interpretation of such oddities (Reade, 2005).

The use of vector error correction models fortified with varied dummies can also be found in Mavrakis and Alexakis (2008), Lütkepohl et al. (2004) and Zhang et al. (2006). Appropriate dummy variables distinguish between dichotomous groups, this enables the variables to introduce the qualitative effects of significant events into the regression model (Rycx, 2012). Also, the asymptotic distributions of Johansen's (1988) cointegration rank tests may be

affected in such a way that the reported rank could be false. Lütkepohl (2004) demonstrates how ignored effects of a level shift can possibly result in biased rank determination in a vector error correction model.

Different types of dummy variables are identified and defined in Rycx (2012) and Hendry and Juselius (2001) discuss the use of transitory dummy variables to capture the effects that the Kuwait war had on gasoline prices in the early 90s. These are used based on the assumption that the effects of an event are short-lived and die out soon after the event has occurred. Prices stabilize to their equilibrium after the event has occurred. Thus these variables are assigned the value one only when the event occurs and zero everywhere else. For the events with more permanent effects to the system, the step / shift dummy variables are used. This type of variable is restricted to the cointegrating space, meaning that the variable will be included in the full lag structure of the final model. “This amounts to setting the residuals equal to zero for the transition from one regime to another (Johansen, Mosconi and Nielsen, 2000)” as cited in Mavrakis and Alexakis (2008: p. 167). To capture the permanent effects, shift variables are assigned the value one from when the event occurs and onward. In investigating the US, Greece, Germany and United Kingdom stock prices from 1991 to 2004, Mavrakis and Alexakis (2008) discover two cointegrated equations which they normalize on the Germany stock prices. In both equations, the structural dummy variable is found to be quite significant. This implies that the introduction of the euro in May 1998 as a single currency of the Member States of the European Council Regulation had a significant impact on the price returns of the Germany stock market.

Thus far, the necessary conditions for cointegration have been discussed. These include stationarity and “integratedness” of time series data, unit root tests and the effects of structural shifts that may bias the necessary analyses. Next, we define and discuss cointegration and its application in the relevant literature. Discussions on heteroscedasticity and statistical arbitrage follow thereafter.

## 2.3 Cointegration

Granger and Newbold (1974) show that traditional modelling tools can produce statistically valid relationships even when regressing independent non-stationary variables on each other. As a result, relationships between certain variables can be mistakenly thought to exist when in fact they do not. Yule (1926) defines such relations as ‘non-sense regressions’, also known as ‘spurious regressions’. To overcome this flaw, Engle and Granger (1987) develop a technique to linearly combine two or more time series that are integrated of the same order such that the resultant process is stationary. This technique is called cointegration. As it will be shown later, this procedure is a more powerful tool when modelling financial time series compared to the traditional methods because it reduces the chances of spurious regressions.

The cointegration methodology captures both short- and long-run system dynamics that characterize time series data (Alexander, 1999; Gujarati, 1999). This is achieved by monitoring historical data to see if the series ‘move’ together and then hypothesising that the data will continue to do so in future also. Alternatively, any linear time trends could be removed from the data and the empirical relationships can be modeled using detrended data. The downside of analysing dynamic systems solely comprising of detrended variables is that any means to detect common trends in the data is foregone. This leads to the approach being only able to characterize short-run dynamics but not the long-term relationships and consequently to loss of valuable information (Alexander and Dimitriu, 2002; Engle and Granger, 2003; Schmidt, 2008). The argument just presented differentiates cointegration from correlation, that is, correlation does not imply cointegration (Alexander, 1999).

Sims (1980) points out that modelling economic series without making unnecessary assumptions can be achieved through the use of unrestricted vector autoregressive models. This idea is then enhanced by Engle and Granger (1987) to what is now called the Engle-Granger two-step cointegration test. This process incorporates the theory of long-run equilibrium into the model so that any disequilibrium is captured and the speed of adjustment back to equilibrium is quantified. Besides the Engle-Granger approach, Johansen (1988) provides the method of maximum likelihood estimators as an alternative way of estimating cointegrated vector autoregressive models. This approach has the advantage of being able to

determine the existence of more cointegration relationships than the Engle and Granger two-step method.

The Johansen (1988) method has since popularized the use of vector error correction models to explain multivariate cointegration. An advantageous feature of such models is that historical information can be used in forecasting (Schmidt, 2005).

Some of the research that has been based on cointegration includes the study of the market equilibrium of the South African forward prime rates (West, 2008). The study uses a two-step cointegration testing method developed by Engle and Granger (1987) to show that between the years 2000 and 2006, the JIBAR and the forward prime rates shared a long-term equilibrium relationship. Maysami et al. (2004) investigate and find that Singapore's macroeconomic variables are cointegrated with the Singapore stock index (STI) between January, 1989 and December, 2001. The scholars use the Johansen and Juselius (1990) method instead of the Engle and Granger (1987) procedure as they argue that the protocols of the former yield more efficient model estimates than the latter.

According to Granger (1987), cointegration implies causality at least in one direction between the variables. To illustrate this point, Christopoulos and Tsionas (2003) show that, not only are financial development and economic growth cointegrated, but there is also causality from the former to the latter. In the test for cointegration, the authors caution on severe distortions suffered by the Johansen and Juselius (1990) procedure in multivariate systems on account of small sample size.

Omran (2003) finds that Egypt's real interest rates are cointegrated with Egypt's stock market performance in the 18-year period since 1980/81. The scholar defines market performance as both market activity and market liquidity. Market liquidity is further defined by two variables, the total value traded to market capitalization and the volume of shares traded to the volume of shares listed. The real interest rates are found to be cointegrated with both variables mentioned in the latter; as a result, the author concludes that the real interest rates are cointegrated with market liquidity. Since market activity and real interest rates are also cointegrated, it is then concluded that real interest rates have an impact on the market performance.

Similarly, the Jakarta Stock Exchange data, the Singapore Stock Exchange data, and the Kuala Lumpur Stock Exchange are found to be cointegrated between 1997 to 2006 (Febrian and Herwany, 2007). Based on the significance of the error correction terms from the vector error correction model, the authors show that the Singapore Stock Exchange is weakly exogenous to the system. As a result, causality is said to be emanating from the Singapore variable to the rest of the system.

Other studies that cover cointegration with respect to Asian markets can also be found in Subramanian (2008) and Guidi (2010). Cointegration based causality is also investigated in the study of the purchasing power parity (Islam and Ahmed, 1999) and in the study of futures and spot prices (Gebre-Mariam, 2011; Alexander, 1999).

Some researchers argue that to a degree, cointegrated systems could suggest some level of predictability. This in turn could upset Fama's (1970) efficient market hypothesis resulting in market inefficiencies, for example, within the foreign exchange market when currency strength and directional influences are considered (Alexander and Dimitriu, 2002). However, others debate that although causality is implied by cointegration, the markets remain efficient (Dwyer and Wallace, 1992; Crowder, 1994). Interestingly, Wilson and Marashdeh (2006) propose that market efficiency might hold in the short-run but not in the long-run.

### **2.3.1 Cointegration tests**

We identify two main classes of tests mostly discussed in literature: error/residual based tests and maximum likelihood ratio tests.

#### *(i) Error/ Residual based tests*

The residual based tests investigate unit roots on residuals of single equation regression models under the null hypothesis of no cointegration. Pesavento (2007) gives a detailed analytical comparison of five such tests based on performance, asymptotic power and size distortions under the null hypothesis of no cointegration. Some of these tests focus on the computation of the asymptotic distribution of the (augmented) Dickey-Fuller (1979) test

(Engle and Yoo, 1987; Phillips and Ouliaris, 1990), but perhaps the most acknowledged in literature is the Engle-Granger (1987) two-step method. The reason is that it is more comprehensible and preferable from a financial view point, and according to Pesavento (2007), no alternative uniformly powerful test exists to give reason to opt otherwise.

### *Engle-Granger two-step*

This test seeks out the linear combination that has minimum variance and is used to estimate vector error regressive processes with cointegration in two steps (Alexander, 1999) below:

#### *Step 1*

- Estimate the cointegrating relationship by ordinary least squares and test for cointegration by applying the augmented Dickey-Fuller (1979) test to the residuals from the regression.

If the test results favour no cointegration then reconsider model specification, otherwise,

#### *Step 2*

- Construct the error correction model and use the lagged residuals from the cointegrating relationship in step 1 as the error correction term.

Though the Engle-Granger (1987) method is easy to implement, it does however have some drawbacks. Firstly, Schmidt (2008) argues that errors considered in the second step of the Engle-Granger (1987) procedure taken from the approximated regression in the first step render this approach subject to two times the estimation error. Secondly, in multivariate situations, at least two cointegrating relationships may exist. As a result, the two step procedure becomes inadequate in such cases (Johansen and Juselius, 1990).

Thus, Phillips and Ouliaris (1988) suggest a multivariate test which determines the rank of non-parametrically estimated system of processes. The upside of this procedure is that it provides the analyst with the ability to determine the number of common trends present in the given system as discussed in Stock and Watson (1998). However, the test statistic is found to be biased where short-run dynamics are concerned. Alternatively, maximum likelihood ratio

tests can also be applied in multivariate analysis where the null of no cointegration is of interest. The likelihood estimator for long-run equilibrium relationships and likelihood ratio tests for cointegration in vector error correction models are derived in Johansen (1988). This approach is said to be asymptotically equivalent to that of Ahn and Reinsel (1988) with the advantage of that the sample statistic can accommodate small sample distortions (Stock and Watson, 1993).

(ii) *Maximum likelihood ratio tests*

Generally, the maximum likelihood ratio tests are used to determine which model best explains the given data. The Johansen procedure uses this property to determine if the variables of interest can be said to share any long-run equilibrium relationships and how many such relationships can be determined.

*Johansen / Johansen and Juselius method*

The Johansen (1988) (or Johansen and Juselius (1990)) testing framework develops the Engle-Granger two-step method further so that multivariate analysis, and hence, the investigation of more than one cointegrating relationship is possible. By using the likelihood estimator for the vector error correction model, the procedure is able to identify a linear combination that is most stationary given a set of variables to be tested for cointegration (Alexander, 1999). Needless to say, the model's error terms must be independent and identically distributed with mean zero and constant variance. In the presence of significant structural changes in the data, the cointegration test would have to be augmented by exogenously including appropriate dummy variables in the vector error correction model (Zhang et al., 2006).

One of the immediate weaknesses of this approach is that most of its small sample properties are unknown (Toda and Yamamoto, 1995).

## 2.4 Cointegration based trading strategies

The Index tracking and Enhanced tracking strategies considered in this sub-section are discussed extensively in Alexander and Dimitriu (2002) which is heavily influenced, in part, by Lucas (1997), particularly on the application of cointegration on asset allocation.

### 2.4.1 Index tracking

This strategy aims to replicate a benchmark in terms of returns and volatility by first specifying the stocks in the tracking portfolio and second, clarifying the weights of each stock in the portfolio by means of the cointegration optimization procedure (Alexander and Dimitriu, 2002).

### 2.4.2 Enhanced index tracking

This strategy extends on the simple tracking strategy by replicating artificial 'plus' or 'minus' indexes. A trader holds a long position in a portfolio tracking the 'plus' benchmark and a short position in a portfolio tracking the 'minus' benchmark. The trader assumes that the asset prices will revert back to the average price given enough time. This is a form of a statistical arbitrage trading strategy (Alexander and Dimitriu, 2002).

### 2.4.3 Statistical arbitrage

Generally, statistical arbitrage is a trading strategy that provides a trader an arbitrage opportunity based on exploiting temporal mispricing of one or more assets with respect to their expected values. It differs from deterministic arbitrage in that it is not riskless, that is, it heavily depends on the asset prices returning to their historical or equilibrium values. However, the upside to the strategy is that it has zero cost, positive probability of a positive payoff and a diminishing variance as time tends to infinity (Hogan et al., 2004). The strategy considers large asset quantities with holding periods that range from seconds to days or even longer.

The dependence of statistical arbitrage to mean reversion strongly relates it to the cointegration technique, hence the interest of this research to investigate it. Hedge funds often implement statistical arbitrage mainly because of the strategy's low volatility and independent returns. These properties show strong dependence on stationary data. Furthermore, the returns are often found to be uncorrelated with the market and relatively high or constant despite economic downturns.

(i) *Risks associated with statistical arbitrage strategies*

The trading strategy is model based; as a result it is not immune to changes in the distribution of returns and other risks specific to the underlying security. As arbitrageurs take advantage of the temporal deviations as they occur, the market is "corrected" and efficiency is improved. This reduces any arbitrage scope and demands constant model re-specification. On the other hand, if the strategy is exercised in the short-run then, it is possible to record enough considerable losses to prompt a trader to default.

(ii) *Statistical arbitrage strategies*

A wide selection of strategies is defined under statistical arbitrage. These strategies share a few commonalities including statistical means for generating excess returns, systemic trading signals and market neutrality (Infantino and Itzhaki, 2010). We discuss some of these strategies below:

a. *Mean-reverting strategies - Pairs trading*

This strategy is also referred to as spread trading. It allows a trader to capitalize on the differences between stock prices while maintaining a risk neutral position. The trader implements a hedge by going long in one asset and going short in another. The spread between the two positions generates the excess returns.

The strategy is based on cointegration since the latter guarantees mean reversion (Hogan et al., 2004; Alexander and Dimitriu, 2002). However, cointegration does not guarantee the success of the strategy (Schmidt, 2008). In a trading interval, a pair can have multiple positive cash flows as long as the pair prices converge each time they diverge (Gatev et al., 2006). The above scholars advise that the returns from the strategy may be biased upward due to the bid-ask relationship since the pairs trading strategy is a contrarian trading strategy. An extended discussion of contrarian strategies can be found in Khandani and Lo (2007) in relation to the 2007 crisis as well as in Sudak and Suslova (2009).

Cheng (2008) simulates the implementation of the pairs trading strategy by tracking 0363.HK and 0882.HK, two assets listed on the Hong Kong stock exchange. The assets were selected on the 26th of September 2007 after having been observed over a 500 days training period prior to the selection day. The author shows that by regressing 0363.HK on 0882.HK, the resultant stationary process can be tracked using a trading rule defined by a two standard error threshold boundary. The threshold forms part of the trading signals which indicates when a long or short position can be held in the portfolio. Thus when the basket of the two assets is overvalued (undervalued), a short (long) position is opened. In addition, the above author demonstrates how the returns accumulated from trading the pair can be substantially eroded by transaction costs which may be incurred from portfolio rebalancing.

The pairs trading strategy is also simulated by Gatev et al. (2003) and Andrade et al. (2005). Gatev et al. (2003) show that the strategy yields annualized excess returns of about 11% based on daily US data from 1962 to 2002. Andrade et al. (2005) on the other hand, find excess annualized returns of 10.18% while using daily stock prices of all stocks listed on the Taiwan Stock Exchange covering a period from 5 January 1994 to 29 August 2002. Furthermore, Andrade et al. (2005) find that the opening of a pairs trading position and the uninformed shocks to the underlying stocks are highly correlated. As a result, the returns of relative value trading strategies may be explained by uninformed trading shocks according to the above authors.

### *b. Multi-factor strategies*

This set of strategies is based on the influence of several chosen factors on the stock return. After running multiple regressions, a portfolio is constructed from stocks according to their respective correlations with the returns. Sudak and Suslova (2009) mention the Arbitrage Pricing Theory as one example of a multi-factor strategy.

## **2.5 Heteroscedasticity**

Engle (1982) points out that when analysing non-stationary time series, it is quite implausible to assume that the forecast variance remains constant over time. To address this issue, the above scholar introduces the autoregressive conditional heteroscedastic processes. These exhibit constant unconditional variances even though they allow for non-constant variances that are conditional on the past. This modelling of time series has proven useful in explaining different economic phenomenon. For example, Engle's (1982) approach is used to show that estimated inflation volatility is related to some economic macro variables in Coulson and Robins (1985).

The hitch with using Engle's (1982) method is that in order to avoid violation of the non-negativity constraints, a fixed lag structure has to be imposed in the variance equation. As a result, Bollerslev (1986) generalizes Engle's (1982) 'processes' so that they allow for a more accommodating lag structure. This is achieved by catering for lagged conditional variances in the variance equation. Since then, different variations of the generalized model have been provided in literature. These include Nelson's (1991) exponential variation, Engle and Ng's (1993) nonlinear variation, the quadratic version by Sentana (1995) and many others.

Over time, many authors have adopted the approach of testing for and considering cointegration in the presence of heteroscedasticity including Lee and Tse (1996). Bauwens et al. (1997) show that accounting for heteroscedasticity when testing for cointegration improves the long-term forecasting power of the vector autoregressive model. The authors develop this argument after modelling long and short term interest rates for five countries taken from 1960-1995. Liu and Shrester (2008) reach a similar conclusion after employing

the Johansen and Juselius (1990) procedure. In this case, the authors investigate the existence of long-term relationships between the Chinese stock market and its macroeconomic variables. Other studies that have considered this type of heteroscedastic cointegration include Franses, Kofman and Moser (1994), Hansen (1992), Boswijk and Zu (2007) and Leykam (2008).

### **2.5.1 How heteroscedasticity can occur**

In summary, a few instances attribute to the occurrence of heteroscedasticity. Model misspecification through omission of relevant variables or using the raw data as it is instead of transforming it into its natural logarithms is one way. Failure to properly account for subpopulation interactions can culminate in under- or overestimated model errors. In some cases simple errors in measurements can also be a contributing factor.

### **2.5.2 Consequences**

Though ordinary least squares estimates may remain unbiased after a regression analysis in the presence of heteroscedasticity, the same cannot be said for the estimates of the standard errors. Failure to acknowledge this may lead to, for example, a type one error when conducting a hypothesis test. However, other scholars, including Gujarati (2009), point out that a good model need not be discredited on account of non-constant variances. Hence Engle (2001) remarks that heteroscedasticity is best modeled rather than handled like a problem to be resolved.

### **2.5.3 Tests**

Several methods are available to an analyst to test for the presence of heteroscedasticity before analysing the data. These include a less formal visual inspection component obtained from plotting error values against the fitted values and analysing the error sizes relative to the predictor variables. If variances are suspected not to be constant, then more formal tests will

need to be employed. These include the Goldfeld-Quandt test (Goldfield and Quandt, 1965), Breusch-Pagan test (Breusch and Pagan, 1979) and White's general test (White, 1980).

White's (1980) test uses the method of regressing the squared residuals from the regression equation onto the regressors. If homoscedasticity is rejected then one employs Bollerslev's (1986) generalized autoregressive conditional heteroscedastic model. The test statistic used in the White test is the Langrange multiplier which is a product of the sample size of the residuals and the squared residuals. Engle (1982) suggests the use of the Langrange multiplier test under the null of no autoregressive conditional heteroscedasticity. The application of this approach can be found in Liu and Shrester (2008). Both authors find that the Chinese stock market is cointegrated with the country's macroeconomic variables in the period form 1992-2001. Furthermore, the Chinese stock market performance is found to be positively related to that of the macro-economy in the long run.



## 2.6 Conclusion

Based on the vast literature covered above, this research adheres to Box and Jenkins' (1976) approach to modelling time series data. This approach comprises three particular steps which are identification, model selection and diagnostic checking. The first step sheds light on the data's properties by first inspecting time series plots and autocorrelation functions of the data, testing for normality and testing for heteroscedasticity according to Pfaff (2008). The data is then tested for unit roots by applying the augmented Dickey-Fuller (1979) and the Kwiatkowski, Phillips, Schmidt and Shin (1992) tests. This choice of tests covers the generally known testing procedures ranging from traditional to conservative testing strategies (Pfaff, 2008).

Upon concluding that the data is indeed unstable in levels, it is then transformed to its stationary form. This will facilitate consistent coefficient estimates to be calculated for the yet to be determined cointegrating equation(s). Next, the data is inspected for structural breaks in levels by observing any "blips" in the charts of the individual stock's returns series. Then formal tests for unit roots with structural breaks are applied in the form of the Andrew-Zivot (1992) test, the supremum  $F$  test and the recursive based cumulative sums test (Zeileis,

2001). From this point on, any all subsequent analysis will include the effects of the structural breaks detected by the tests and appropriately dated.

The second step according to Box and Jenkins' (1976) follows with the application of trace and maximal eigenvalue tests for cointegration and rank determination to the data (Johansen and Juselius, 1990, Johansen, 1991, 1995). The series considered in this step can either all be integrated of the same order or mixed with stationary variables (Johansen, 1991, 1995; Mavrakakis and Alexakis, 2008; Hjalmarsson and Österholm, 2007). When applying the Johansen and Juselius (1990) procedure, care is taken to test, identify and capture the effects of any breaks in the data. This is because the test is known to have its asymptotic properties that can be affected by the presence of such structural break(s) in the data. The dummy variables incorporated into the rank determination phase will give qualitative descriptions of the effects caused by the extraordinary events in the long-run.

Finally, to facilitate a dynamically stable equilibrium relation, model testing is conducted. Three testing procedure are carried out, the Jarque and Bera test (1980) for normality, the Breusch-Godfrey (1980) test for serial correlation of the model residuals and Engle's (1982) autocorrelated conditional heteroscedasticity test for volatility clustering in the stock returns.

## CHAPTER 3

### Methodology

Before we explore the numerous and various tests and other procedures we use in this research, we provide the reader with a diagram summarizing the methodology discussed in this chapter and followed in the analysis of the data. All the variables and other notations are detailed and explained in the discussion to follow. We begin the summary by following the three basic steps given below:

Step 1: Transform all data to log-values to stabilize sample variability

Step 2: Difference all log-transformed data to track price changes over time

Step 3: Estimate the Andrew-Zivot (1992) model C

$$\Delta y_t = \mu + \eta t + \gamma DU_t + \varphi DT_t + \delta y_{t-1} + \theta_i \sum_{i=1}^p \Delta y_{t-i} + \varepsilon_t,$$

UNIVERSITY of the  
WESTERN CAPE

Step 4: Test for unit roots as displayed in the flow chart below:

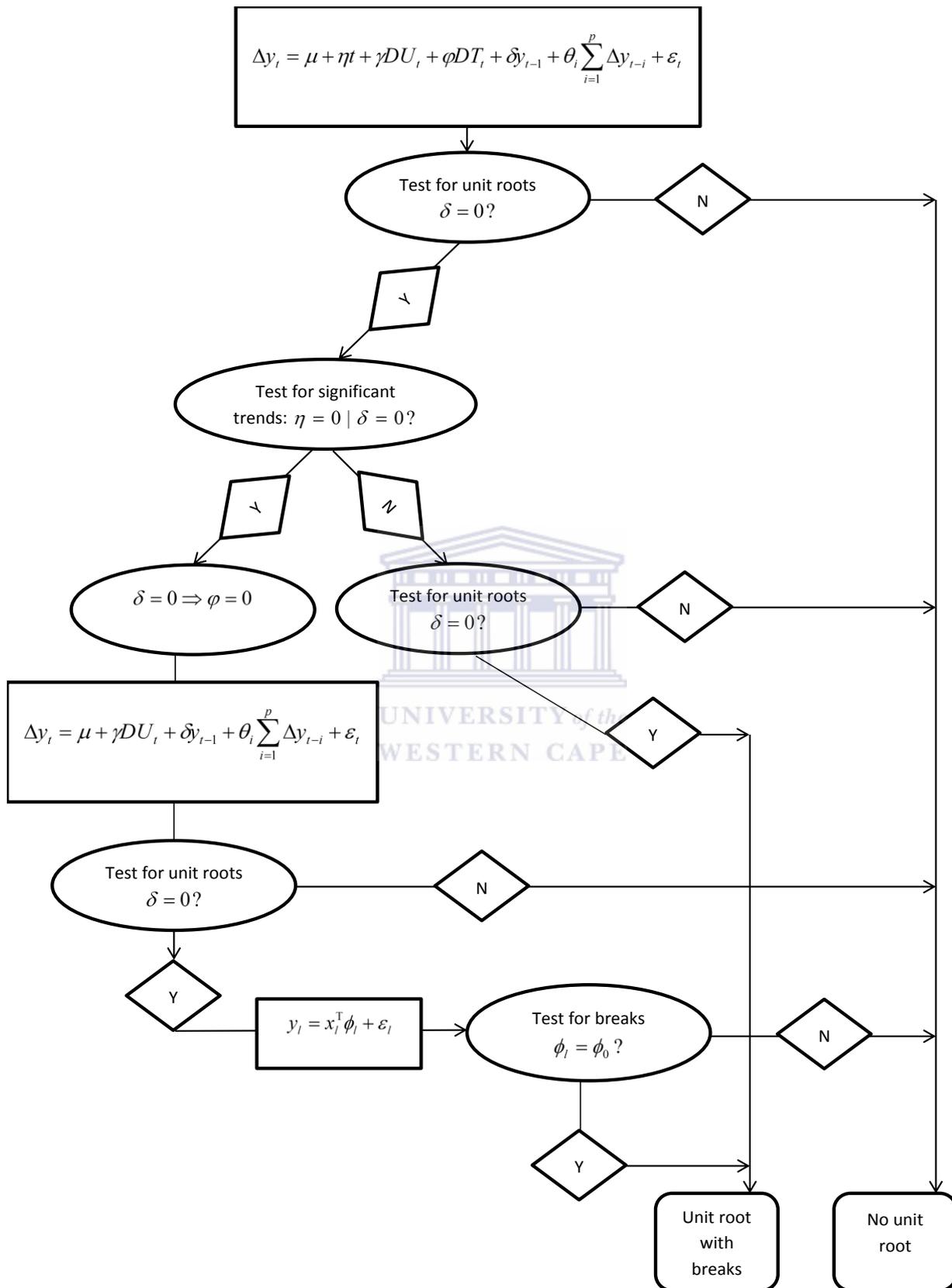


Figure 1: Testing sequence for unit roots; with and without structural breaks.

This chapter outlines the statistical modelling exploited in the analysis and testing stage that is demonstrated in the next chapter. The chosen tests and procedures used in the analysis are sequentially laid out to give direction of how the background theory covered in Chapter 2 is put into practice in chapter 4.

Alexander (2008) states that financial time series data have stochastic trend and not deterministic trend because they are generated by efficient financial markets. Therefore, in view of our stock price data, we begin by considering a simple time series regression given by

$$y_t = y_{t-1} + \varepsilon_t. \quad (1)$$

This regression, (1), is called a first order autoregressive process, abbreviated AR(1). The variable  $y_t$  represents the price of some stock  $y$  at time  $t$ . The stock price  $y_t$  is determined by the joint influence of a single past price  $y_{t-1}$  (hence “autoregressive”) of  $y$  at time  $t - 1$  (hence “first order”). A random variable,  $\varepsilon_t$ , captures the instantaneous impact of new information arriving into the market at time  $t$ . Therefore, this new information is incorporated into the price  $y_t$  such that new information coming in at time  $(t + 1)$  is independent of price  $y_t$ . It follows then that the best prediction of any future price beyond time  $t$  is  $y_t$ . Because of that, the price process in (1) is called a random walk. To facilitate the understanding of the underlying data generating process(s), we analyze the stock prices as log transformed data.

The error terms,  $\varepsilon_t$ , of a random walk process are independent and identically generated such that

$$\begin{aligned} E(\varepsilon_t) &= 0, \quad \forall t \\ \text{Var}(\varepsilon_t) &= \sigma_\varepsilon^2, \quad \forall t \\ \text{Cov}(\varepsilon_t, \varepsilon_s) &= E(\varepsilon_t - E(\varepsilon_t))(\varepsilon_s - E(\varepsilon_s)) \\ &= E(\varepsilon_t \varepsilon_s) \\ &= 0, \quad \text{for } t \neq s. \end{aligned} \quad (2)$$

Henceforth, we denote an independent and identically distributed process as i.i.d. As long as  $E(\varepsilon_t)$  and  $Var(\varepsilon_t)$  are finite,  $\varepsilon_t$  remains stationary.

When a mean ( $\nu$ ) and a slope coefficient ( $\omega$ ) are included in (1), the equation can be written as follows

$$y_t = \nu + \omega y_{t-1} + \varepsilon_t. \quad (3)$$

Equation (3) is called a random walk process with drift provided  $|\omega|=1$ . (To be explained further in section 3.1).

A generalization of model (3) is defined as follows

$$y_t = \nu + \omega_1 y_{t-1} + \dots + \omega_p y_{t-p} + \varepsilon_t \quad (4)$$

where the  $\omega_i$ 's are parameters for  $i=1 \dots p$  and some  $p$ . Equation (4) is then called an autoregressive process of order  $p$  abbreviated AR( $p$ ). The process  $y_t$  is generated by only its past values and the random error satisfying the set of conditions in (2).

Pfaff (2008) shows that from regression (3), the process  $y_t$  can also be explained by past shocks with decaying weights as follows:

$$\begin{aligned} y_t &= (\nu + \varepsilon_t) + \omega(\nu + \varepsilon_{t-1}) + \omega^2(\nu + \varepsilon_{t-2}) + \dots \\ &= \left(\frac{\nu}{1-\omega}\right) + \varepsilon_t + \omega\varepsilon_{t-1} + \omega^2\varepsilon_{t-2} + \dots \end{aligned} \quad (5)$$

This representation is called an infinite order moving average process. Letting  $\mu = \left(\frac{\nu}{1-\omega}\right)$  and  $\omega^i = \vartheta_i$  for  $i = 0, 1, 2, \dots$ , then  $y_t$  can be modeled as a finite moving average of its past errors given by

$$y_t = \mu + \vartheta_0 \varepsilon_t + \vartheta_1 \varepsilon_{t-1} + \vartheta_2 \varepsilon_{t-2} + \dots + \vartheta_q \varepsilon_{t-q} \quad (6)$$

such that coefficients  $\vartheta_i$ , for  $i = 0, 1, 2, \dots, q$ , can be any real numbers. The parameter  $\mu$  is the mean of  $y_t$  and is equivalent to the initial value  $y_0$  when  $t = 0$ . Model (6) represents a moving average process of order  $q$  and is abbreviated MA( $q$ ).

In process (6) each  $\vartheta_i$ , for  $i = 0, 1, 2, \dots, q$ , is called a coefficient of autocorrelation such that  $-1 < \vartheta_i < 1$ . In a higher order regression such as (6)  $y_t$  is stationary when the effects of previous shocks die out with time, that is, when  $|\vartheta_i| < 1$ . Therefore  $y_t$  is non-stationary when these effects persist indefinitely, that is, when  $\vartheta_i = 1$  (Hendry and Juselius, 2000).

Literature has it that in other instances, data may exhibit combined characteristics of both an autoregressive process and a moving average error process (Pfaff, 2008; Alexander, 2008). In such cases, the resultant model is called an autoregressive moving average (ARMA henceforth) model. If such a process has  $p$  autoregressive terms and  $q$  moving average terms then the model is denoted ARMA( $p, q$ ) and is defined as

$$y_t = \mu + \omega_1 y_{t-1} + \dots + \omega_p y_{t-p} + \vartheta_0 \varepsilon_t + \vartheta_1 \varepsilon_{t-1} + \dots + \vartheta_q \varepsilon_{t-q}, \quad \varepsilon_t \sim \text{i.i.d.}(0, \sigma^2). \quad (7)$$

Finally, each subsequent time period between observations is called a time lag. The number of lags is determined by minimizing various information criteria such as the Akaike Information Criterion (AIC), Schwarz-Bayesian Information Criterion (SBIC), the Hannan-Quinn information criterion (HQC) or the final prediction error (FPE) (Akaike, 1970; Schwarz, 1978; Hannan and Quinn, 1979 respectively).

If a statistically and economically sound model is to be interpreted as having generated the sample data, the model must be stable and be unaffected by time (Pokorny, 1987). To achieve this, the mean and variance of the sample data must be time invariant, that is, the sample data must be stationary and should not reflect heteroscedasticity. Therefore, model (7) is

stationary if and only if the coefficients of the moving average terms are finite and the roots of the characteristic equation of (4) lie within the unit circle (Alexander, 2008).

### 3.1 Unit root tests, stationarity and structural change

We explore the stationarity of a process by considering the process' unit roots. This is because a unit root process is non-stationary. For example, using regression (1) as a template, the  $E[y_t] = 0$  and  $Var[y_t] = t\sigma^2$  implies that though the expected value of the price data is constant the variance of the process grows with time (Pfaff, 2008). The importance of discerning the stationary nature of a particular time series is that it gives helpful information when it comes to cointegration and model specification that involves that particular variable of interest (Herlemont, 2004). Since the stationarity of a series is determined through testing for the presence of unit roots in the series of choice, we proceed to discuss the methods and testing procedures we utilize in this research to determine the stationarity of each of our five data series. These tests are conducted with and without considering structural shifts to showcase some of the differences in the two test cases and the consequences of ignoring structural breaks in data that are found to have breaks. With respect to our data, the testing methods outlined in this section will show how we intend to address the following questions;

- i) Do the series exhibit unit roots?
- ii) In what order are the series integrated?
- iii) Do the series exhibit significant structural breaks?
- iv) How does accounting for these structural breaks affect the stationarity of the series?
- v) What dangers can be faced in testing for stationarity without taking the breaks into consideration?

### 3.1.1 Unit root tests assuming no structural change

Consider the AR(1) model (3) given above. The path of the given process is dependent on the value of  $\omega$  such that if  $|\omega| \geq 1$  then  $y_t$  is said to be non-stationary as system shocks accumulate over time and the process grows unbounded. On the other hand, if  $|\omega| < 1$  then  $y_t$  is stationary. However, when  $|\omega| = 1$  then  $y_t$  has a unit root; this is characteristic of a random walk process.

If regression (3) is found to be non-stationary, then the process can be transformed to a stationary series by conducting the following manipulation;

$$y_t - y_{t-1} = \omega y_{t-1} - y_{t-1} + \varepsilon_t. \quad (8)$$

(Please note equation (8) omits the mean from equation (3) for simplicity and clarity of the explanation to follow; no ambiguity is suffered as a result.)

Equation (8) can then be represented as:

$$\begin{aligned} \Delta y_t &= (\omega - 1)y_{t-1} + \varepsilon_t \\ &= \delta y_{t-1} + \varepsilon_t \end{aligned} \quad (9)$$

where  $\Delta y_t = y_t - y_{t-1}$  and  $\Delta$  denotes the first difference. Understanding differenced data is akin to understanding how the data changes over time. Either equation (3) or (9) can be tested for unit roots by checking whether  $|\omega| = 1$  in (3) or that  $\delta = 0$  in (9). Therefore, both cases can be simultaneously accounted for by testing the following hypotheses,

$$H_0 : \delta = 0 \text{ (unit root)} \quad (10a)$$

and

$$H_1 : \delta < 0 \text{ (stationary)}. \quad (10b)$$

If it is found that  $\delta = 0$  then regression (9) becomes,

$$\Delta y_t = \varepsilon_t. \quad (11)$$

This implies that the first difference of the random walk process,  $y_t$ , is stationary because  $\Delta y_t$  is defined in terms of  $\varepsilon_t$  which is known to be purely random and stationary, denoted I(0). Since  $y_t$  had to be differenced only once to realize stationarity in (9), the series is said to be an integrated process of order one, denoted I(1).

In order to properly characterize the trend properties of the data, three possible trend cases of the test regression can be considered when testing for unit roots,

*Case 1:* A random walk process with drift,

$$\Delta y_t = \mu + \delta y_{t-1} + \varepsilon_t. \quad (12)$$

The hypotheses to be tested are:

$$H_0 : \delta = 0 \text{ (unit root without drift, } \Rightarrow y_t \sim I(1)) \quad (13a)$$

$$H_1 : \delta < 0 \text{ (stationary with non-zero mean, } \Rightarrow y_t \sim I(0)). \quad (13b)$$

*Case 2:* A random walk process with drift around a stochastic trend

$$\Delta y_t = \mu + \eta t + \delta y_{t-1} + \varepsilon_t. \quad (14)$$

The hypotheses to be tested are:

$$H_0 : \delta = 0 \text{ (Unit root with drift, } \Rightarrow y_t \sim I(1)) \quad (15a)$$

$$H_1 : \delta < 0 \text{ (stationary with deterministic time trend, } \Rightarrow y_t \sim I(0)). \quad (15b)$$

The testing procedure valid to test *Cases 1* and *2* is the Dickey-Fuller (DF) unit root test (Dickey and Fuller, 1981). However, this test has the disadvantage that it overlooks the possibility of serial correlation in the error terms. This may lead to biased conclusions. Therefore, to overcome this short-fall, Dickey and Fuller (1981) suggest the replacement of the AR(1) processes in the first two cases with an ARMA(p, q) process so that the assumption of white noise error terms is not violated. The improved test regression is called the augmented Dickey-Fuller (1981) test (ADF henceforth).

*Case 3:* The augmented Dickey-Fuller test

The ADF test regression is defined as follows:

$$\Delta y_t = \mu + \eta t + \delta y_{t-1} + \theta_i \sum_{i=1}^p \Delta y_{t-i} + \varepsilon_t \quad (16)$$

Where  $p$  is selected so that serial correlation in the error terms is removed. *Case 3* is then used to test the null hypothesis that  $\delta = 0$  against the alternative that  $\delta < 0$  using the following test statistic:

$$\tau = \frac{\hat{\delta}}{s.e.(\hat{\delta})} \quad (17)$$

where  $\hat{\delta}$  is the least squares estimate of  $\delta$  and  $s.e.(\hat{\delta})$  is the usual estimate of the coefficient standard error. The critical values used are taken from Dickey and Fuller (1981) and Hamilton (1994). We choose to use the ADF test because it is relatively easy to implement and its results can be easily interpreted.

The number of lagged regressors,  $p$ , is empirically determined by adopting the general-to-specific method. This method allows for an arbitrary lag value,  $p$  max, to be chosen and set as an upper bound so that this value is scaled down by dropping the last insignificant regressor from the test regression (16). The final test equation would then only possess significant endogenously lagged regressors. This selection process is refined by using it jointly with the AIC (Akaike, 1970) and the SBIC (Schwarz, 1978).

To complement the ADF test results, we chose as a second unit root test the Kwiatkowski, Phillips, Schmidt and Shin (1992) test.

The Kwiatkowski, Phillips, Schmidt and Shin (1992) test, abbreviated KPSS henceforth, is a Lagrange Multiplier (LM) test which tests trend and/or difference stationarity. This is a conservative-type test as Pfaff (2008) presents it, which simply means that contrary to most other unit root tests, including the ADF test, the KPSS test has the null hypothesis as a stationary process and the alternative hypothesis as the unit root. According to Pfaff (2008), rejecting the null hypothesis in this case will most assuredly suggest that the series investigated has a unit root. This is to say that where the ADF test favours a unit root while the KPSS test favours a stationary process, results from the latter are used for the final conclusion (Pfaff, 2008). This is the guideline we use where the ADF and KPSS test results conflict.

Kwiatkowski et al. derive their test by considering the following model:

$$y_t = \psi t + v_t + \varepsilon_t, \quad (20a)$$

$$v_t = v_{t-1} + u_t \quad (20b)$$

where  $t$  is the deterministic trend with coefficient  $\psi$  and  $v_t$  is a random walk process. The error process  $\varepsilon_t$  is a  $I(0)$  process from equation (20a) and  $u_t$  is the error process of equation (20b) assumed to be i.i.d.  $(0, \sigma_u^2)$ . The initial value  $v_0$  is a constant corresponding to the level/intercept. Under the null hypothesis, the stationarity of  $y_t$  is investigated by setting  $H_0 : \sigma_u^2 = 0$ . This implies that  $v_t$  is constant. The alternative hypothesis is set as  $\sigma_u^2 > 0$  and the test statistic is defined as follows:

$$LM = \frac{\sum_{i=1}^T S_i^2}{\hat{\sigma}_\varepsilon^2}. \quad (21)$$

The term  $\hat{\sigma}_\varepsilon^2$  is the long-run error variance estimate from the regression of  $y_t$  on a constant and time, and  $S_t$  is the partial sums of the residuals  $\hat{\varepsilon}_t$  from the same regression. Kwiatkowski et al. (1992) estimate the long-run variance by the following formula,

$$s^2(k) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2 + 2T^{-1} \sum_{s=1}^k w(s,k) \sum_{t=s+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} \quad (22)$$

where  $w(s,k)$  is an optimal weighting function. The authors use the Barlett window as in Newy and West (1987) given by  $w(s,k) = 1 - s/k + 1$  so that  $s^2(k)$  is non-negative. This ensures that  $\hat{\sigma}_\varepsilon^2$  is also non-negative as it is estimated by

$$\hat{\sigma}_\varepsilon^2 = s^2(k) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2 + 2T^{-1} \sum_{s=1}^k \left(1 - \frac{s}{k+1}\right) \sum_{t=s+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}. \quad (23)$$

The partial sums of the residuals  $S_t$  is calculated as

$$S_t = \sum_{i=1}^t \hat{\varepsilon}_i, \quad t = 1, 2, \dots, T. \quad (24)$$

The critical values for the KPSS test can be found in Kwiatkowski et al. (1992).

In trading, when a pair of stocks share a constant price ratio then gains can be realized when the stock prices deviate from their equilibrium means. This can be attributed to the fact that the trading strategy implemented in such a case will be based on the idea that if the stocks share historically similar trading patterns then this will be so also in future (Herlemont, 2004). What makes the strategy successful is the mean reversion property. We give a brief discussion of the mean reversion property next. The discussion is mainly based on the works of Alexander (2008), Exley et al. (2004) and various other texts covering the same subject.

The reason stationary processes are so attractive in asset investment is that they exhibit the mean reversion property. This property simply says that when stock prices increase (decrease) to their maximum (minimum) level then the market forces them down (up)

towards an equilibrium constant. It then becomes intuitively clearer to define the return series of stock  $y$  at period time  $t$  as  $\Delta y_t$ , so that an AR (1) process, equation (12) for example, can be thought of as a discretized version of the Ornstein-Uhlenbeck (1930) process. The Ornstein-Uhlenbeck (1930) process is a continuous time model of a stochastic variable based on a mean-reverting diffusion of the form

$$dy(t) = \varphi(\theta - y(t))dt + \sigma y(t)dz \quad (25)$$

where  $y(t)$  is a logarithmic random walk process based on the stock prices. The parameter  $\theta$  is the equilibrium level which is the mean price to which the stock prices revert to in the long-run,  $\varphi$  is the rate at which the prices mean-revert,  $\sigma$  is the time dependent volatility and  $dz$  is a standard Brownian motion.

The process (25) is mean-reverting in the absence of stochastic fluctuations. This is shown by modelling (25) as a first-order differential equation given by

$$\frac{dy(t)}{dt} = \varphi(\theta - y(t)). \quad (26)$$

Solving (26) gives

$$y(t) = y(0)e^{-\varphi t} + \theta \quad (27)$$

where the process is subject to the initial condition  $y(0)$  and  $\varphi > 0$ . As time increases, the process in (27) decays exponentially at a rate of  $\varphi$ . Eventually  $y(t)$  settles on a constant  $\theta$ . Therefore, the process given in (11) has the most rapid mean-reversion (Alexander, 2008).

Methods used in determining the stationarity of a process may be biased by the presence of structural breaks in the data. Perron (2006) cautions that sometimes failure to reject the unit root hypothesis may be due to the presence of significant structural changes in the trend component of the test function. Therefore, in the following section, we look at unit root testing procedures that account for breaks in the data.

### 3.1.2 Unit root tests accounting for structural change

Consider testing the stability of a standard regression model given as:

$$y_l = x_l^T \phi_l + \varepsilon_l, \quad (1 \leq l \leq n) \quad (28)$$

with  $y_l$  being the observation of the dependent variable at time  $l$  influenced by  $x_l$ . The latter is a vector of observations of the independent variables with the first entry equal to one. The disturbance is  $\varepsilon_l$  and  $\phi_l$  is a  $k$ -dimensional vector of regression coefficients.

The null and alternative hypothesis of no structural shift will be given by

$$H_0 : \phi_l = \phi_0 \quad (29a)$$

$$H_a : \phi_l \neq \phi_0. \quad (29b)$$



Testing the null hypothesis given in (29a) addresses the question of whether or not the data exhibit any significant structural breaks. Next, we identify the respective break dates of the most significant breaks in each series. One reason for this approach is that we can test for stationarity of each series around these breaks. The second reason is that we can get a sense of any common event(s) that may have occurred and afflicted the series around the same time period. However, though we test and show that each series exhibits potentially multiple significant breaks, we do not investigate these further for lack of time dedicated to this research. We therefore, chose to limit our investigation to only one significant structural break in each series without serious consequence to the goals set for this research.

The null hypothesis given in (28) is tested by applying a supremum  $F$  test (sup  $F$  henceforth) to the data (Zeileis, 2001). The null hypothesis of no structural shift in the average stock price is tested against the alternative that there exists a significant break. The sup  $F$  test statistic is defined as

$$\sup F = \sup_{t1 \leq z \leq t2} F_z \quad (30)$$

where  $t1$  and  $t2$  define a closed interval of potential break points such that  $j \leq t1 \leq z \leq t2 \leq T - j$ . For a potential breakpoint,  $z$ , model segments can be estimated by ordinary least squares of the data before and after the break point, that is,  $\phi^1$  and  $\phi^2$  respectively. The resulting residuals  $\hat{r} = (\hat{\varepsilon}^1 \hat{\varepsilon}^2)^T$  can then be compared to the model residuals ( $\hat{\varepsilon}$ ) through the  $F$  statistic given by

$$F_z = \frac{(\hat{\varepsilon}^T \hat{\varepsilon} - \hat{r}^T \hat{r}) \setminus j}{\hat{r}^T \hat{r} \setminus (n - 2j)} \quad (31)$$

with  $j$  and  $(n - 2j)$  degrees of freedom (Chow, 1960). The largest value from all possibilities calculated in (27) gives the  $\sup F$  (Zeileis, 2001). Then the null hypothesis is rejected if  $\sup F$  is greater than some critical value tabulated in Andrews (1993).

The alternative of one single shift about point  $z$  is formulated as

$$\phi_t = \begin{cases} \phi^1, & (1 \leq t \leq z) \\ \phi^2, & (z \leq t \leq T), \end{cases} \quad (32)$$

where  $j \leq z \leq T - j$ .

On the other hand, to test the alternative hypothesis of multiple breaks, we use the method of cumulative sums (CUSUM) of the recursive errors (Brown et al., 1975). This method is denoted Rec-CUSUM henceforth and is given by

$$W_n(t) = \frac{1}{\tilde{\sigma} \sqrt{\gamma}} \sum_{l=j+1}^{j+\lfloor t\gamma \rfloor} \tilde{\varepsilon}_l, \quad 0 \leq t \leq 1 \quad (33)$$

where  $\tilde{\varepsilon}_l$  is the  $l$ -th recursive error term,  $\tilde{\sigma}$  is the corresponding estimate of the standard deviation and  $\gamma = n - j$  is the number of recursive residuals.

Under the null hypothesis, the limiting process for  $W_n(t)$  is the standard Wiener process  $W(t)$  and under the alternative hypothesis the residuals will have zero mean up to point  $z$ , a structural break point, where after the process will deviate from its mean.

In both the sup  $F$  test and the Rec-CUSUM test, each of the five stock price random variables is modeled according to equation (28) with  $x_l = 1$  for all  $l$ . This allows each series to be tested for any significant changes in their average market prices over time.

To test for unit roots while accounting for breaks endogenously, we use the Andrew and Zivot (1992) unit root test. We abbreviate this test AZ henceforth.

Andrew and Zivot (1992) suggest the use of three possible models in this test procedure. The first model, A, allows for a one time change in the mean, model B captures a shift in the trend function and model C allows for a change in both the level and the slope of the trend function. The models are given below:

$$\text{Model A:} \quad \Delta y_t = \mu + \eta t + \gamma DU_t + \delta y_{t-1} + \theta_i \sum_{i=1}^p \Delta y_{t-i} + \varepsilon_t \quad (34a)$$

$$\text{Model B:} \quad \Delta y_t = \mu + \eta t + \varphi DT_t + \delta y_{t-1} + \theta_i \sum_{i=1}^p \Delta y_{t-i} + \varepsilon_t \quad (34b)$$

$$\text{Model C:} \quad \Delta y_t = \mu + \eta t + \gamma DU_t + \varphi DT_t + \delta y_{t-1} + \theta_i \sum_{i=1}^p \Delta y_{t-i} + \varepsilon_t \quad (34c)$$

where  $DU_t$  is the shift dummy variable to capture the shift in the mean at every possible break point and  $DT_t$  is the shift dummy variable to capture the shift in the trend function.

Perron (1989) says that most economic series are adequately modeled by either model A or C. Therefore, we opt to use model A as our test equation. This choice is motivated by that

since we do not presume stock prices to be predictable, then the need to consider a break in the trend function is eliminated.

The tested null hypothesis is that the series  $\{y_t\}$  has a unit root with a structural break, that is,  $\delta = 0$  and the alternative hypothesis says that if  $\delta < 0$  then the process is broken and trend stationary. When the test statistic is larger than some critical value, then the null hypothesis is rejected. The critical values are taken from Andrew and Zivot (1992).

We use the results of the AZ test and the differenced data to estimate the common break date in the data. The identification of the common break point and other significant common events to the series facilitate the creation of appropriate dummy variables to include in the cointegration analysis.

In November 2007, a structural shift is observed in all series and the proper shift dummy variable to capture its effects is given by

$$D_{yyyy(mm)} = \begin{cases} 1 & t \geq yyyy : mm : dd(z) \\ 0 & t < yyyy : mm : dd(z). \end{cases} \quad (35)$$

In September 2004, an event with temporal interference to the price movements is also recorded and is properly captured by an impulse variable defined as

$$D_{yyyy(mm)} = \begin{cases} 1 & t = yyyy : mm : dd \\ 0 & otherwise \end{cases} \quad (36)$$

Where  $z$  denotes either November 2007 or September 2004 in (35) and (36) respectively. The year and month when the significant event occurred is denoted  $D_{yyyy(mm)}$  and  $yyyy : mm : dd$  denotes time as year: month: day and thus  $yyyy : mm : dd(z)$  denotes time with respect to the structural break. The events of September 2004 and November 2007 are further detailed in Chapter 5.

Another set of information that we characterize by dummy variables is that of seasonal variations. We take this step to further stabilize variability in the stock returns thereby reducing the chances of heteroscedastic residuals from the final model. In essence, we try to maintain the efficient market hypothesis by attempting to do away with market predictability suggested by seasonal anomalies in the data (Sah, 2009). We achieve this by incorporating eleven centered seasonal dummy variables for each month of each year. The twelfth month is represented by the constant term included in the final model. A similar approach can also be found in Alexakis and Mavrakis (2008). Using centered seasonal dummy variables ensures that only the mean is shifted and the trend remains unaffected (Johansen, 1988, 1991). The incremental effects of every other month are measured relative to the constant month used as the benchmark. Each seasonal dummy variable is defined as unity for a given month and zero everywhere else.

### 3.2 Cointegration

After determining that the series are at least integrated of the first order we then proceed to check if they share any long-term equilibrium relationship. For this, we employ the technique of cointegration pioneered by Granger (1981).

We define cointegration according to Granger (1981) as follows:

Suppose  $y_t$  and  $x_t$  are both I(1) time series and  $w_t$  is some stationary process, then

$$y_t = \alpha + \beta x_t + w_t \quad (37)$$

where  $\alpha$  and  $\beta$  are parameters. Re-arranging (37) gives

$$y_t - \alpha - \beta x_t = w_t \quad (38)$$

which implies that  $y_t - \alpha - \beta x_t$  is stationary because  $w_t$  is an I(0) process.

If  $\beta$  is unique then  $y_t$  and  $x_t$  are said to be cointegrated and (38) is called a cointegrating regression where  $\beta$  is the cointegrating parameter or cointegrating vector when in the form  $[1, -\beta]$ . Thus the relationship purported by regression (37) is not spurious and any valuable long-term information is retained (Gujarati, 1995). In general, any I (1) variables that can be linearly combined in their levels to produce a I (0) series are said to be cointegrated.

### 3.2.1 Granger representation theorem, VAR and VECM

Now consider how a system of two or more cointegrated variables in a state of disequilibrium manages to adjust itself to attain equilibrium in the long-run. This means that the term that corrects for error becomes quite significant. To illustrate this, Granger and Weiss (1983) first formulated what is now known as the Granger representation theorem of such a model characterized by an inbuilt error correction mechanism. This theorem says that a bivariate system of cointegrated I (1) variables  $y_t$  and  $x_t$  given by

$$\begin{aligned} y_t &= \sum_{i=0}^p \eta_{1i} x_{t-i} + \sum_{i=1}^p \omega_{1i} y_{t-i} + \varepsilon_{1t} \\ x_t &= \sum_{i=0}^p \eta_{2i} x_{t-i} + \sum_{i=1}^p \omega_{2i} y_{t-i} + \varepsilon_{2t} \end{aligned} \quad (39)$$

where  $\varepsilon_{1t}$  and  $\varepsilon_{2t}$  satisfy conditions given in (2). The equations can be represented in differenced form as

$$\begin{aligned} \Delta y_t &= \alpha_1 (y_{t-1} - \beta x_{t-1}) + \sum_{i=0}^{p-1} \eta_{1i}^* \Delta x_{t-i} + \sum_{i=1}^{p-1} \omega_{1i}^* \Delta y_{t-i} + \varepsilon_{1t} \\ \Delta x_t &= \alpha_2 (y_{t-1} - \beta x_{t-1}) + \sum_{i=0}^{p-1} \eta_{2i}^* \Delta x_{t-i} + \sum_{i=1}^{p-1} \omega_{2i}^* \Delta y_{t-i} + \varepsilon_{2t} \end{aligned} \quad (40)$$

where both parameters  $\alpha_1$  and  $\alpha_2$  measure the speed of adjustment given that either one or both parameters are significantly different from zero. The notation and presentation in both systems (39) and (40) are adopted from Engle and Granger (2003). The short and the long-term effects of  $x$  on  $y$  are measured by  $\eta_i^*$  and  $\beta$ . Since  $y_{t-1} - \beta x_{t-1}$  is stationary (see

regression (38)), it can thus be viewed as a measure of the degree of disequilibrium under the supposition that the equilibrium relationship between  $x$  and  $y$  is defined by  $y_{t-1} - \beta x_{t-1} = 0$ .

The system given in (40) is said to be in error correction form. This means that it is a representation of the short-run dynamic relationship between  $x$  and  $y$  which is in disequilibrium at any given time  $t$ . The system has an inbuilt tendency to stabilize itself in the long-run (Engle and Granger, 2003).

Assuming that all variables in regression (40) are endogenous, a vector representation of an auto regression model of the same order  $p$  can be defined as

$$y_t^* = A + \sum_{i=1}^p \Gamma_i y_{t-i}^* + \varphi D_t + u_t^* \quad (41)$$

Regression (41) is called a vector auto regression (VAR) model. This model is then transformed into a vector error correction model (VECM) by including an error correcting mechanism,  $\Pi y_{t-1}^*$ , courtesy of the Granger representation theorem. In addition, in order for the VEC model to validly account for all the data, outliers that can be traced to significant world events need to be included in this model through the use of intervention dummy variables defined in the previous section 4.1. The resulting VEC model is given by

$$\Delta y_t^* = \Pi \begin{bmatrix} y_{t-1}^* \\ D_{yyyy}(mm)_{t-1,z} \\ 1 \end{bmatrix} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i}^* + \sum_{j=0}^{p-1} \theta_j \Delta D_{yyyy}(mm)_{t-1,z} + \varphi D_t + u_t^*, \quad t = 1, \dots, T \quad (42)$$

where  $\Pi$  is a  $(K \times (K+2))$  coefficient matrix of rank  $r$ ,  $K$  is the number of endogenous variables, and the shift dummy variable is denoted  $D_{yyyy}(mm)_{t-1,z}$ . The adjustments parameters are given in a  $5 \times r$  matrix  $\alpha^*$  and the cointegrating vectors are columns of matrix  $\beta^*$ . Where only one cointegrating relation is concerned, the equation is normalized on BOVESPA based on the December, 2007 market capitalization figures. This approach

follows Febrian and Herwany (2007). Vector  $y_t^*$  is a 5 x 1 column of I (1) variables, [ $l_{jse}, l_{sse}, l_{sensex}, l_{micex}, l_{bovespa}$ ] (see Chapter 4 for description of the data) where the stock variables are denoted and defined as:

$l_{jse}$ : Johannesburg Stock Exchange All Share (JSE - in logs)

$l_{sse}$ : Shanghai Stock Exchange (SSE - in logs)

$l_{sensex}$ : SENSEX (in logs)

$l_{rtsi}$ : Russian Trading System Index (RTSI - in logs)

$l_{bovespa}$ : BOVESPA (in logs).

If  $r$ , the number of cointegrating relationships, is defined in the interval  $[0, 5]$ , then  $\beta^{*T} y_t^*$  is stationary. Vector  $D_t$  contains 11 seasonal dummy variables, an impulse dummy variable and the constant. Matrices  $\Gamma_i$  hold parameters for  $i=1 \dots p-1$  and vector  $u_t^*$  is a column of white noise error terms. Similar setups to the one given in equation (42) are used by Lutkepohl (2004) and Alexakis and Mavrakis (2008).

The OLS estimator of  $\beta^*$  is said to be super consistent since Stock (1987) shows that in the presence of cointegration this estimator is not only consistent but also rapidly converges to its true value at a rate of  $T^{-1}$ . Typically, error correction models are used to shed light on the short-run dynamics of a system they characterize and can be used in forecasting. Evidently, the coefficient of the error correction term can be used in testing for cointegration (Banerjee et al., 1996). Also, we use the results from the estimated VECM (42) to check whether the left hand side variable is endogenous or weakly exogenous. According to Febrian and Herwany (2007), it can be implied that a variable is weakly exogenous if the coefficient of the corresponding error correction term is not significantly different from zero, otherwise, the variable is endogenous. Other inferences that can be drawn from the coefficients of the error correction term include the speed and direction of adjustment following a temporal disturbance in the system (Febrian and Herwany, 2007).

In addition, since the long run relations are estimated based on log-transformed values, caution must be exercised when interpreting the elasticities. Augmenting the results with impulse response functions can help to improve the interpretation of the long run equation relating the dependent variable to the explanatory variables (Lutkepohl, 2004). Of course, the cointegrating equation is valid if the variable it is normalized on has a nonzero coefficient. This ensures that the variable is indeed part of the cointegrating space given in (42). If this is true, then the dependent variable, BOVESPA in this research, must be cointegrated with at least one of the other variables (Lutkepohl, 2004).

Regression (42) is adapted from Mavrakis and Alexakis (2008) and Lutkepohl (2004). The function already models I(1) processes that are found to be cointegrated and assigns the number of cointegrating relationships as  $r$ . However, what the model does not show is how the variables are found to be cointegrated and how the count of these cointegrated relationships,  $r$ , is carried out. These questions are addressed and answered by the Likelihood ratio testing procedures developed by Johansen (1988) and Johansen and Juselius (1990). We collectively refer to these procedures as the JJ tests henceforth.

### 3.2.2 Johansen and Juselius' cointegration procedure

The JJ procedures are based on two maximum likelihood ratio (MLR/LR) tests called the Maximal eigenvalue test and the Trace test (Johannsen, 1988, 1991). These tests determine whether or not the I(1) variables are cointegrated and how many such relationships exist. Let  $r$  denote the cointegration rank, then the Maximal eigenvalue test and the Trace test are defined as follows:

#### *a. Maximal eigenvalue test*

The test investigates the null and alternative hypotheses given by

$$H_0 = r \text{ is tested against } H_a = r + 1,$$

the LR statistic is given by

$$J_{\max} = -T \ln(1 - \lambda_{r+1}), \quad r = 0, 1, \dots, n-1. \quad (43)$$

*b. Trace test*

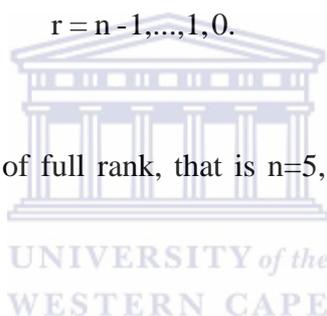
Under this test the null and alternative hypotheses are

$$H_0 = r \text{ is tested against } H_a > r.$$

The Trace test statistic is given by

$$J_{\text{trace}} = -T \sum_{i=r+1}^n \ln(1 - \lambda_i), \quad r = n-1, \dots, 1, 0. \quad (44)$$

If the cointegrating relations are of full rank, that is  $n=5$ , then this would imply that  $y_t$  is trend stationary.



In both cases, (43) and (44),  $T$  is the sample size and  $\lambda_i$  is the  $i$ -th canonical correlation of the  $r$  largest canonical correlations of  $\Delta y_t$  with  $\Delta y_{t-1}$ . This follows from (42) after correcting for any lagged differences and deterministic variables defined by the likelihood estimator of  $\beta$  for a given  $r$  (Österholm and Hjalmarsson, 2007). The critical values are given in Johansen and Juselius (1991).

After realizing the VEC model, we then consider the model's ability to efficiently relate the five stock variables over the considered time frame. Consequently, we wonder if the VEC model's error terms show constant conditional variance over time or if the stock variables are best modeled by Bollerslev's (1986) generalized autoregressive conditional heteroscedastic (GARCH) model. On the other hand, we question whether the standard errors of the estimated regression parameters are within reason or not, that is, are the model error terms uncorrelated or not? The significance of uncorrelated error terms is mainly in providing as accurate t-values as possible for the estimated parameters as this will lead to more reliable

inferences given at a later stage. Finally, we enquire about any systemic errors that the estimated model might exhibit because such errors could result in excessive skewness and kurtosis. We test for all these listed concerns as given in the next section.

### 3.3 Diagnostic testing

After estimating the VECM model (42) (see section 3.2.1), we then check for its adequacy and efficiency by conducting a series of tests applied to the residuals of the model's OLS estimated VAR form. This process and its selection of testing formulae used are adapted mainly for Pfaff (2008).

We investigate the efficiency of the OLS estimates by testing the model for *serial correlation*. This test is based on the premise that the error terms of a correctly specified model are a white noise process. For this, we employ the Breusch-Godfrey (1980) Lagrange multiplier test (BG henceforth). This choice is motivated by the fact that the test accommodates higher order serial correlation. This is a more powerful test compared to, say, the Durbin and Watson (1950, 1951) test which specifies error terms simulated by an AR(1) model in its alternative hypothesis. The tested null hypothesis is that there is no correlation up to lag order  $h$ , that is, the errors are serially independent and the alternative is that the errors are serially correlated.

Based on the VAR ( $p$ ) model given in (41), consider the following auxiliary equation:

$$\hat{u}_t = \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \dots + \rho_h \hat{u}_{t-h} + \varepsilon_t \quad (45)$$

where  $\varepsilon_t$  is a white noise process. The BG tests the following null hypotheses based on regression (45):

$$H_0 : \rho_1 = \dots = \rho_h = 0, \quad (46a)$$

$$H_a : \exists \rho_i \neq 0, \quad i = 0, 1, \dots, h. \quad (46b)$$

The test statistic is defined as

$$LM = T(K - tr(\Xi_R^{-1}\Xi_{UnR})) \quad (47)$$

where  $\Xi_R$  and  $\Xi_{UnR}$  are the covariance matrices of the restricted and unrestricted models respectively. The sample size and the number of endogenous variables are denoted  $T$  and  $K$  respectively. The test statistic is distributed  $\chi^2(hK^2)$  and  $H_0$  is rejected when  $LM$  exceeds the upper chi-square critical value with  $hK^2$  degrees of freedom (Pfaff, 2008).

Next, we test the VAR model for *goodness-of-fit*. The null hypothesis is that the residuals follow a normal distribution. We apply the Jarque-Bera (1980) normality test and skewness and kurtosis tests as well. We abbreviate the test JB henceforth. The JB test statistic is given as follows:

$$JB = \frac{T}{6} \left( s^T s + \frac{1}{4} \left( (k-3)^T (k-3) \right) \right) \quad (48)$$

where  $T$  is the sample size,  $s$  and  $k$  are sample measures of skewness and kurtosis respectively. The JB test follows a  $\chi^2(2K)$  and the skewness and kurtosis tests are distributed  $\chi^2(K)$ , where  $K$  is the number of endogenous variables. Under the assumption for normality the coefficients of skewness and excess kurtosis are 0 and 3 respectively.

Next, we test the constancy of the model's variance by applying the Lagrange multiplier (LM) test proposed by Engle (1982). The LM test is applied to the residuals of the cointegrating relationship to test the null hypothesis of *homoscedasticity*. If constant conditional variance is rejected, then the data is best modeled by Bollerslev's (1986) GARCH model. The null is tested by the test statistic

$$V(q) = \frac{1}{2} TK(K+1) \left( 1 - \frac{2tr(\hat{\Theta}\hat{\Theta}_0^{-1})}{K(K+1)} \right) \quad (49)$$

where  $V(q) \sim \chi^2 \left( q \left( \frac{K(K+1)}{2} \right)^2 \right)$  and  $\hat{\Theta}$  defines the covariance matrix of the regression equation which  $V(q)$  is based on. The latter is given as

$$v(\hat{u}_t \hat{u}_t^T) = c + \alpha_1 (\hat{u}_{t-1} \hat{u}_{t-1}^T) + \dots + \alpha_q (\hat{u}_{t-q} \hat{u}_{t-q}^T) + \varepsilon_t \quad (50)$$

where  $c$  is a vector of constants of dimension  $\frac{K(K+1)}{2}$  and  $\alpha_i$ , for  $i = 1, \dots, q$ , are coefficient matrices of dimensions  $\frac{K(K+1)}{2} \times \frac{K(K+1)}{2}$ . The operator  $v$  stacks columns of symmetric matrices from the main diagonal on downward and  $\varepsilon_t$  is a vector of spherical error processes.

After performing the necessary diagnostics on the estimated VEC model (42), we now focus on the cointegrating equation. The interest we have in this equation is its applicability in investment strategies because of its stationarity. One such strategy is the pairs trading investment technique. The strategy is based on the mean reversion property that is exhibited by stationary processes as discussed before. Hence we proceed to check which two of the five stock variables can be used in the pairs trading strategy.

### 3.4 Statistical arbitrage

The pairs trading strategy is a statistical arbitrage technique implemented mainly to exploit short-run price deviations from a pair of stocks believed to share long-run price equilibrium. This generalized definition of pairs trading gives us the motivation to investigate the trading strategy within our five variable cointegrated system. According to Lutkepohl (2004), since the five stock markets are cointegrated then there must be at least one pair of stocks from the five that is cointegrated. To find the cointegrated pairs, we test through 10, that is  ${}^5C_2$ , possible pairs of stocks. The cointegrated pairs are then taken as candidates in the pairs trading strategy outlined in Herlemont (2004), Gatev et al. (2006), Schmidt (2008) and Triantafyllopoulos and Montana (2011). The execution of the strategy depends on trading signals created to indicate when to open and close market positions. To evaluate the

performance of each traded pair, we calculate and compare the accumulated returns over time. We utilize the “PairTrading” package provided by Takayanagi and Ishikawa (2012) in the R programming language (v2.15.0) to implement the strategy.

Consider a portfolio,  $P$ , consisting of two stocks,  $u$  and  $v$  such that at time  $t$  we can represent such a portfolio as  $P_t = (u_t, v_t)$ . The stock prices  $u_t$  and  $v_t$  are observed at discrete time points  $t = 1, 2, \dots$ . In order to successfully implement the pairs trading strategy involving the stocks in portfolio  $P$  at time  $t$ , then  $P_t$  must be mean reverting. That is, there must exist a process  $w_t$  such that

$$w_t = v_t - \alpha - \beta u_t \quad (51)$$

where  $\beta$  is unique (see regression (37) in section 3.2). The process  $w_t$  is called the price spread of  $P_t$  and is expected to be stationary for the relationship in (51) to be meaningful. We estimate regression (51) by using the JJ cointegration procedure so that we can accommodate both the structural shifts and seasonal changes in the data. The parameters  $\alpha$  and  $\beta$  are called the premium and hedge ratio respectively. The long-run equilibrium value of the portfolio is the resultant sum of the premium and the error term. Since  $E(w_t) = 0$  and  $w_t$  is mean-reverting, then as time progresses, the portfolio fluctuates around the equilibrium value with dynamics determined by  $w_t$ . Given enough time after each temporal deviation from the equilibrium price, the value of the portfolio can always be expected to revert back to the premium. The cointegrating parameter  $\beta$  represents the amount of units of stock  $u$  that are held short at time  $t$  for one unit of  $v$  held long.

We determine the stationarity of the process  $w_t$  by administering the ADF unit root test (Dickey and Fuller, 1981) discussed in section 3.1.1 above and the Phillips and Perron (1988) test. We define the Phillips and Perron (1988) test (PP henceforth) below.

Pfaff (2004) states that because the PP test does not impose the i.i.d assumption in its testing procedure, this leads to the test ignoring any serial correlation exhibited in the error terms of the test regression. The test regression referred to is estimated by equation (37) with error

term,  $u_t$ , that is stationary but not necessarily homoscedastic. Therefore, to correct for any serial correlation and heteroscedasticity in the error term, non-parametric test statistics modified from the DF (Dickey and Fuller, 1981) test statistics are used and given below:

$$Z_\tau = \left( \frac{\hat{\sigma}^2}{\hat{\lambda}^2} \right)^{1/2} \cdot \tau_{\delta=0} - \left( \frac{\hat{\lambda}^2 - \hat{\sigma}^2}{\hat{\lambda}^2} \right) \cdot \left( \frac{n \cdot se(\hat{\delta})}{\hat{\sigma}^2} \right) \quad (52a)$$

$$Z_\delta = n\hat{\delta} - \frac{1}{2} \left( \frac{n^2 \cdot se(\hat{\delta})}{\hat{\sigma}^2} \right) (\hat{\lambda}^2 - \hat{\sigma}^2)$$

where  $\hat{\sigma}^2$  and  $\hat{\lambda}^2$  are consistent estimators of the following variance parameters,

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n E[u_t^2]$$

$$\lambda^2 = \lim_{n \rightarrow \infty} \sum_{t=1}^n E \left[ \frac{U_n^2}{n} \right] \quad (52b)$$



where  $U_n = \sum_{t=1}^n u_t$ . Though under the null hypothesis  $H_0: \delta = 0$ , the PP statistics have the same asymptotic distributions as the ADF statistics, the PP test has the added advantage of being robust to general forms of heteroskedasticity in  $u_t$  and thus does not require specification of the lag length for the test regression.

The trading strategy simply says that if the spread  $w_t$  reaches a given threshold then open a position in the market. As the spread stabilizes again, close the held position. We begin by setting the spread threshold level at 5%, or equivalently, at 2 historical standard deviations. This ensures that our analysis closely follows the example given in Takayanagi and Ishikawa (2012). Thus, when there is a price divergence of 5% or more between the two stock prices then the strategy is implemented.

The implementation of the strategy is based on two basic rules outlined in Herlemont (2004). First, the market entry point is best when the price spread hits the threshold as it begins to

approach its equilibrium state. Herlemont (2004) advises against entering the market when the spread crosses the threshold for the first time. This is because it is not necessarily profitable betting on prices that have a spread that gets continuously wider with time. We also investigate how signals generated by 1% and 10% spread thresholds compare to those generated by a threshold of 5%. The results will indicate which threshold level is optimal in the successful simulation of the trading strategy with respect to our data sample. The second rule is to close the position when the spread reaches its equilibrium state. This is achieved by reversing the long-short positions that were previously held.

To illustrate when to open a position in the market, how long to hold it for and when to close it, we use Herlemont's (2004: p. 8) graphical example given in figure 2a below:

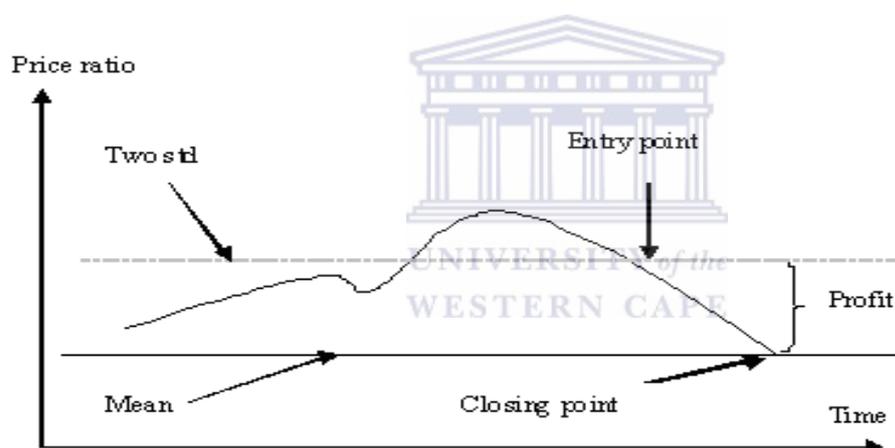


Figure 2a: Pairs Trading Rules.

The example given above is seemingly self-explanatory so we do not give any further details about it from this point onward.

Figure 2b below shows how the trading signals are generated, presented and interpreted. The illustration is taken from Gatev et al. (2006: p. 804) and shows a pair of cointegrated stock prices observed over a six month period. The paired stocks comprise of Kennecott and Uniroyal daily stock data starting from August 1963 to January 1964. The authors also use a threshold of two historical standard deviations to create their trading signals and they represent these by the “up-and-down” dashed lines at the bottom of the graph.

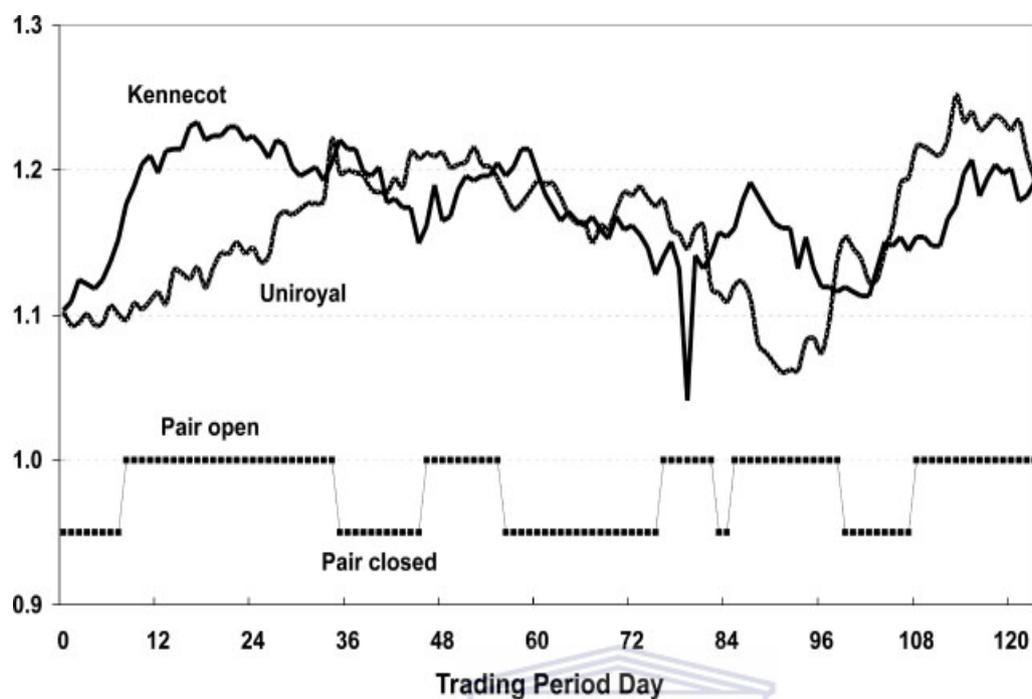


Figure 2b: Trading Signal.

According to figure 2b, the first position was opened around day 8 and remained open until day 36. From day 37, the position is closed until around day 46, then a second consecutive position around day 47 is opened and so on. According to the authors, figure 3b shows five consecutive positions that can be opened during the six month period though not always in the same direction.

Finally, we estimate the accumulated returns over the 9-year time period based on the 1%, 5% and the 10% threshold levels to track the performance of each stock portfolio. Takayanagi and Ishikawa (2012) define the return on the portfolio,  $r_p$ , as

$$r_p = \frac{r_u + r_v \beta}{1 + |\beta|} \quad (53)$$

where  $r_u$  and  $r_v$  are returns on stocks  $u$  and  $v$  which we approximate by differenced logs (see equation (9)).

We determine the risk adjusted performance(s) of our structured portfolio(s) by using the Sharpe Ratio (Sharpe, 1966, 1975). This ratio will basically indicate whether the calculated returns on a portfolio(s) are due to excess risk or sound investment decisions. The ratio is calculated as follows:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}. \quad (54)$$

The parameters  $R_p$  and  $R_f$  are the expected return on the portfolio and the risk free interest rate. The portfolio standard deviation is denoted  $\sigma_p$ .

In conclusion, we realize that from the vast amount of information on the pairs trading strategy, efficiency of the strategy can be improved by using high frequency data. This can be attributed to the ability to promptly execute trades on the precise day a trading signal is generated. However, for this study, we prefer to continue utilizing monthly data. This is because we retained the ability to account for structural shifts and seasonality in the data when testing for the cointegrated pairs. Therefore, we interpret our trading signals on a “month-to-month” basis. For example, a signal can suggest an opening and closing of a position within the same month. Or on the other hand, the interpretation could be to open a position in one month and close it in another. In this way, our interpretation does not reveal the exact daily trades but generalizes over the months of interest. We acknowledge that transaction costs are incurred in pairs trading particularly through portfolio rebalancing. We also acknowledge the negative effect the transaction costs have on the cumulative returns from the trade. However, we do not explicitly calculate any transaction costs in our simulation.

## CHAPTER 4

### Data description

The data we analyze includes monthly closing prices for the JSE All Share index, the SSE index, the BSE SENSEX index, the RTS Index and the BOVESPA stock market index. We acquired the data from [www.yahooofinance.com](http://www.yahooofinance.com) and Plexus Holdings (South Africa). The data sample is made up of 600 observations, that is, 120 observations per dataset. The data covers the sample period from 02:01:2001 to 31:12:2010.

We analyze the stock prices in their local denominations following Febrin and Herwany's (2007) approach. The scholars argue that converting all currencies to the same denomination may result in some complications. Examples include the cross-country interest rate fluctuations suffered under a common currency conversion and the complications brought on by the restrictive assumptions held by the purchasing power parity.

Our sample period covers some very volatile time periods such as the year 2006 when the US economy suffered a mortgage crisis. This crisis later evolved into the subprime mortgage crisis sometime in 2007 which then turned into a global financial catastrophe in 2008. The financial meltdown was characterized by slowed production of consumer goods, increased unemployment rates and sharp drops in commodity prices. In October 2007, the world stock indices had peaked in prices, but only to substantially fall from 2008 onwards. The MSCI Emerging markets recorded a drop of 40.5%, with the largest detractors being the financial sector which returned 61% and the energy sector which returned 53% (Behar and Hest, n.d.). By the end of 2008, the MSCI removed the Pakistani Karachi Stock Index from the emerging market index due to the nation's deteriorating market conditions brought about by the credit crisis.

Below, we give brief summaries of some of the reports based on the BRICS markets.

**Brazil:** After the demutualization in 2007, the Brazilian stock exchange became a for-profit company. By November 2007, the Brazilian Mercantile and Futures Exchange (BM&F) went public which resulted in a cash inflow of BRL 5.98 billion in its opening public offering. In October 2007 the São Paulo Stock Exchange (BOVESPA)

followed suite and raised around BRL 6.6 billion from its public offering. In May 2008, the BM&F and the BOVESPA merged to create the BM&FBOVESPA, the third largest stock exchange in the world (BM&FBOVESPA, n.d.).

Russia: The early signs of the financial crisis in Russia were recorded around May 2008. Later that year, prices of export product fell along with the RTSI and MICEX stock indices. This saw the local banks enter a liquid assets crisis. Around the fourth quarter of 2008, the economy fell into recession and the index recorded a weight decline within the MSCI Emerging Markets Index of more than 50% (Behar and Hest, n.d).

India: The SENSEX recorded a 1 408 point loss on the market in January 2008 following investor unrest about the US recession. In December 2008, the Indian economy witnessed the devaluation of the Rupee (Ahmed , 2008), lowering of the basic interest rates and infusion of state funds into the economy as measures to insulate the economy from the looming global financial crisis. The Mumbai terrorist attacks did not help the economy either. The attacks negatively influenced investor confidence and resulted in a sharp drop in local stock prices. The economy entered into recession around December 2012.

China: By October 2008, the SSE had declined by almost 70%, resulting in a loss of about two-thirds of its value by December 2008 (Yao et al., 2008; Chow, 2010).The economy recorded interest rate cuts for the first time since 2002. In November of the same year, forced closures of factories resulted in critical levels of the unemployment rate. In the same month, a \$586 billion stimulus package was released by the central government as a means to insulate against the impending crisis.

South Africa:

In 2004, the JSE launched the triple bottom line based on the Socially Responsible Investment (SRI) index. The index was the first of its kind in an emerging market and served as a benchmark to South African companies to the benefit of the investors monitoring the behaviours of these companies. As of September 2005, the JSE was the 18<sup>th</sup> largest stock exchange in the world by market capitalization with around 400 companies listed and a market liquidity of 31.2% (SouthAfrica.info, n.d).

## CHAPTER 5

### Data analysis and modelling

The Box-Jenkins (Box and Jenkins, 1976) 3-step modelling approach is assumed in this chapter. These steps include identification, model selection and diagnostic checking. We carry out the first step by analyzing different graphical representations of our data.

#### 5.1 Graphical analysis

The objective of analyzing the time plots is to get some idea about the underlying data generating process governing each of the five series. Understanding these processes will aid in determining what formal tests need to be administered in the subsequent analysis paving way to the main goal of this research.

Three time graphs are given. The first plot is of the data in its original (raw) pricing, the second plot is of the log transformed data and the third plot is of the differenced data. The plots are labeled figures 3, 4 and 5 below respectively, and each plot is augmented by data summaries labeled table 1a, 1b and 1c respectively. All time plots have the monthly prices on the vertical axis and time from January 2001 to December 2010 given on the horizontal axis.

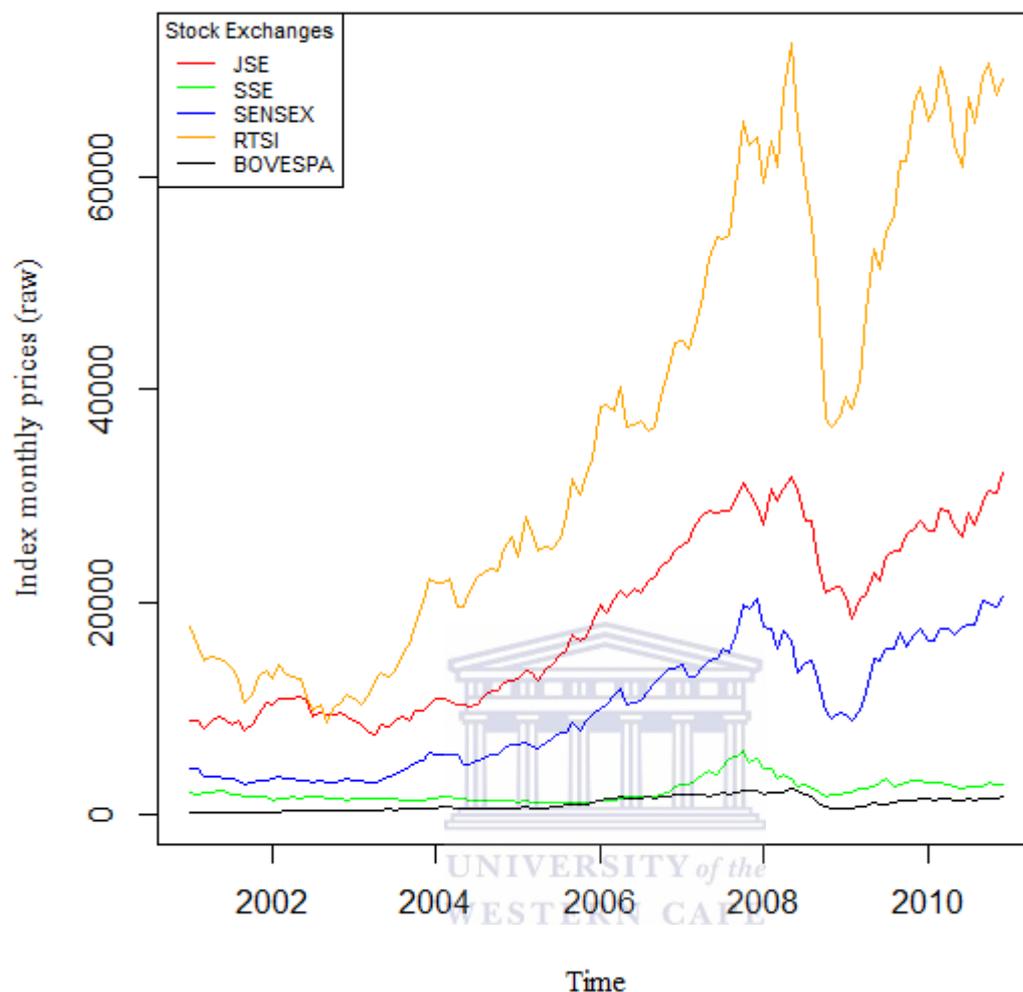


Figure 3: Time series plots of the raw monthly prices.

In figure 3 above, BOVESPA appears to be quite erratic in its progression compared to the other series and RTSI seems very stable by contrast. On the other hand, both SSE and RTSI do not show much trend compared to BOVESPA, SENSEX and JSE. The plots in figure 3 are simply realizations of the true underlying processes. In order to understand the influential forces behind each series, we “smooth” the data by transforming it into logarithms as shown in figure 4 below.

Table 1a: Descriptive statistics for the raw monthly index prices.

Statistic	JSE	SSE	SENSEX	RTSI	BOVESPA
Maximum Value	32123	5955	20509	2459.9	72593
Minimum Value	7510	1061	2812	164.8	8623
Mean	18278.56	2203.67	9745.11	1019.42	35922.33
Standard Deviation	8175.07	1049.97	5657.73	643.91	20358.60
Skewness	0.1743	1.4425	0.3165	0.4127	0.3112
Kurtosis	1.4456	4.8136	1.6663	1.8613	1.6682
Coefficient of Variation (%)	44.72	47.65	58.06	63.16	56.67
Jacque-Bera	12.6892	58.016122	10.8968	9.8888	10.8059
Probability	0.0018*	0.0000*	0.0043*	0.0071*	0.0045*
Number of Observations	120	120	120	120	120

Notes:  $\chi^2_{0.01}(2) = 9.21$  and  $\chi^2_{0.05}(2) = 5.99$ . \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

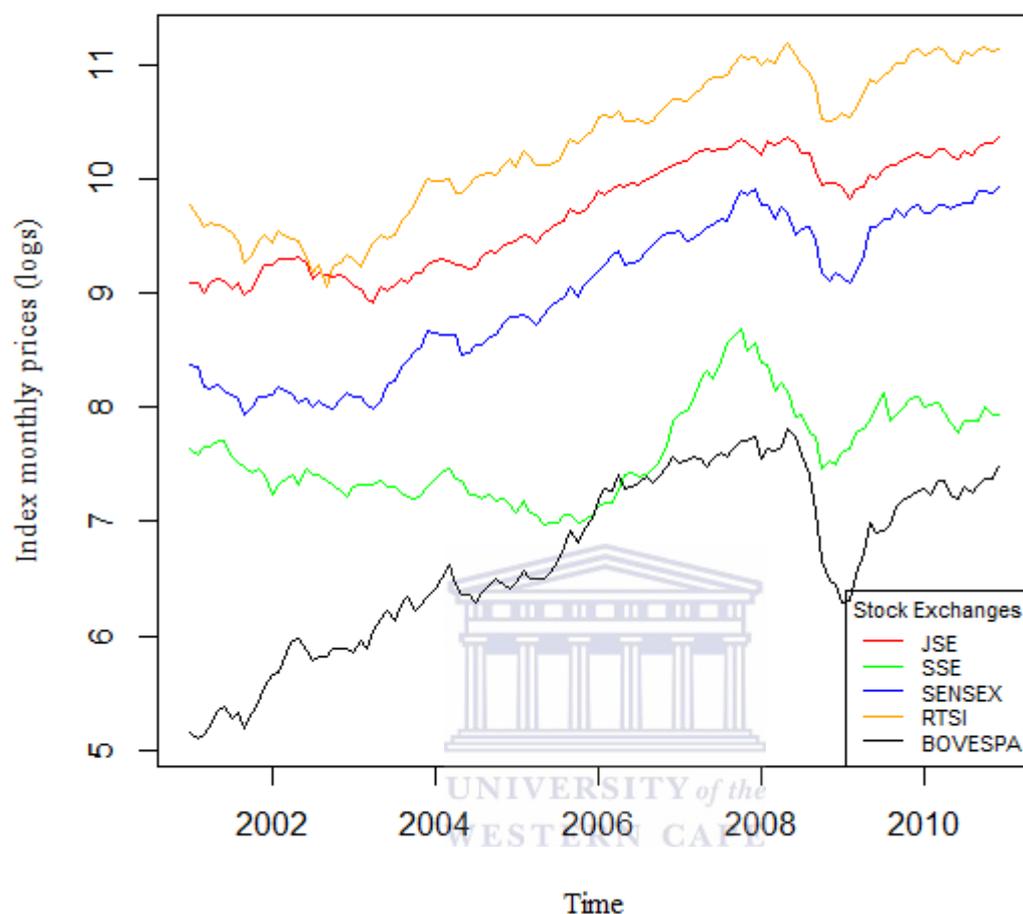


Figure 4: Time series plots of logged monthly prices.

The variability within the prices is clearly stabilized in figure 4 above due to the application of logarithmic transformations to the data. An obvious decline in stock prices around 2008-2009 is observed which can be attributed to the world recession recorded from 2007 to 2009. Also, in 2004, a small, yet very short-lived, price increase in the stock prices was recorded. The stock prices show no signs of reversion to any average prices; this is further indication of the data's non-stationarity.

Table 1b: Descriptive statistics for the log-transformed monthly prices.

Statistic	JSE	SSE	SENSEX	RTSI	BOVESPA
Mean	9.70	7.60	8.99	6.68	10.30
Standard Deviation	0.4798	0.4187	0.6563	0.7516	0.6441
Skewness	-0.0997	0.6151	-0.1776	-0.3939	-0.2457
Kurtosis	1.3922	2.5415	1.5369	2.0568	1.6864
Coefficient of Variation (%)	4.94	5.51	7.30	11.25	6.25
Jacque-Bera	13.1242	8.6173	11.3334	7.5514	9.8357
Probability	0.0014***	0.0135**	0.0035***	0.0229**	0.0073***
Number of Observations	120	120	120	120	120

Notes:  $\chi^2_{0.01}(2)=9.21$  and  $\chi^2_{0.05}(2)=5.99$ . \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

When we consider the absolute values of the skewness statistics, table 1b shows reduced skewness compared to values of table 1a. Consequently, normality is improved though the data is still not normally distributed at 5% significance level as shown by the JB statistic.

We compare the risk-return relationship of investing in one market compared to the other by using the coefficient of variation. This quantity is dimensionless. According to table 1b, the JSE stock is a less risky investment asset as it offers a more attractive risk-return ratio compared to the relatively riskier BOVESPA and SENSEX indices. The SSE borders on average while RTSI shows to be the riskiest investment.

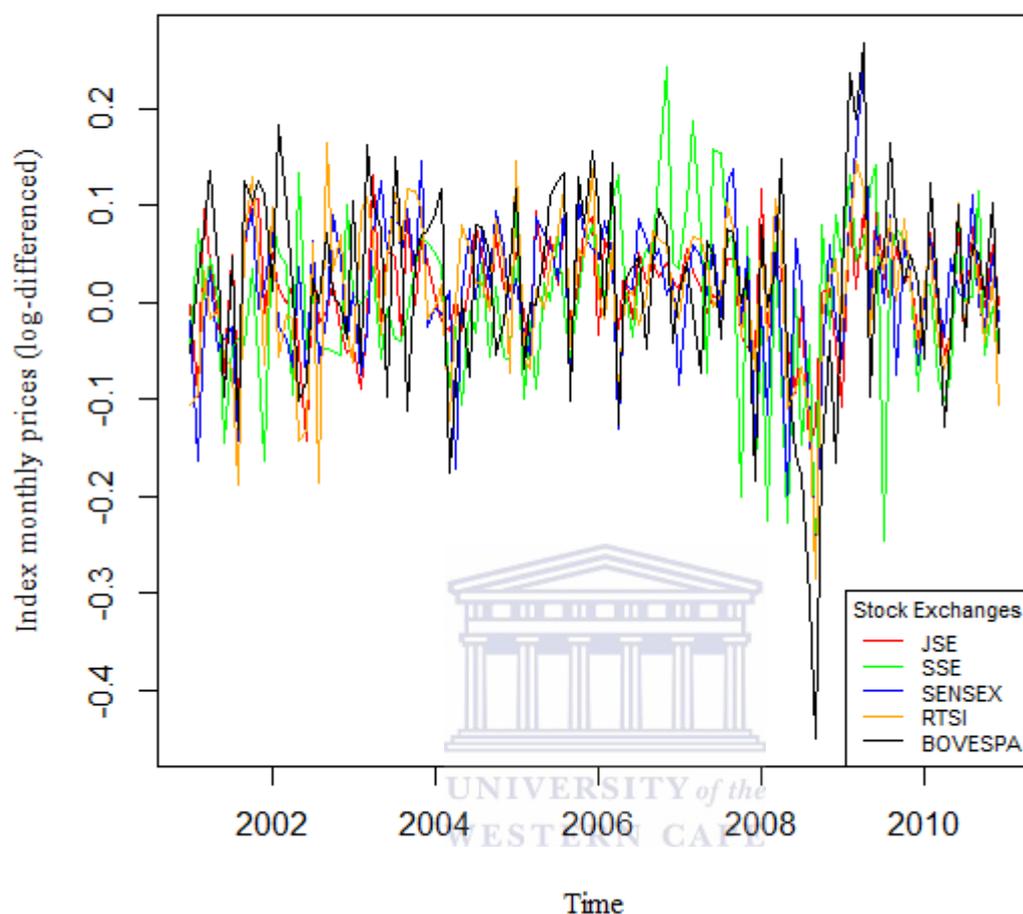


Figure 5: Time series plots of log-differenced monthly prices.

Figure 5 shows the data fluctuating around a mean of zero and the variance relatively stable. According to the conditions given in (2) (see Chapter 3), we argue that the data is non-stationary in its levels until it is log-differenced. Having differenced the data only once to achieve stationarity, we conclude that the data is integrated of order one. These results are formally tested and confirmed in the next section.

Table 1c: Descriptive statistics for the log-differenced monthly prices.

Statistic	JSE	SSE	SENSEX	RTSI	BOVESPA
Mean	0.0107	0.0026	0.0131	0.0195	0.0115
Standard Deviation	0.0559	0.0900	0.0778	0.1053	0.0771
Skewness	-0.4084	-0.5327	-0.5961	-0.9827	-0.6849
Kurtosis	3.1651	3.8946	4.4565	5.7023	4.0299
Jacque-Bera	3.5926	9.7997	18.1427	55.8537	13.8319
Probability	0.1659**	0.007448***	0.0001***	0.0000***	0.0010***
Number of Observations	119	119	119	119	119

Notes:  $\chi^2_{0.01}(2)=9.21$  and  $\chi^2_{0.05}(2)=5.99$ . \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

UNIVERSITY of the  
WESTERN CAPE

It is known that the use of log transformed data will render seasonal effects in the data additive. This may lead to improved effectiveness of our seasonal dummy variables (Reade, 2005) and reduce chances of model misspecification (Zhou and Zhou, 2005; Leykam, 2008).

Table 1d: Correlation matrix of the log-differenced monthly prices.

	JSE	SSE	SENSEX	RTSI	BOVESPA
JSE	1.000 (0.0)	0.7016** (0.0)	0.9747** (0.0)	0.9137** (0.0)	0.9634** (0.0)
SSE	0.7016** (0.0)	1.000 (0.0)	0.6714** (0.0)	0.5562** (0.0)	0.6658** (0.0)
SENSEX	0.9747** (0.0)	0.6714** (0.0)	1.000 (0.0)	0.9202** (0.0)	0.9892** (0.0)
RTSI	0.9137** (0.0)	0.5562** (0.0)	0.9202** (0.0)	1.000 (0.0)	0.8960** (0.0)
BOVESPA	0.9634** (0.0)	0.6658** (0.0)	0.9892** (0.0)	0.8960** (0.0)	1.0000 (0.0)

Notes: The values in the given correlation matrix are calculated based on log-differenced stock prices.

\*\* denotes significant correlations at 5% significance level and p-values are given in ().

The correlation matrix above shows strong positive correlation between the BRICS stock prices. This indicates that there is statistical evidence that the BRICS move in tandem with one another as they respond to changing economic and market conditions. These results are depicted in figure 4 above. SENSEX and BOVESPA show strongest correlation followed by SENSEX and JSE, BOVESPA and JSE, SENSEX and RTSI, RTSI and JSE and RTSI and BOVESPA with a correlation of 0.8960. All correlations are significant at a 5% significant level as shown by the given p-values.

These results provide the earliest evidence in this research that a portfolio consisting of only the BRICS major stock indices is not well diversified. This implies that in this portfolio risk cannot be diversified away. Alternatively, we can reduce investment risk by hedging. This is when we short the positively correlated stocks. We achieve the capacity to carry out this operation by providing further statistical evidence showing that SENSEX and RTSI and RTSI

and BOVESPA remain correlated in the long-run. This will be shown later in the analysis since all data must first be tested for unit roots, i.e. stationarity.

The reader is reminded that the main goal of this research is to investigate long-term relationships between the BRICS. Therefore, the implications of the results in table 1c on portfolio diversification are not the key focus of this research but rather, observations noted as our analysis unravels.

## 5.2 Tests for stationarity

As mentioned before, stationarity is investigated by testing the data for the presence of unit roots. The tests are conducted on both logged and differenced values. When the series reveal unit roots, then we conclude that the data is non-stationary otherwise, stationary. The order of integration for each series is determined with the aid of the analysis of the autocorrelation functions.

Two tests are employed; the ADF test (Dickey and Fuller, 1979) and the KPSS test (Kwiatkowski et al., 1992). The hypotheses for each test are

ADF test:

$H_0$ : The log prices contain a unit root with drift.

$H_a$ : The data is stationary with deterministic time trend.

The critical values are taken from Dickey-Fuller (1981) and Hamilton (1994). The optimal lag structure of each series is determined by the general-to-specific method where  $p$  max is arbitrarily chosen as 13 (see case 3, section 3.1.1, Chapter 3).

KPSS test:

$H_0$ : The process is level-stationary with drift.

$H_a$ : The series is non-stationary.

The critical values are provided in Kwiatkowski et al. (1992). The procedure followed by the KPSS test is “in accordance with a conservative testing strategy” says Pfaff (2008: p103).

The test results for the five time series in both levels and in differences are summarized below.

Table 2: Unit root tests in the absence of structural breaks.

Index	ADF Test Statistic	Lags	KPSS Test Statistic	Lags
JSE	-0.5168	1	1.1706***	9
$\Delta$ JSE	-10.6738*	0	0.1164	1
SSE	-1.2436	13	0.6103**	9
$\Delta$ SSE	-3.6364*	12	0.1703	1
SENSEX	-0.3677	1	1.1923***	9
$\Delta$ SENSEX	-9.3652*	0	0.1322	1
RTSI	-1.925	1	1.1204***	8
$\Delta$ RTSI	-7.7944*	0	0.1827	1
BOVESPA	-0.5169	1	1.318***	8
$\Delta$ BOVESPA	-9.3648*	0	0.1473	1

Notes: The ADF and KPSS critical values for the 10%, 5% and 1% significance levels are -2.57, -2.88, -3.46 and 0.347, 0.463, 0.739 respectively. Critical values are provided in Dickey-Fuller (1981), Hamilton (1994) for the ADF test and in Kwiatkowski et al. (1992) for the KPSS test. \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% respectively.

Both tests show that all series are non-stationary in levels but stationary after they are differenced once at 5% significance level. This suggests that all the series are I(1) as assumed by figures 4 and 5. We complement these tests with the ACF and PACF plots given in Appendix A labeled A1-A5. All the ACF plots show slow decay with increasing number of lags. This is also characteristic of non-stationary series. In each PACF plot, a single spike is evident in the first lag. This is indicative of an AR(1) processes.

Next, we test the series for structural breaks.

### 5.3 Tests for structural change

We test for structural uniformity of the data in two test phases;

*Phase 1.* Each series is tested under the null hypothesis of no breaks. The alternative hypotheses are set specific to each applied test. We then re-test for stationarity while accounting for breaks endogenously. This exercise is carried out by using the AZ test (Andrew and Zivot, 1992) (see section 3.1.2, Chapter 3).

*Test Phase 2.* We estimate the break date of the event that caused the break in the data. This is important because the effects of the break, as and when it occurred, have to be accounted for when estimating the final cointegration model. Hence it is necessary that the break date be traced to its causal event.

A brief note to point: we appreciate that some test results will suggest the possibility of multiple breaks. However, in order to determine the significance of these breaks, we propose that perhaps the best approach will be Bai and Perron's (2003) dynamic programming procedure. Since this technique is not fully investigated in this research, we choose to only consider the single most dominant break in each data set for the sake of progress. Perron (2006) suggests that first preference can be given to the large shifts when all other detected breaks are small enough. We use this as support for our decision. As a result, we appreciate the likelihood that the conclusion drawn at the end of this section may vary from future research conducted in the same respect with similar, if not the same, data.

#### *Test Phase 1*

The average stock prices from each series are tested for any significant change between regimes. This is carried out by applying the Recursive-CUSUM test based on the estimates of the errors and the  $\text{sup}F$  test (see section 3.1.2, Chapter 3). Both tests propose a null hypothesis of no structural break. The  $\text{sup}F$  tests for a single shift alternative while the

Recursive-CUSUM approach tests against multiple breaks in its alternative hypothesis. The results are summarized in table 5 below.

Table 3: Structural break tests for each time series.

Index	Recursive-CUSUM Test Statistic	sup $F$ Test Statistic
JSE	4.4872***	780.2488***
SSE	1.3273***	272.9000***
SENSEX	4.6373***	534.2057***
RTSI	5.8757***	278.5730***
BOVESPA	4.4699***	440.9790***

Notes: All p-values for the Recursive-CUSUM test and the sup $F$  test are  $< 2.2e-16$ , sse has a p-value=0.001648 for the Recursive-CUSUM test, \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

Both tests reject the null hypothesis. This suggests that all data series contain structural breaks. The Recursive-CUSUM test suggests that the data may possibly be contaminated by multiple breaks. Then it follows that the sup $F$  test would favour the single shift alternative.

Next, we estimate the break dates through the Andrew-Zivot (1992) test. For each series, the null and alternative hypotheses are:

$H_0$ : The process has a unit root with drift excluding a break in the mean.

$H_a$ : The process is trend stationary with a break in the mean.

The optimal number of lags used in the testing process is determined by observing the most significant t-ratio of the lagged endogenous variable. Hence we choose  $p$  max=13 as an adequate starting point. The minimum t-statistic of the coefficient of the lagged variables defines the test statistic. The test results are given below in table 4.

Table 4: Unit root test in the presence of structural breaks.

Index	Test Statistic	Lags
$\Delta I_{jse}$	-11.2088***	0
$\Delta I_{sse}$	-4.8332**	4
$\Delta I_{sensex}$	-9.9593***	0
$\Delta I_{rtsi}$	-5.4275***	3
$\Delta I_{bovespa}$	-10.1134***	0

Notes: The critical values for the 1%, 5% and 10% significance levels are -5.34, -4.80, -4.58 respectively. \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively. Estimated Break Dates are presented as year(month).

In table 4 above, all indices are transformed to differenced log values and the unit root tests strongly reject the null hypothesis for all five indices at 5% significance level. The series are only stationary after accounting for a break in the mean. Since we do not intend to segment the data in the test for cointegration, we therefore use the results above to conclude that though the series are I(1), they contain significant breaks in levels.

### *Test Phase 2*

In Appendix B, we give individual plots of the differenced data extrapolated from figure 5. The plots are labelled B1 to B5. Each plot exhibits a “blip” around November 2007, which is identified as position 83 on the graphs. We confirm the position of this blip by following the approach used in Lutkepohl et al. (2004). The “blip” in the differenced data corresponds to a structural shift in levels (Mavrakakis and Alexakis, 2008).

The event that occurred prior to the break in 2007 can be attributed to the frenzy that gripped the world markets as the prospect of a global recession became undeniable. In October 2007, the Chinese stock market recorded a high of 6 124.044 points due mainly to the flooding of the speculative investors into the market ahead of the US slipping into its recession in

December 2007. The SENSEX witnessed high price volatility particularly from July 2007 which saw the market undergo huge corrections due to selling by Foreign Institutional Investors following the heavy selling in the international markets which led to a single day 615 point loss in August of 2007. The Russian stock market suffered delayed effect from the crisis as it only seemed to record a decreasing trend in prices around the end of May 2008.

Next, we estimate the VEC model making provisions for the 2007 and 2004 system shocks.

## 5.4 Model specification

### 5.4.1 Specification of dummy variables

We begin by defining the dummy variables to qualitatively describe the effects of the extraordinary world events. The estimated VEC model accommodates the structural shift recorded in 2007 and the transitory effects recorded in 2004. The structural break of 2007 occurred in November and the temporary price increase of 2004 happened in September. We denote the dates of these two occurrences as 2007(11) and 2004(09) respectively. We then identify the permanent effects of the break by the capital letter  $S$  so that the dummy variable capturing the break is denoted by  $S_{2007(11)}$ . On the other hand, we denote the temporal price increase of 2004 by the small letter  $d$  so that the transitory dummy variable describing this effect is denoted by  $d_{2004(09)}$ .

The shift dummy variable assumes the value 0 before the break and 1 from November 2007 onwards. The transitory dummy variable takes on the value 1 for September 2004, -1 for October 2004 and zero everywhere else.

### 5.4.2 Lag length determination

The lag structure of the VAR model is determined through the information criteria defined in Chapter 3. The criteria are minimized based on an initial lag value of 10 chosen arbitrarily. The optimal lag length selected for the underlying VAR model is 1. The results are summarized in table below.

Table 5: Lag length determination.

Lags	Criteria			
	AIC (e+01)	HQ (e+01)	SC (e+01)	FPE (e-12)
1	-2.7264	-2.6966	-2.6528	1.4437
2	-2.7146	-2.6598	-2.5796	1.6294
3	-2.6913	-2.6116	-2.4949	2.0723
4	-2.6791	-2.5745	-2.4212	2.3727
5	-2.6742	-2.5447	-2.3550	2.5472
6	-2.6790	-2.5246	-2.2985	2.5080
7	-2.6603	-2.4816	-2.2184	3.1678
8	-2.6485	-2.4443	-2.1452	3.7994
9	-2.6592	-2.4302	-2.0945	3.7146
10	-2.6723	-2.4183	-2.0463	3.6381
Selected Lags	1	1	1	1

Notes: For each criterion, the number of lags is selected such that the order of the VAR minimizes the respective criterion.

Mavrakis and Alexakis (2008) point out that a 1 lag selection for the VAR model has some negative impact on the properties of the model's residuals. As a result, we opt for a lag length of 2 instead. This decision is based on the equivalence test performed on OLS estimated coefficients of both the 2- and 1-lagged restricted VECMs. In both cases, the VECMs were estimated by the JJ (Johansen and Juselius, 1990) procedure. This way, we avoid increasing the mean-square forecast errors of the underlying VAR structure as Lutkepohl (2004) cautions. The equivalence test mentioned above is not demonstrated in this research; however

the coding procedure is adapted from the R source code found at *R Graphical Manual* (Matthieu, n.d.).

Next we estimate the VEC model and the number of cointegrating equations it has.

### 5.4.3 Rank determination

The cointegrating rank,  $r$ , is determined for two cases; with and without the inclusion of the structural dummy variable. We adopt Lutkepohl's (2004) approach to show the shift dummy's potential to change the asymptotic properties of the test equation and its subsequent influence on the choice of  $r$ . The test equation also includes one transitory dummy variable and 11 seasonal dummy variables. Tables 6 and 7 below give the results of the JJ trace and eigenvalue tests respectively.

For the trace test the hypotheses are:

$$H_0 : r = r_0.$$

$$H_a : r > r_0, \quad r_0 = 0, 1, \dots, 4.$$



For the eigenvalue test the hypotheses are:

$$H_0 : r = r_0.$$

$$H_a : r = r_0 + 1, \quad r_0 = 0, 1, \dots, 4.$$

Table 6: Rank determination ignoring the shift dummy variable.

<u>Trace Test</u>				
<u>Rank</u>	<u>Test Statistic</u>	<u>Critical Values</u>		
		<u>10%</u>	<u>5%</u>	<u>1%</u>
$r \leq 4$	2.50	7.52	9.24	12.97
$r \leq 3$	10.80	17.85	19.96	24.60
$r \leq 2$	23.06	32.00	34.91	41.07
$r \leq 1$	39.37	49.65	53.12	60.16
$r = 0$	73.48*	71.86	76.07	84.45
<u>Maximal Eigenvalue Test</u>				
<u>Rank</u>	<u>Test Statistic</u>	<u>Critical Values</u>		
		<u>10%</u>	<u>5%</u>	<u>1%</u>
$r \leq 4$	2.50	7.52	9.24	12.97
$r \leq 3$	8.30	13.75	15.67	20.20
$r \leq 2$	12.26	19.77	22.00	26.81
$r \leq 1$	16.31	25.56	28.14	33.24
$r = 0$	34.10*	31.66	34.40	39.79

Notes: The critical values are adopted from Johansen and Juselius (1990), table A2,

\*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

In table 6 above the trace test does not reject the null of no cointegration at 5% significance level while the eigenvalue test does.

Table 7: Rank determination including the shift dummy variable.

<u>Trace Test</u>				
<u>Rank</u>	<u>Test Statistic</u>	<u>Critical Values</u>		
		<u>10%</u>	<u>5%</u>	<u>1%</u>
$r \leq 4$	2.64	7.52	9.24	12.97
$r \leq 3$	13.95	17.85	19.96	24.60
$r \leq 2$	29.46	32.00	34.91	41.07
$r \leq 1$	53.88**	49.65	53.12	60.16
$r = 0$	92.40***	71.86	76.07	84.45
<u>Maximal Eigenvalue Test</u>				
<u>Rank</u>	<u>Test Statistic</u>	<u>Critical Values</u>		
		<u>10%</u>	<u>5%</u>	<u>1%</u>
$r \leq 4$	2.64	7.52	9.24	12.97
$r \leq 3$	11.31	13.75	15.67	20.20
$r \leq 2$	15.51	19.77	22.00	26.81
$r \leq 1$	24.42	25.56	28.14	33.24
$r = 0$	38.52**	31.66	34.40	39.79

Notes: The critical values are adopted from Johansen and Juselius (1990), table A2,

\*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

In table 7 above both tests reject the null of no cointegration at 5% significance level. The trace test suggests two cointegrating equations and the eigenvalue test decides on one cointegrating equation. According to Lutkepohl et al. (2000), the two tests show no major differences in local power. On the other hand, the trace test is often preferred because of its robustness to skewness and excess kurtosis (Cheung and Lai, 1993). Hence we chose to proceed with the both rank outcomes and determine from diagnostic testing which of the two models is statistically superior. Based on the results above, we can conclude that the BRICS are not cointegrated when the structural break is ignored. However, at least one long-run equilibrium relationship exists when the shift is accounted for.

#### 5.4.4 VECM estimation

In the case where  $r=1$ , we normalize the long-run relation on the variable with the smallest market capitalization following the approach in Febrian and Herwany (2007). We use market capitalization values for the year ended 31 December 2007 given below in table 9. The variable should also have a non-zero coefficient according to Lutkepohl (2004). We determine this by testing for cointegrated pairs using the JJ approach and give the results in table 8.

In the second case were  $r=2$ , we build on specifics of the first case by normalizing the second cointegrating equation on the second variable in succession with the smallest market and a non-zero coefficient.

Table 8: Testing for cointegrated pairs.

Stock Pairs			
JSE-SSE Trace: $r=0$ Eigenvalue: $r=0$	SSE-SENSEX Trace: $r=0$ Eigenvalue: $r=0$	SENSEX-RTSI Trace: $r=1^{***}$ Eigenvalue: $r=1^{**}$	RTSI-BOVESPA Trace: $r=1^*$ Eigenvalue: $r=1^*$
JSE-SENSEX Trace: $r=0$ Eigenvalue: $r=1^{**}$	SSE-RTSI Trace: $r=0$ Eigenvalue: $r=0$	SENSEX-BOVESPA Trace: $r=0$ Eigenvalue: $r=1^*$	
JSE-RTSI Trace: $r=0$ Eigenvalue: $r=0$	SSE-BOVESPA Trace: $r=0$ Eigenvalue: $r=0$		
JSE-BOVESPA Trace: $r=0$ Eigenvalue: $r=0$			

Notes:

The critical values are adopted from Johansen and Juselius (1990), table A2,

\*\*\*, \*\* and \* denote significance at 1%, 5% and 10% significance levels respectively.

The results given in table 8 above only indicate whether or not there is cointegration between the tested pairs at 5% significance level. We overcome any disparity between the results of the trace and the eigenvalue tests by choosing the results of the former.

Table 9: Market capitalization data as of 31 December, 2007.

	Stock Market Index				
	JSE	SSE	SENSEX	RTSI	BOVESPA
Market Capitalization (Trillion/US\$)	0.828	3.7	1.09	0.196	0.0013

Notes: Data sources: [www.marketswiki.com](http://www.marketswiki.com), [www.wikininvest.com](http://www.wikininvest.com) and [www.rts.ru](http://www.rts.ru)

This research deems data usable if readily available within a six-month period prior to cut-off date, i.e. “as of 31 December, 2007” implies data used is found to be readily available from 30 June, 2007 - 31 December, 2007.

Both tables 8 and 9 above show that the best variable to normalize the first cointegrating equation on is BOVESPA. The second equation is normalized on RTSI. We do not base the normalization of the second equation on any obvious merit except that we simply follow the trend we use for the first equation.

The coefficient and loading matrices are estimated and given in the tables below with the t-statistics given in the brackets.

Table 10a: The cointegrating vector and loading parameters (r=1).

Matrix	JSE	SSE	SENSEX	RTSI	BOVESPA	Constant
$\hat{\beta}^{*T}$	0.0849 [0.43]	0.0636 [0.96]	-1.6347 [-10.95]	0.4446 [5.30]	1.000	0.0224 [0.02]
$\hat{\alpha}^{*T}$	-0.0976 [-2.51]	-0.2154 [-3.51]	0.0143 [0.26]	-0.2646 [-3.89]	-0.1229 [-2.29]	
Multiple R-squared: 0.197, Adjusted R-squared: -0.0081 F-statistic: 0.9606 on 24 and 94 DF, p-value: 0.5235						

Notes: t-statistics are given in [].

The long-run relation given in table 10a can be summarized (to two decimal places) as:

$$l_{bovespa} = -0.02 - 0.08l_{jse} - 0.06l_{sse} + 1.63l_{sensex} - 0.44l_{rtsi} + \varepsilon_t \quad (51)$$

[0.02] [0.43] [0.96] [-10.95] [5.30]

where  $l_{bovespa}$ ,  $l_{jse}$ ,  $l_{sse}$ ,  $l_{sensex}$  and  $l_{rtsi}$  denote log values of the respective stock prices and  $\varepsilon_t$  is the error term.

The parameter estimates in regression (51) measure elasticity. That is, if the JSE, SSE or RTSI stocks appreciate by just 1% in price, then while holding all else constant, the price of the BOVESPA stock will decrease by 0.08%, 0.06% or 0.44% respectively. However, a 1% increase in the price of the Shanghai SENSEX will lead to a 1.63% increase in the BOVESPA stock price.

Based on the given low adjusted R-squared value, it can be argued that among other alternatives, one of two things could be done to improve on the model's performance: 1) other external factors can be included in the model such as more dummy variables to compensate for, say, the omitted structural breaks that prove to be significant. 2) On the other hand, the statistic might be indicating that perhaps some of the variables in equation (51) may serve the study best if they are excluded. In that case, it could be further argued that by

referring to the t-values the given system of stock prices might be largely influenced by the Russian and Indian markets as compared to the JSE and SSE indices.

This can be explained by noting that South Africa only became a full member of the BRICS group in 2011. As a result it can be hypothesised that in time, the inclusion of this state in bi- and multilateral arrangements and other such acts, could see South Africa gaining considerable influence within the group. In China's case, perhaps a definitive stand on the Yuan/Dollar peg should be addressed. If the Yuan is allowed to appreciate against the dollar, this might bring to pass some of the fears and expectations surrounding this issue on an international level. In addition, if China agrees to import not only commodities but also value added goods from the other group members, then its influence within the BRICS could be notably significant.

The RTSI's high t-value from regression (51) can be thought of as a result of Russia's abundance of natural gas and mineral reserves. In particular, the country has the largest oil deposits compared to the other member states. This wealth of natural resources can explain the country's relatively significant influence on the five variable system (Ministry of Finance, 2012). On the other hand, Brazil and India have a shared history dating back to the Portuguese Empire. This relationship has considerably developed into the India-Brazil-South Africa (IBSA) dialogue forum over the time. Co-operation, particularly between India and Brazil, has included such areas as the permanent participation of developing states in the United Nations Security Council (UNSC), the reform of the United Nations (UN), also areas in science and technology, international trade, space programs, etcera ("Brasilia Declaration", 2003).

Considering the p-value of 0.5235 for the F-test, future research might reconsider model (51) under a different set circumstances and other influential factors not addressed in the estimation of regression but inferred in the immediate discussion. Alternatively, an entirely different approach might be taken.

The opposite signs in  $\hat{\alpha}^{*T}$  and  $\hat{\beta}^{*T}$  indicate that the equation (51) tends to move towards its long-run equilibrium following temporal deviations caused by shocks to the system.

According to the t-values, the SENSEX and RTSI variables are most significant in regression (51).

Table 10b: The cointegrating vector and loading parameters (r=2).

Matrix	JSE	SSE	SENSEX	RTSI	BOVESPA	Constant
$\hat{\beta}_1^{*T}$	0.2988 [1.86]	-0.1233 [-2.27]	-0.9504 [-7.73]	0.000 [0.00]	1.000	-3.6231 [-4.79]
$\hat{\beta}_2^{*T}$	-0.4812 [-1.17]	0.4203 [2.88]	-1.5392 [-2.53]	1.000	0.000 [0.00]	8.200 [4.03]
$\hat{\alpha}_1^{*T}$	-0.0154 [-0.37]	-0.0592 [-0.91]	0.1223 [2.12]	-0.2739 [-3.68]	-0.1270 [-2.16]	
$\hat{\alpha}_2^{*T}$	-0.0529 [-3.74]	0.1139 [-5.15]	-0.0062 [-0.32]	-0.1166 [-4.60]	-0.054 [-2.70]	
Multiple R-squared: 0.197, Adjusted R-squared: -0.008077						
F-statistic: 0.9606 on 24 and 94 DF, p-value: 0.5235						

Notes: t-statistics are given in [].

The adjustment coefficients  $\hat{\alpha}_1^{*T}$  and  $\hat{\alpha}_2^{*T}$  given in table 10b correspond to the equations (52a) and (52b) below respectively:

$$l_{bovespa} = 3.62 - 0.30l_{jse} + 0.12l_{sse} + 0.95l_{sensex} + \varepsilon_{t1} \quad (52a)$$

[-4.79] [1.86] [-2.27] [-7.73]

$$l_{rtsi} = -8.20 + 0.48l_{jse} - 0.42l_{sse} + 1.54l_{sensex} + \varepsilon_{t2} \quad (52b)$$

[4.03] [-1.17] [2.88] [2.53]

where  $\varepsilon_{t1}$  and  $\varepsilon_{t2}$  are the respective error terms

In order to maintain equations (52a) and (52b) stationary, a 1% increase in the SENSEX or in the SSE stock prices has to yield 0.12% or 0.95% increase in the BOVESPA. On the other hand, a 1% increase in the JSE stock prices must result in a 0.3% decrease in the BOVESPA stock prices. Similarly, an increase in either the JSE or the SENSEX stock prices by 1% must stimulate a 0.48% increase in the RTSI stock or a 1.54% increase respectively. Almost all the coefficients are significantly different from zero in both equations (52a) and (52b) with the possible exception of the JSE.

#### 5.4.5 Impulse response functions

As discussed in section 3.2.1, the interpretation of a dynamic multivariate system can be improved upon by also considering impulse response functions. As a result, we plot and analyse the response functions given in Appendix C where each figure depicts the response of each variable to an exogenous shock to the system.

For example, in figure C5, when a one-time external shock is introduced to the Brazilian market, the remaining four markets are not significantly affected by this occurrence. We argue this by observing that the other functions seem to remain within two-standard deviations from zero. The error bounds are given as dashed lines around the functions. Most noticeably, is the seemingly lasting effect that the shock appears to have on the Brazilian stock market itself. This behaviour is due to the unit root present in the system, consequently, the response function does not return to zero. Lastly, each function is analysed over 12 periods representing the monthly data used in this research. Thus, figure C5 provides the evidence for supporting the endogeneity of the variable BOVESPA.

In the same manner, we argue that SENSEX is truly exogenous to the system. This can be seen from figure C3. A shock to the Indian market appears to be transmitted to the entire system as all stock variables are seen to be significantly affected by this event. The BOVESPA response function is positively affected as expected from equation (51), however, the unit shock does not actually effect a 1.63% change in the BOVESPA stock prices over the observed 12 periods. This goes to show some of the shortcomings of solely basing long-run relations on elasticities alone.

Unit shocks to the other markets appear to have significant effects on the Brazilian market. It is possible that the lack of detectable effect in BOVESPA from a shock in the Shanghai market can be due to the SSE also being an endogenous variable. This would imply that the relation between these two variables can be through some third factor, for example, any of the remaining three stock variables in the system.

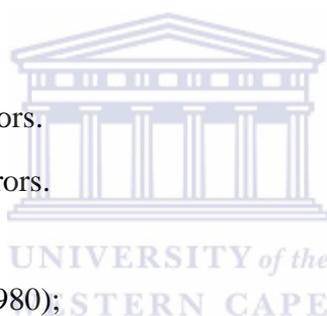
## 5.5 Model misspecification tests

We test for model misspecification by checking if the model's errors are normally distributed or that they show any ARCH effects or serial correlation. All tests carried out are discussed in section 3.3, Chapter 3. The hypotheses for each test are summarized below:

LM ARCH test (Engle, 1982);

$H_0$ : Homoscedastic errors.

$H_a$ : Heteroscedastic errors.



JB normality test (Jacque-Bera, 1980);

$H_0$ : Error terms normally distributed i.e. skewness = 0 and excess kurtosis = 3.

$H_a$ : Error terms not normally distributed.

BG serial correlation test (Breusch-Godfrey, 1980);

$H_0$ : No serial correlation up to order  $h$ .

$H_a$ : First order first order serial correlation.

The test results are given below:

Table 11a: Model testing (r=1).

Test	Critical Value 5%	Statistic	p-value	Degrees of Freedom
BG	124.34	101.41	0.023**	75
JB	18.31	29.12	0.001*	10
Skewness	11.07	13.14	0.022**	5
Kurtosis	11.07	15.97	0.007*	5
ARCH		1198.92	0.062***	1125

Notes: \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% respectively.

The JB test statistic is distributed  $\chi^2(2K)$ , Skewness and Kurtosis are distributed  $\chi^2(K)$ .

The BG statistic is distributed  $\chi^2(hK^2)$ .

The ARCH test is distributed  $\chi^2(qK^2(K+1)^2/4)$ ,

K=5, h=3, q=5.

At 5% significance level table 11a above suggests that the underlying VAR model has no significant residual autocorrelation up to order 3. Normality is rejected at 5% level of significance probably because of excess kurtosis since the residuals from the VAR do not exhibit skewness. Cited in Alexakis and Mavrakakis (2008), Juselius (2001) points out that violation of non-normality is less serious when caused by kurtosis than when caused by skewness because in the case of the latter estimates can be biased. Lastly, homoscedasticity is strongly rejected at 1% level of significance. This is expected as it was implied in earlier discussions (see Chapter 2, section 2.5) that it is impractical to assume constant variance in stock prices particularly data taken from emerging markets. As a result, a GARCH (1, 1) model may serve best in explaining price volatility in the BRICS.

Table 11b: Model testing (r=2).

Test	Critical Value 5%	Statistic	p-value	Degrees of Freedom
BG	124.34	100.13	0.028**	75
JB	18.31	36.04	0.000*	10
Skewness	11.07	13.09	0.023**	5
Kurtosis	11.07	22.95	0.000*	5
ARCH		1205.048	0.048**	1125

Notes: \*\*\*, \*\* and \* denote significance at 1%, 5% and 10% respectively.

The JB test statistic is distributed  $\chi^2(2K)$ , Skewness and Kurtosis are distributed  $\chi^2(K)$ .

The BG statistic is distributed  $\chi^2(hK^2)$ .

The ARCH test is distributed  $\chi^2(qK^2(K+1)^2/4)$ ,

K=5, h=3, q=5.

Table 11b above shows that when the data is modeled by two long-run equations, the underlying VAR model does not exhibit serial correlation or ARCH effects at 5% significance level. The normality violation is clearly a result of excess kurtosis as the results indicate that skewness is not significant in the model error terms.

From both tables 11a and 11b, we conclude that a VEC model with two cointegrating equations best captures the BRICS long-run equilibrium relationship with the necessary volatile times in the sample period accounted for. However, the results are not without room for improvement as both tables show that excess kurtosis is still prevalent in the model residuals, and hence normality is violated. Although the cointegration tests are robust to ARCH effects (Rahbek et al., 2002), the results still indicate that heteroscedasticity within the error terms is quite significant.

We also visually analyze the residual series from the VAR(2) processes tested in tables 11a and 11b. The series are given in Appendix D where figure D1 and D2 denote residual plots of VAR(2) processes with one and two cointegrating equations respectively. The graphs indicate that both VAR(2) processes are stable and volatility clustering is quite evident. The differences between the graphs are very subtle, but upon careful inspection we do confirm that the former ( $r=1$ ) model exhibits volatility clustering to a lesser extent than the former.

Next, we use the cointegration results obtained above to transition into the framework of implementing a statistical arbitrage trading strategy.

## 5.6 Statistical arbitrage

We investigate statistical arbitrage through the pairs trading technique. As discussed in section 3.4, Chapter 3, pairs trading is highly reliant on the mean reversion property. Therefore, we put the strategy into practice through a series of four steps which use the results from equation (51) and table 10 given above. The steps outlined are combined and tailored from the following literature: Herlemont (2004), Gatev et al. (2006), Schmidt (2008), Triantafyllopoulos and Montana (2011) and Takayanagi and Ishikawa (2012).

*Step 1:* Find a pair of stocks that have historically moved together.

From table 8 above, two stock pairs SENSEX-RTSI and RTSI-BOVESPA show evidence of cointegration between 2001 and 2009. We abbreviate these pairs henceforth as S-R and R-B respectively. The historical price movements of each stock pair are shown in figures 6 and 7.

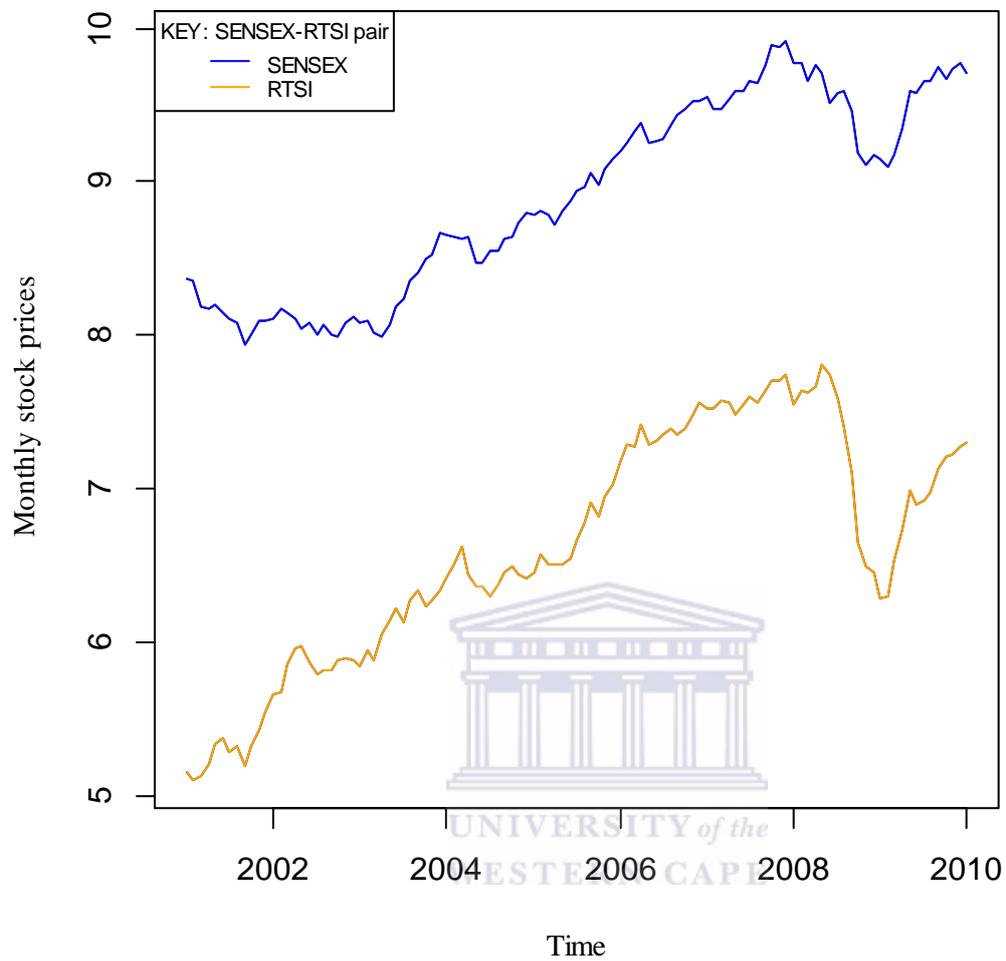


Figure 6: SENSEX and RTSI monthly stock prices.

The SENSEX and RTSI price movements show obvious convergence from the beginning of the sample period until the end of 2003. Between mid-2004 and mid-2005 the prices seem to diverge slightly.

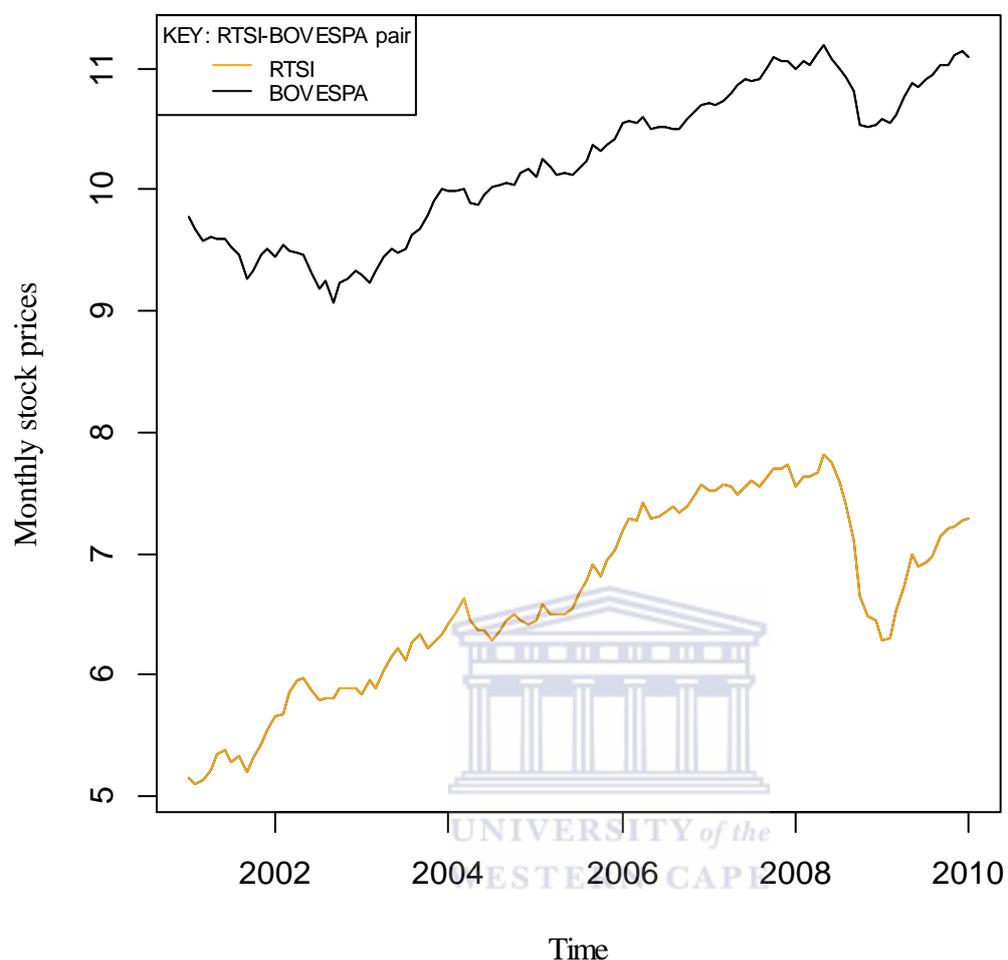


Figure 7: RTSI and BOVESPA monthly stock prices.

Similarly, The RTSI and BOVESPA prices show some convergence from the beginning of the sample period until the end of 2003, and slight divergence between mid-2004 and mid-2005. Other such relations are not easily detectable with a naked eye.

*Step2:* Determine the price spread and test for stationarity.

The price spread is simply the respective error terms from each pair's cointegrating equation. Applying the pairs trading algorithm provided by Takayanagi and Ishikawa (2012), the estimated cointegrating equations are given as:

$$\text{S-R: } l_{rtsi} = -2.79 + 1.05l_{sensex} + w_{S-R} \quad (53a)$$

[-7.5] [25.5]

$$\text{R-B: } l_{bovespa} = 5.17 + 0.77l_{rtsi} + w_{R-B} \quad (53b)$$

[21.95] [21.92]

where  $w_{S-R}$  and  $w_{R-B}$  are the spreads for the S-R cointegrating equation (53a) and R-B cointegrating equation (53b). The t-values are given in the square brackets.

The p-value for the F-test and the adjusted R-squared for equation (53a) are 0.00 and 0.8454 respectively. Similarly, for equation (53b) the adjusted R-squared is 0.8029 and the p-value for the F-test is 0.00. From these statistics, it is safe to assume that the respective data is adequately modeled in both systems given by equations (53a) and (53b). Moreover, the absolute values for the given t-statistics suggest that in both regressions, the coefficients of the independent variables are significantly different from zero. These results follow those given by regression (51) and the relationships between the paired variables can be argued as before (see discussion given in pages 85-86).

In both equations above, normalization is based on market capitalization. The rationale is that indices with large market capitalization are more economically stable compared to the indices with relatively smaller market capitalization. As a result, it can be hypothesized that the direction of causality goes from the former to the latter. Hence using the results in table 9 to determine the role of each variable in each of the regressions (53a) and (53b), we find that premiums are -2.79 and 5.17 and the hedge ratios are 1.05 and 0.77 respectively. The evolution of  $w_{S-R}$  and  $w_{R-B}$  over time is shown in figure 8 below respectively:

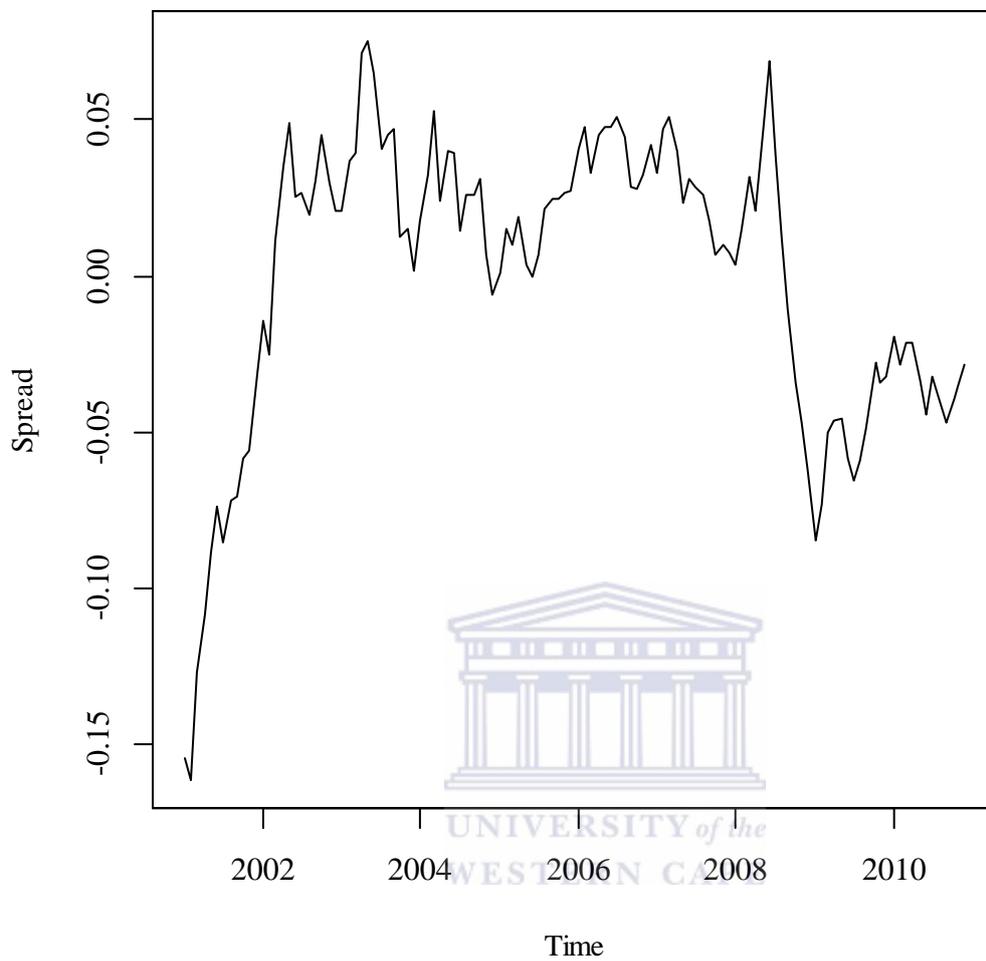


Figure 8: Price spread between the SENSEX and RTSI stocks.

In figure 8 above, the price spread calculated from equation (53a) shows to be slightly stationary between 2002 and 2008. The sudden decline after 2008 can be argued to have been the result of the structural shift in prices discussed in section 5.3.

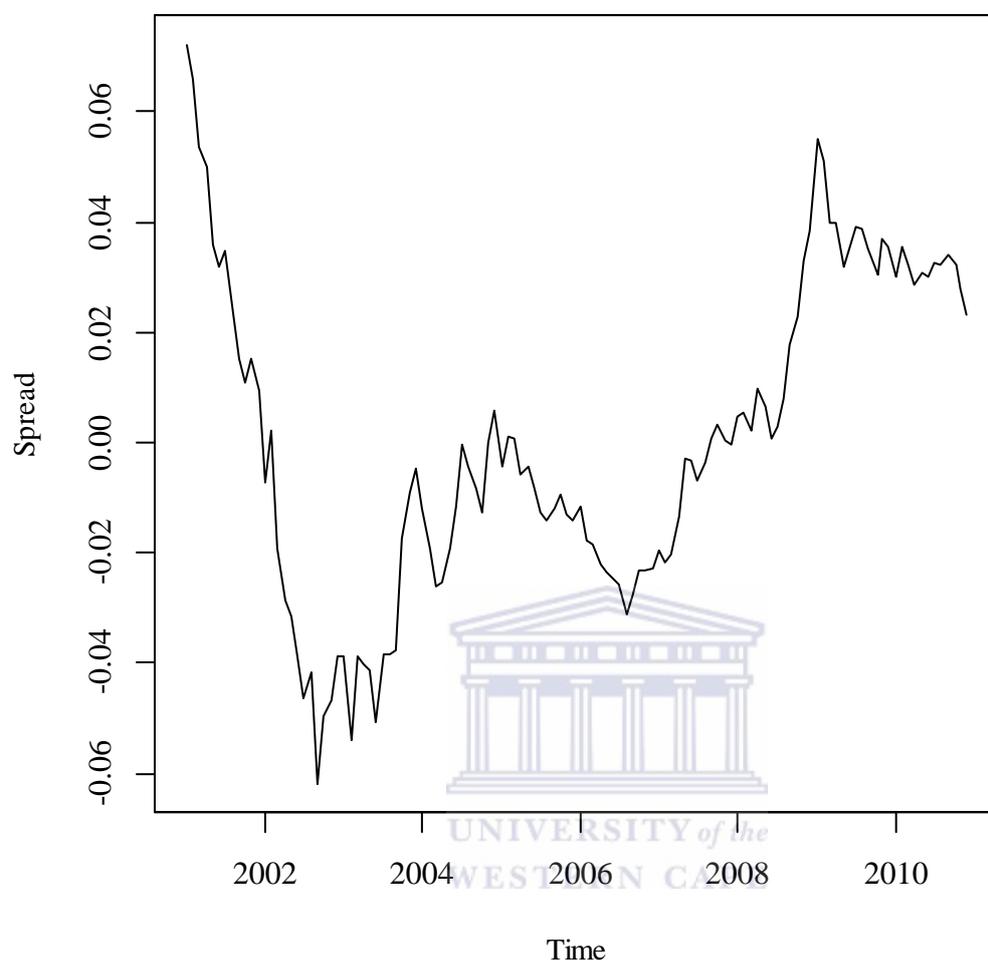


Figure 9: Price spread between the RTSI and BOVESPA stocks.

The spread in figure 9 is calculated from equation (53b). Its stationarity is not readily observable from the graph. However, the effects of the shift in prices around 2008 is fairly obvious around 2009.

Since it is assumed that both equations (53a) and (53b) show long-run equilibriums of the pairs, then the spreads must prove to be stationary. We test this by applying the ADF and PP unit root tests. Both tests separately investigate the following hypotheses:

$H_0$ : Unit root process.

$H_a$ : Stationary process.

The test results are given in the following table.

Table 12: Unit root tests.

Stock Pair	Unit root test	
	ADF	PP
S-R	Do not reject $H_0$	Reject $H_0$
R-B	Reject $H_0$	Reject $H_0$

Notes: Since both the ADF and PP tests are asymptotically equivalent, critical values are taken from Dickey-Fuller (1981) and Hamilton (1994).

UNIVERSITY of the  
WESTERN CAPE

We adhere mainly to the PP test results over the ADF results since the PP test has the advantage of being robust to heterogeneity of the respective error processes (Pfaff, 2008). We therefore conclude that both price spreads are mean reverting. This is a pre-condition that has to be met for successful implementation of the pairs trading. If we convert the cointegrated stock prices to the dollar denomination then the relationship in (52) can be described as: for every 1 dollar long position held in  $l_{rtsi}$ , there is a 0.77 dollar short position held in  $l_{senssex}$ . Similarly, the relationship in (53) can be described as: for every 1 dollar long position held in  $l_{bovespa}$ , a corresponding 1.05 dollar short position should be held in  $l_{rtsi}$ . Since the implementation of the strategy depends on the spreads reverting back to their respective equilibriums, then it is vital to know when to enter and exit the market. Hence the creation of trading signals.

We now estimate parameters for the back testing. The parameters are estimated from a 30 month-period arbitrarily chosen to start from January 2001 to June 2003. The spread, trading

signals and returns for the remaining 90 months will then be based on the parameters estimated from our historical sample.

*Step 3: Create trading signals.*

We create the trading signals for each trading pair based on whether each respective spread is over a 5% threshold. We also provide 1% and 10% thresholds as contrasts. The trading signals created for each pair are given below:

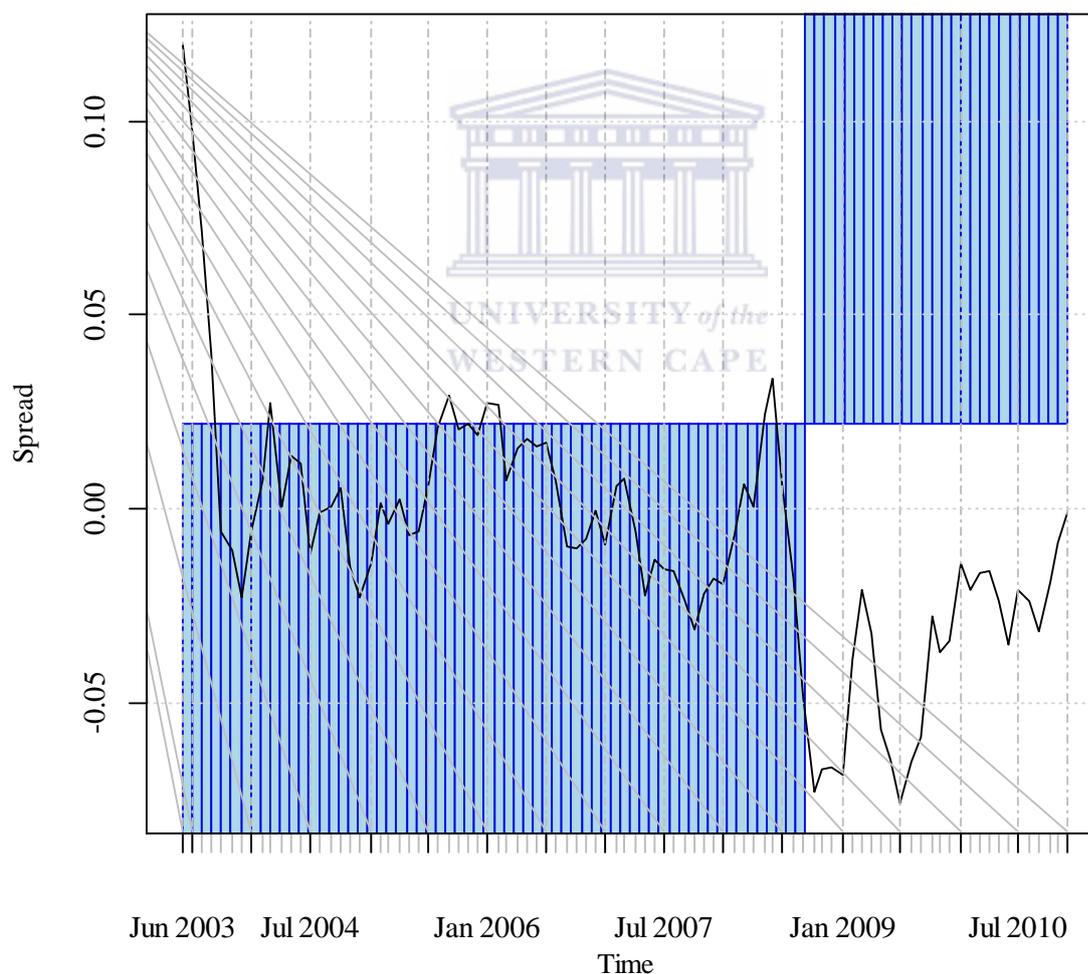


Figure 10a: Trading signals at 5% spread (SENSEX-RTSI).

Figure 10a suggests that we hold the S-R stock pair until August 2008. From then on enter into a trade by going long in the RTSI and shorting the SENSEX stocks. Both transactions must be of the same dollar amounts.

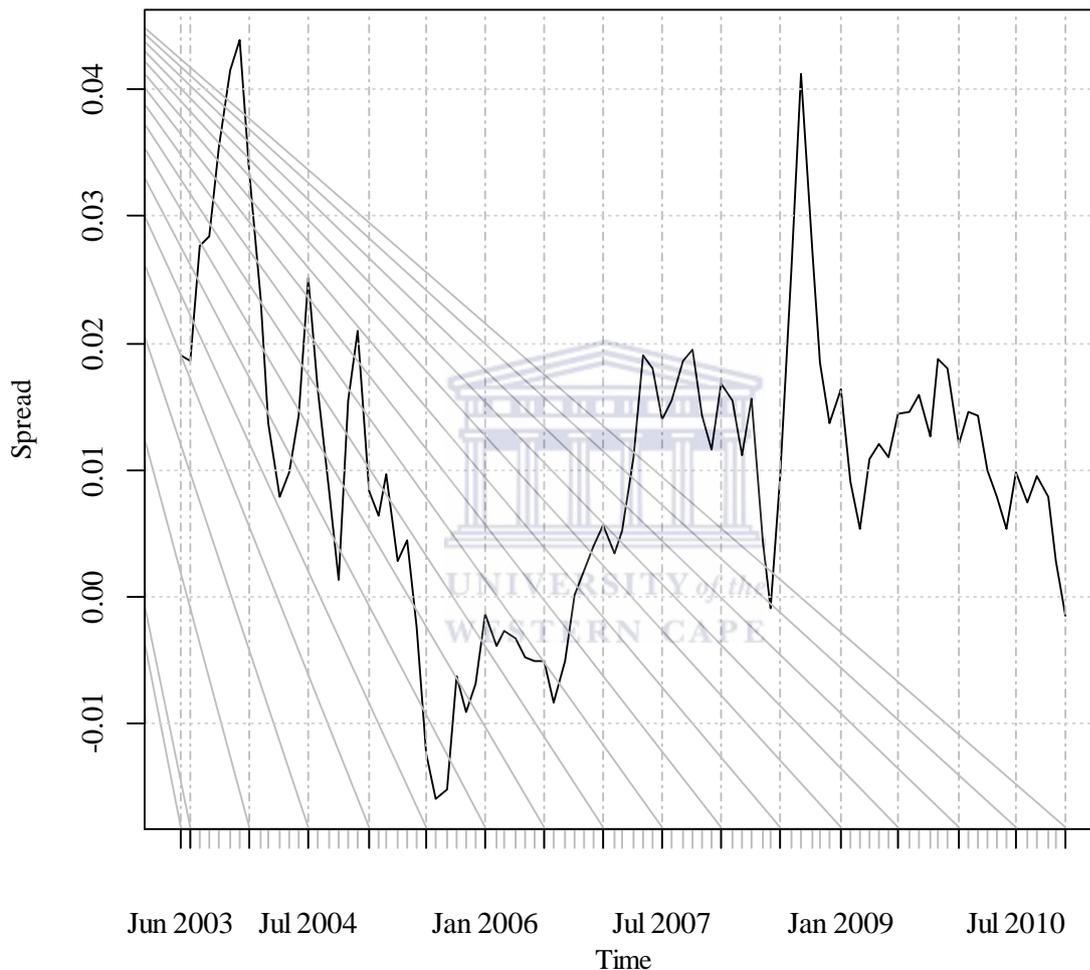


Figure 10b: Trading signals at 5% spread (RTSI-BOVESPA).

At the 5% spread threshold, an S-R pair trade could be implemented in the month of September, 2008. The strategy would be to buy long the winning stock (RTSI) and sell short on the losing stock (SENSEX). This position is held until 2010. The position is closed by selling short the stock that was previously bought and buying long the stock that was previously short. On the other hand, the R-B pair spread shows no market entry and exit points at the 5% threshold. This could imply that the price spread between the two stocks

does not diverge wide enough to be significant at 5%. Hence, we look also at spreads with respect to 1% and 10% significance levels.

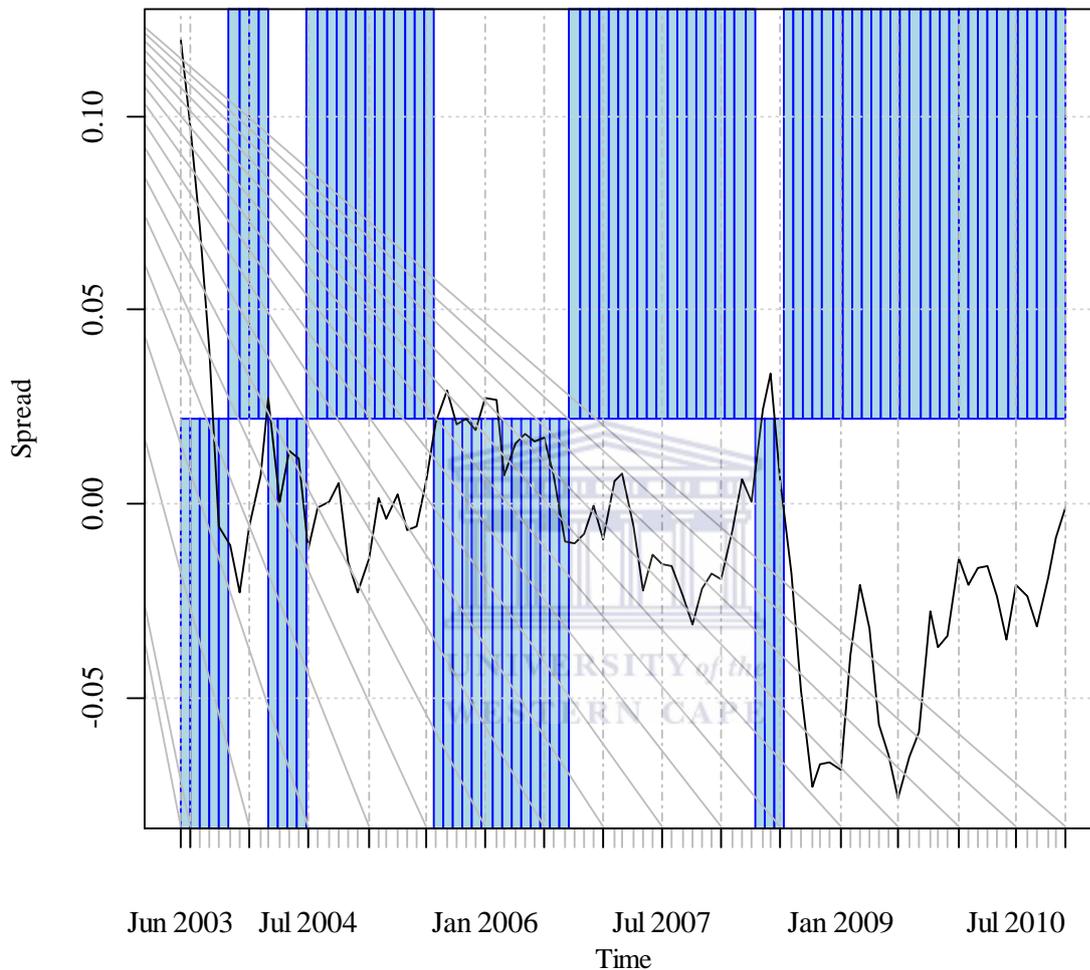


Figure 10c: Trading signals at 1% spread (SENSEX-RTSI).

The 1% spread threshold in figure 10c reveals multiple market entry points for the S-R pair. Four trade entry points can be identified: October 2003, July 2004, October 2006 and July 2008. For each of these entries the positions are closed on the following dates; February 2004, July 2005 and April 2008 respectively. The last entry point holds the longest for over 2.5 years before closing.

Using monthly closing prices, we convert the stock prices to USD denominations. The tabulated prices and their conversion rates for each entry and exit trade points are given below. We do this so that we can calculate and effectively compare estimated percentage gains realized on each trade based on opening and closing positions. The returns are calculated on single “round-trip” pairs trades. In the words of Karvinen (2012), a single round-trip pairs trade is defined by market entry and exit points, i.e. opening and closing positions.

The opening and closing positions for the SENSEX-RTSI pairs trade are as follows:

Entry position: Buy long in the RTSI and sell the SENSEX short.

Exit position: Cover the short in the SENSEX stock by going long and short RTSI.

Table 13a: Entry position for the SENSEX-RTSI pairs trade.

	Opening Position – Go long in the RTSI and shorting the SENSEX					
	Local Currencies		RUBLE/\$	RUPEE/\$	USD converted	
	RTSI (RUBLE)	SENSEX (RUPEE)			RTSI	SENSEX
Entry Dates						
October 2003	506.12	4906.87	30.1411	45.569	16.79	107.68
July 2004	540.27	5170.32	29.07894	46.035	18.58	112.31
October 2006	1613.57	12961.9	26.86045	45.77	60.07	283.20

Notes: RUPEE/\$ acquired from Plexus Holding Limited (South Africa) and RUBLE/\$ acquired from <http://www.gocurrency.com>.

In table 13a only “dollar value” entry positions are calculated. The exit dates and dollar value positions are given in table 13b below.

Table 13b: Exit position for the SENSEX-RTSI pairs trade.

	Closing Position- Go long in the SENSEX and shorting the RTSI					
	Local Currencies		RUBLE/\$	RUPEE/\$	USD converted	
	RTSI	SENSEX			RTSI	SENSEX
Exit Dates						
February 2004	670.14	5667.51	28.49862	45.305	23.51	125.10
July 2005	779.18	7635.42	28.68897	43.465	27.16	175.67
April 2008	2122.5	17287.3	23.51208	39.88	90.27	433.48

Notes: RUPEE/\$ acquired from Plexus Holding Limited (South Africa) and RUBLE/\$ acquired from <http://www.gocurrency.com>.

Table 13c: Profit and loss on SENSEX-RTSI pairs trade.

RTSI	Gain Ratio (%)	SENSEX	Loss Ratio (%)	Average Gain (%)
+6.72	40.02	-17.42	13.92	13.05
+8.58	46.18	-63.36	36.07	5.06
+30.2	50.27	-150.28	34.67	7.8

Notes: +/-=Profit/Loss, Profit/Loss=Opening position price-Closing position price,

Gain Ratio=Profit/Buying price, Loss Ratio=Loss/Selling price

Average Gain= (Gain Ratio-Loss Ratio)/2.

As expected, losses are recorded on all short positions held in the SENSEX. These are then offset by the gains realized from the long positions held in the RTSI. On average, the first trade performed better compared to trades entered into in the subsequent times. It is interesting to note that the first trade cycle closed in February, 2004. Earlier analysis (see Chapter 5, section 5.1 figure 4) has shown that a price spike was recorded in 2004, therefore

it follows that the performance of the first trade would do well. The news about the US credit crisis can be assumed to have been the cause of the poorer performance in the second trade and the onset of the recession may have shaken investor confidence enough to result in the lesser performance of the final trade which closed out in 2008.

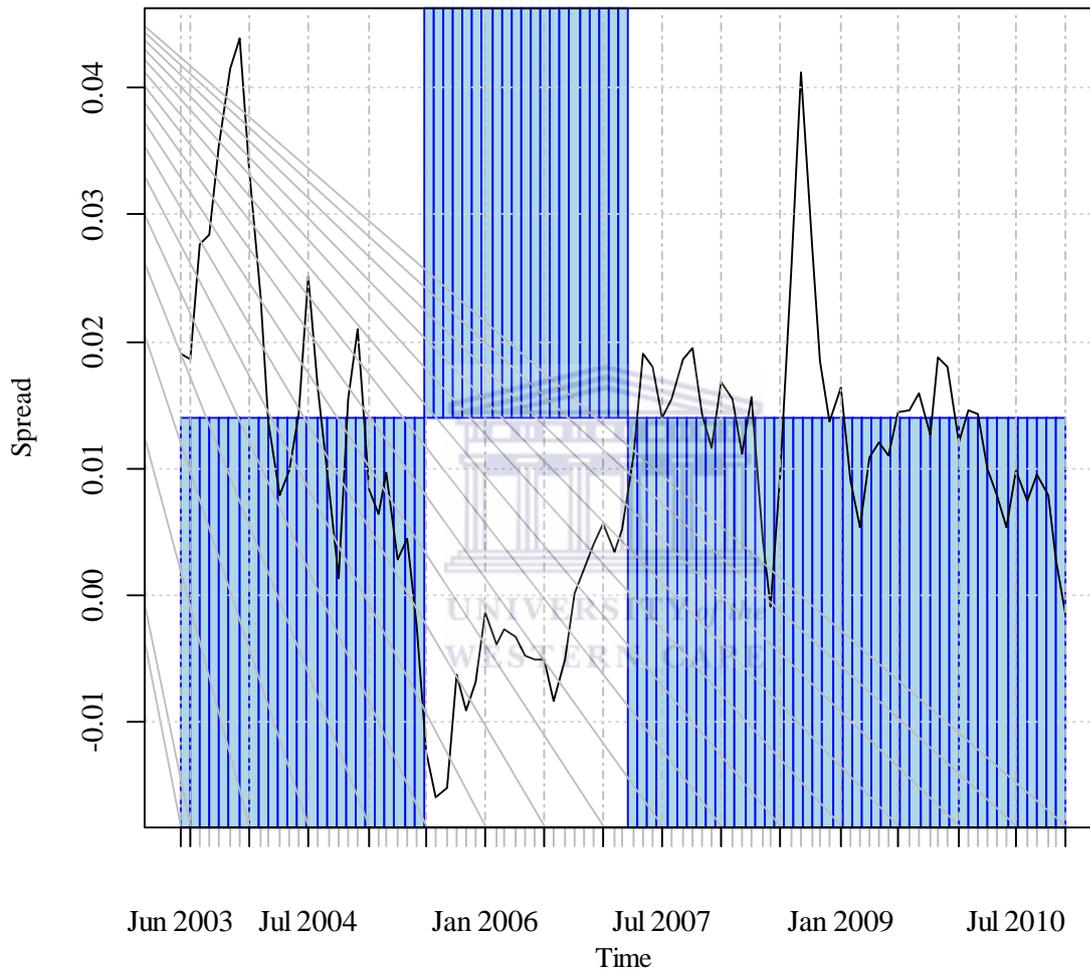


Figure 10d: Trading signals at 1% spread (RTSI-BOVESPA).

In comparison, figure 10d shows far less market entry points for the R-B pair compared to the S-R pair at the 1% spread threshold. The R-B pair indicates one trade entered on the June 2005 and closed on March 2006. We calculate the profit and loss of this trade as we did for the S-R pair. The entry and exit positions are:

Entry position: Buy long in the BOVESPA stock and short RTSI.

Exit position: Go long in the RTSI and short sell BOVESPA.

Table 14a: Entry position for the RTSI-BOVESPA pairs trade.

	Opening Position – Go long in the BOVESPA and short sell the RTSI stock					
	Local Currencies		RUBLE/\$	REAL/\$	USD converted	
	RTSI (RUBLE)	BOVESPA (REAL)			RTSI	BOVESPA
Entry Dates						
June 2005	698.4	25051	30.1411	2.41323	23.17	10380.69

Notes: RUPEE/\$ and REAL/\$ ratios acquired from <http://www.gocurrency.com>.

The market entry date to trade in the R-B stock pair is given in table 14a as June 2005. The closing date for this position is given in table 14b as March, 2006.

Table 14b: Exit position for the RTSI-BOVESPA pairs trade.

	Closing Position – Go long in the RTSI and sell the BOVESPA short					
	Local Currencies		RUBLE/\$	REAL/\$	USD converted	
	RTSI (RUBLE)	BOVESPA (REAL)			RTSI	BOVESPA
Exit Dates						
March 2006	1434.99	37952	27.87645	2.14374	51.48	17703.64

Notes: RUPEE/\$ and REAL/\$ ratios acquired from <http://www.gocurrency.com>.

We are able to only determine one trading cycle for the R-B stock pair in our sample period. The gain realized in this trade is calculated and given in table 14c below as 7.78% of the initial investment capital.

Table 14c: Profit and loss on RTSI-BOVESPA pairs trade

RTSI	Loss Ratio (%)	BOVESPA	Gain Ratio (%)	Average Gain (%)
-28.31	54.99	7322.95	70.54	7.78

Notes: +/-=Profit/Loss, Profit/Loss=Opening position price-Closing position price,

Gain Ratio=Profit/Buying price, Loss Ratio=Loss/Buying price

Average Gain= (Gain Ratio-Loss Ratio)/2

The three trading opportunities presented by the S-R stock pair show accumulated returns of about 26% averaging about 8.64% per trade. On the other hand, the R-B pair trade realized only a 7.8% return. The lower returns from the R-B pair may stem from the relatively weak relationship between RTSI and BOVESPA which showed to be cointegrated only at 10% significance level. Since the S-R showed strong signs of cointegration at 5% significance level under the trace test (see table 8 above), this may be a significant attribute to the higher returns calculated.

Below we observe the effects of setting the spread threshold too wide for the two traded pairs, S-R and R-B.

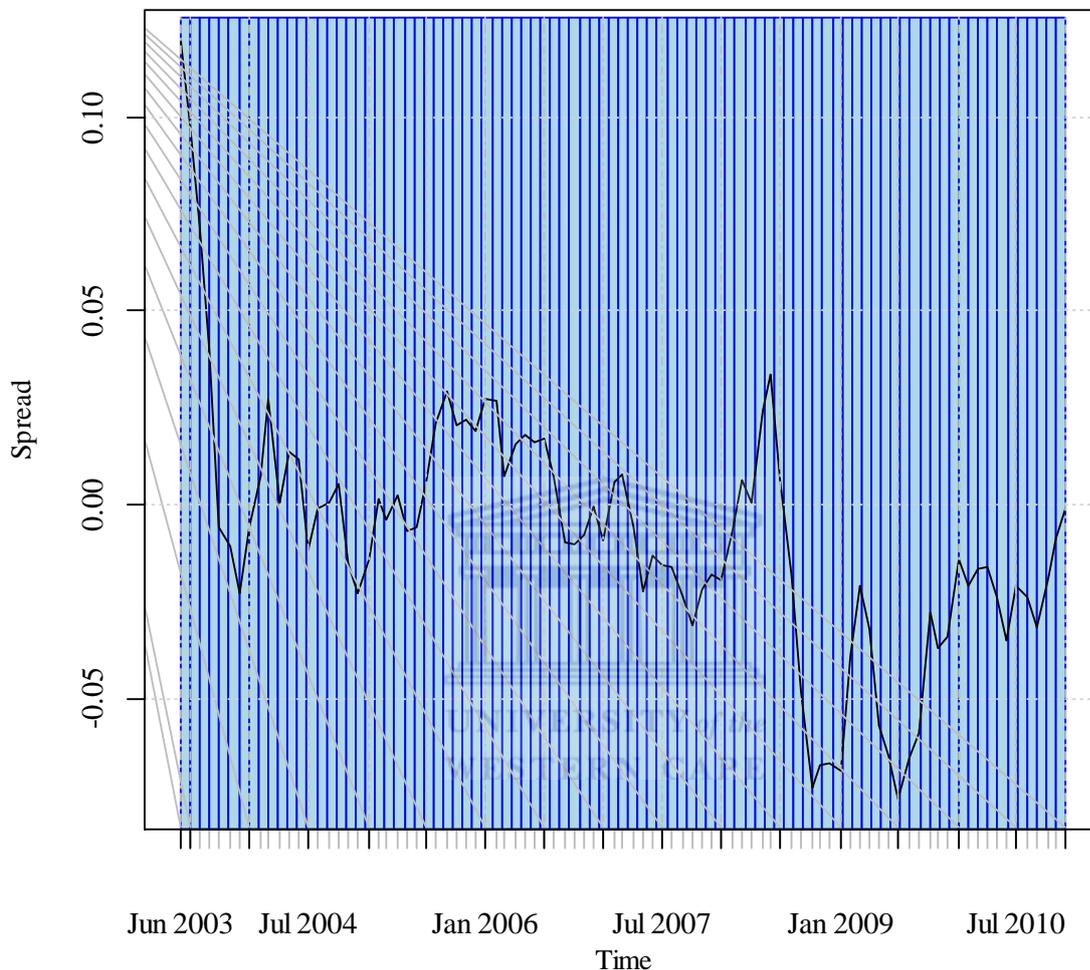


Figure 10e: Trading signals at 10% spread (SENSEX-RTSI).

At a 10% threshold, the S-R pair yields inconclusive results since the signals produced seem to suggest simultaneously entering into trades and holding positions in the stocks. On the other hand, the graph of the R-B pair below shows that no trading signals are generated at all at a 10% threshold. This suggests that no trades can be implemented given such boundary parameters.

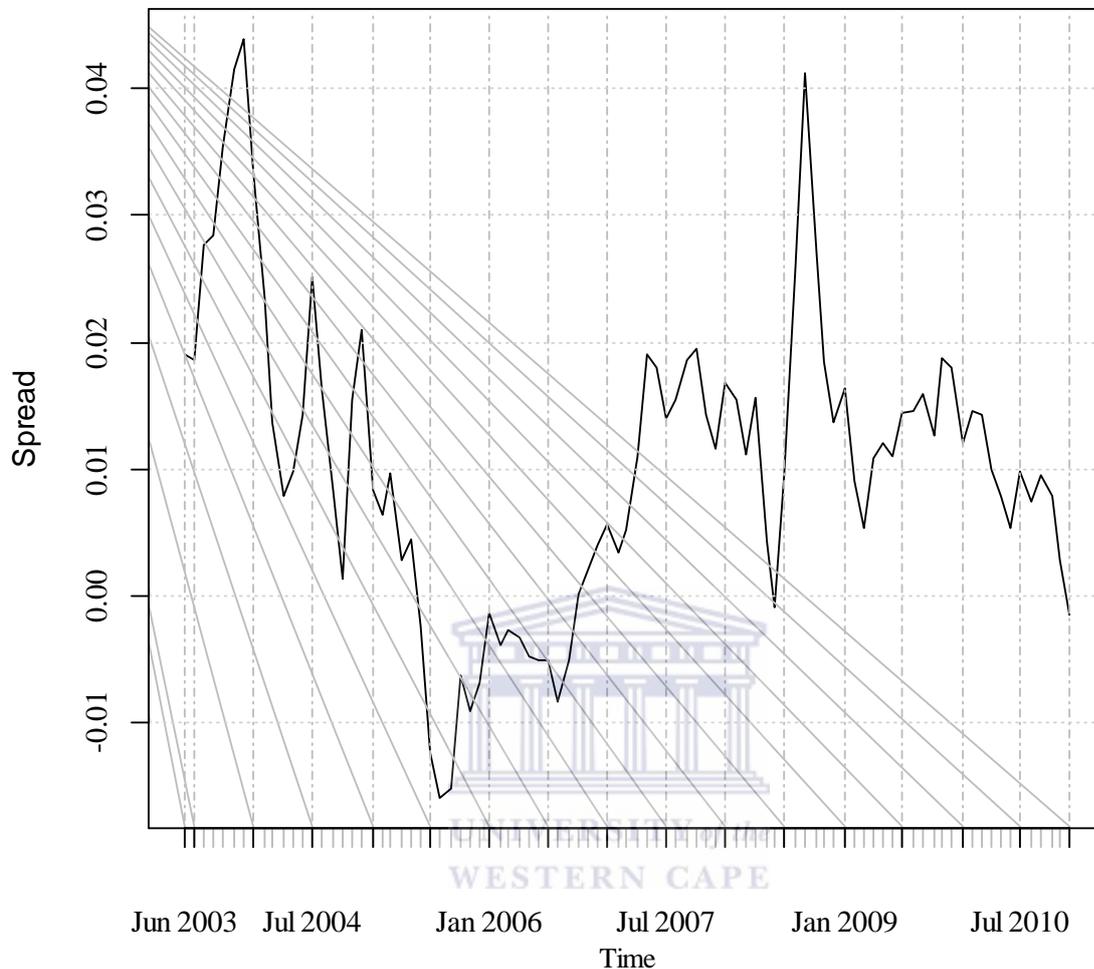


Figure 10f: Trading signals at 10% spread (RTSI-BOVESPA).

In conclusion, we realize that optimal pair trading positions can be realized when we consider a 5% threshold to each of our pair price spreads. Setting the threshold to this level produces an optimal number of entry positions for each pair trade. This leads to an investment strategy that allows for the minimization of transaction costs that can be incurred from entering into trades. Entering into numerous trades on the same stock pair will most likely culminate in transaction costs overwhelming the trade gains (Gatev et al., 2006). This could be the case, for example, when the spread threshold is set to 1% even though more trade entry points are created compared to the 5% threshold.

*Step 4: Performance of each pair trade.*

The performance of each trade is determined from its estimated returns based on the estimated hedge ratio and the respective trading signals. Since some of the trade conditions did not produce trading signals, we therefore deem it only necessary to evaluate the performance of those trades that yielded trading signal.

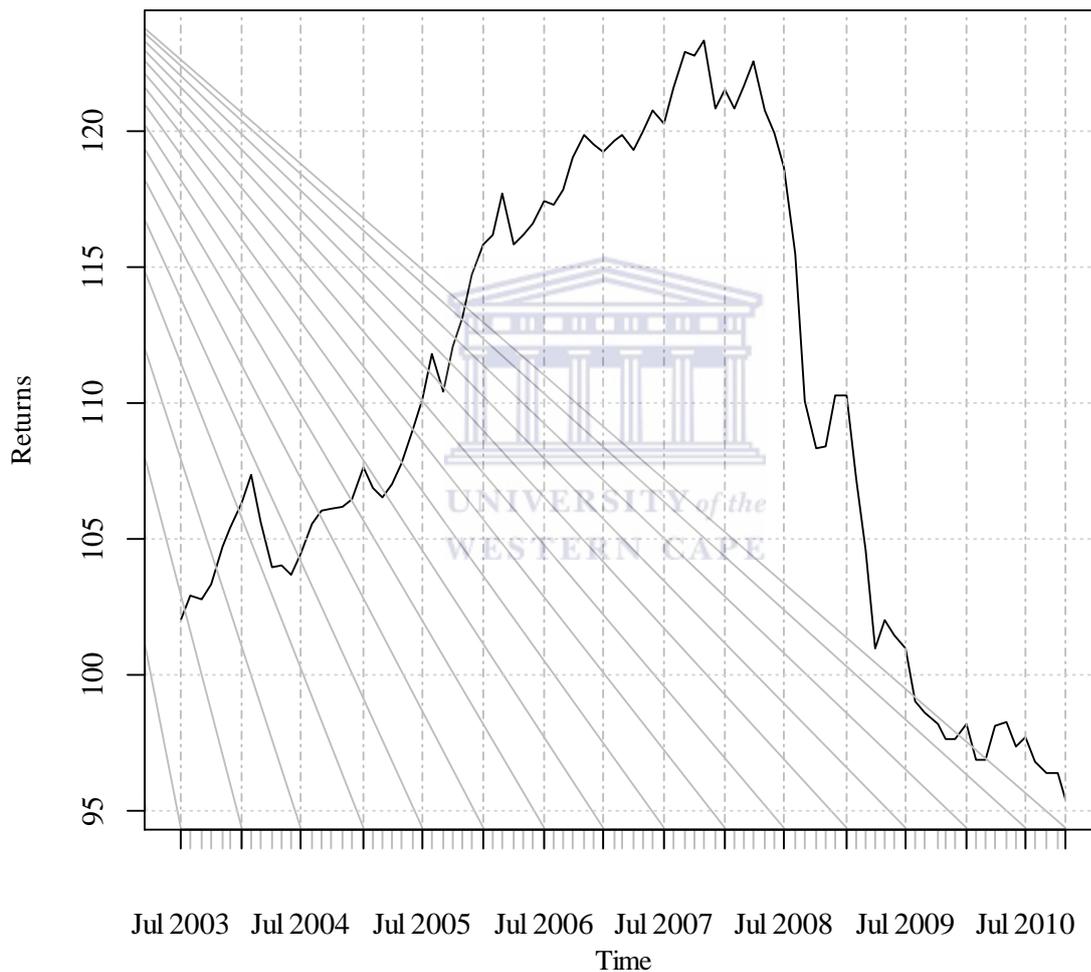


Figure 11a: Performance of the SENSEX-RTSI pairs trading at 5% spread.

Figure 11a shows increasing excess returns from 2003 until around mid-2007. For this series, the S-R pair trade is based on the 5% spread threshold. This performance could be attributed to the different speeds at which the Indian and Russian markets entered the recession. Combined reports from [erube.info](http://erube.info) and [www.economist.com](http://www.economist.com) say that the world stock indices peaked around October 2007 before plummeting soon after.

The MSCI emerging markets index is reported to have fallen 40.5% around the same time and not long after that both India and Russia slipped into recession around December 2008 and June 2008 respectively. The effects of this event can be seen dropping quite fast between 2008 well into 2010.

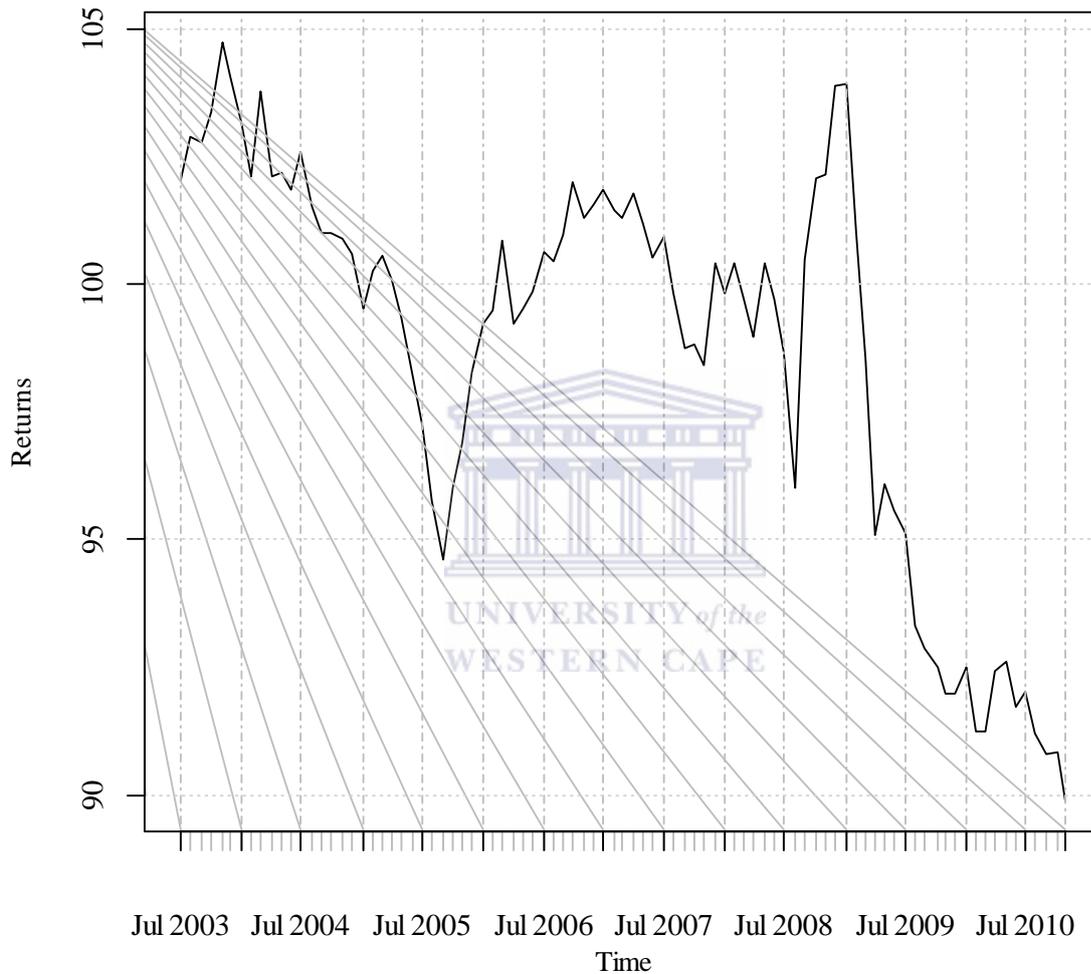


Figure 11b: Performance of the SENSEX-RTSI pairs trading at 1% spread.

The return function of the SENSEX-RTSI shows two sharp declines in 2005 and 2008 followed by a prominent peak in 2009 and a sharp decline thereafter just as in figure 11a. The R-B portfolio return is relatively steady from 2003 to about 2008. There is a sharp drop in the late 2008 followed by a sudden increase from 2009 which seems to settle on a familiar level beyond 2010.

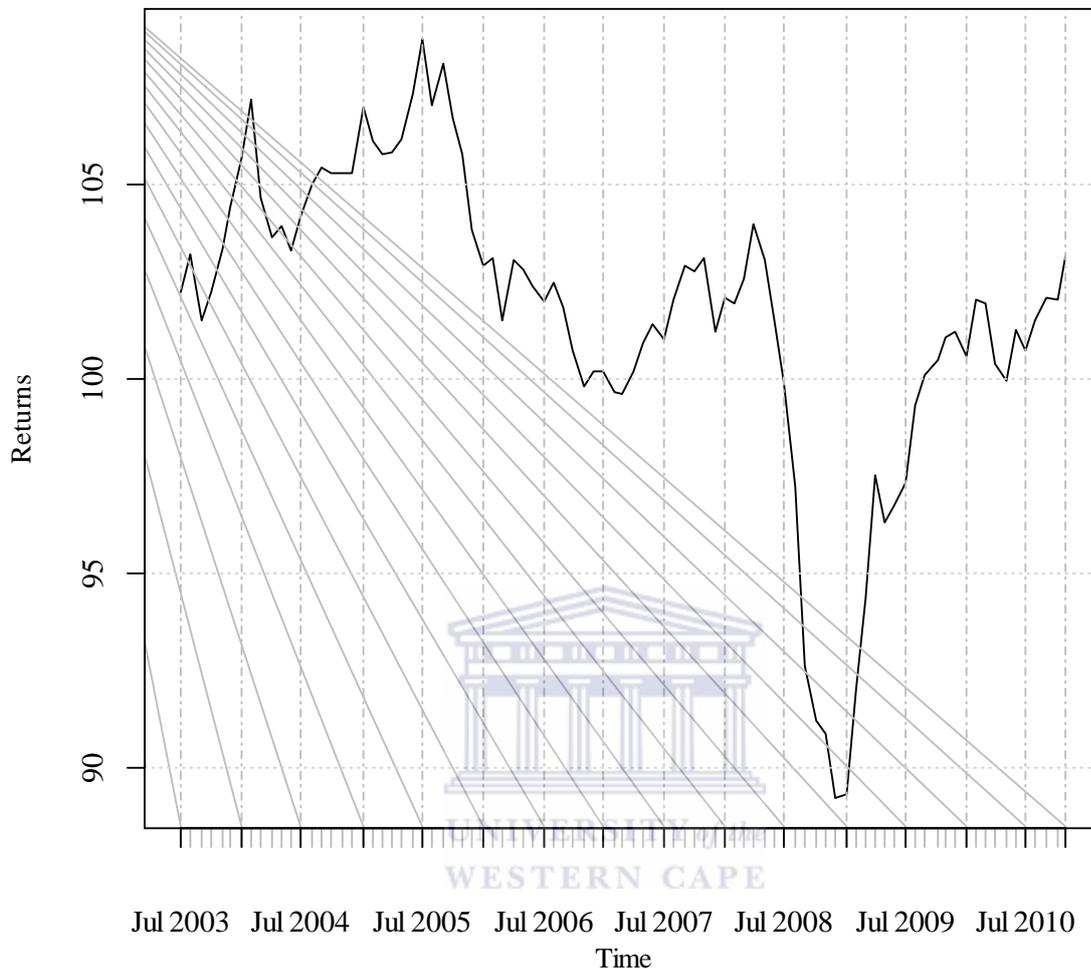


Figure 11c: Performance of the RTSI-BOVESPA pairs trading at 1% spread.

The RTSI-BOVESPA stock pair shows highest performance around 2005. The series also shows the clear evidence of the effects of the recession; this is evidenced by the sharp drop in returns between 2008 and 2009. Beyond 2009 the pair seems to exhibit steady improvement.

Assuming a zero risk free rate, the Sharpe ratios of both S-R and R-B are calculated as 1.166183 and -0.91 using equation (54) (see section 3.4). These results suggest that investment in the S-R pair does not come with too much risk compared to investing in the R-B pair. Furthermore, the negative sign on the R-B Sharpe ratio suggests that a risk-less asset would perform considerably better than the pair.

## CHAPTER 6

### Conclusion and recommendations

By analyzing the major stock market indices, we have managed to show that the BRICS economies are cointegrated between January, 2001 and December, 2010. This suggests that the indices historically move together. Therefore, from an investment point of view, the markets can also be expected to continue moving together in future. These results speak to the first part of the research question given in Chapter 1 and will be discussed further below.

However, we found that the BRICS are only cointegrated when the structural shift of 2007 is taken into consideration. The trace test suggested that the BRICS share two cointegrating equations while the eigenvalue test concluded on only one cointegrating relation. These results address the second part of the research question. Thus it is clear that structural breaks can have significant influence on how the data is analysed and as a result, the conclusion of the analysis.

For example, ignoring the breaks in our analysis could have suggested that it is possible to internationally diversify our portfolio. This would have led to the investigation significantly diverting from addressing the key issue of long-run relationships between the BRICS to portfolio diversification. In addition, the data would have had to be modeled by a VAR form with differenced variables instead of the VEC model. In this case, we hypothesize that the VAR model's ability to predict market performance, to any level, would have had to be brought to question. This is because the model would have been technically inadequate to handle large economic shifts since the data clearly presented with very volatile time periods within the sample period investigated.

According to the unit root test results, considering structural breaks in the tests provided further information on the underlying processes of the individual series. The data were found not only to be non-stationary in levels but also to contain significant breaks in the mean stock prices. Therefore, we can speculate that the effects of the 2007-2009 recession had profound effects on the average prices of the BRICS stock markets. Hence, the inclusion of a structural shift dummy variable in the estimation of the VECM was justified. This course of action

proved to have adverse effects on the cointegration test. As pointed above, the data showed only to be cointegrated after concluding that the effects of the economic recession were significant enough to consider in the analysis. The cointegration rank also proved to be affected as both the trace and eigen value tests gave different results in that respect. However, we note that these results may possibly not be entirely attributed to the effects of the structural shift.

Having satisfactorily concluded that the BRICS are cointegrated, we proceeded to test the estimated VECM for adequacy. We carried out this task by testing the underlying VAR structure for ARCH effects, normality and serial correlation. The VAR model with two long-run relations showed no serious signs of ARCH effects and the results of the normality tests were within acceptable boundaries. However, the VAR with one cointegrating equation exhibited adverse ARCH effects. As a result, we argued that perhaps the data are best modeled by a VAR with two estimated long-run relations instead of one.

Finesse is required in addressing the first part of the research question. The results of both correlation and cointegration suggest that the inclusion of the South African JSE into a portfolio consisting of only the BRICs major stock indices results in a portfolio that is poorly diversified. Since the cointegration results revealed that the BRICS share at least one long-run relation, it was also equally important to give insight into the short-run dynamics of the data. This was vital since the definition of the VEC model given in the Chapter 3 was that it not only captures the long-run relations of the data but also short-term effects and responses of the variables to each other and other external stimuli.

So, the high correlation coefficients served only to show that diversification in a portfolio consisting of just the BRICS major stock indices was very poor. After acknowledging this outcome, the key issue became about addressing the question “...what is the upside of holding a portfolio of stocks that have been found to historically move together?” It was therefore imperative to guide this research to chiefly investigate this question and not the question of how to diversify our portfolio.

A consequence of the cointegration test results indicated that it would be possible to implement the pairs trading strategy, a statistical arbitrage technique. This would only be

feasible because of the mean reversion property covered by cointegration. This meant that some of the portfolio's investment risk could be diversified away through hedging. That is to say, since the stocks were found to be highly correlated, a long position in one stock would be expected to be offset (hedged) by a short position in another. The pairs used in the implementation of the trading strategy were SENSEX-RTSI and RTSI-BOVESPA. These were selected based on their long-term relationships, that is, SENSEX was found to be cointegrated with RTSI and RTSI was found to be cointegrated with BOVESPA.

The results showed that if a 1% spread threshold is set for both stock pairs, the second pair (RTSI-BOVESPA) outperformed the first pair (SENSEX-RTSI). However, if provision is made for trading costs incurred in each trade, it followed that entering into numerous trades on the same pair would highly erode the gains. This led to the realization that perhaps the optimal spread ceiling would be best set at 5% instead of 1%. However, the results showed that the 5% spread threshold yielded unsatisfactory trading signals over the sample period. Only one market entry point for the SENSEX-RTSI pair was generated while no entry points were generated for the RTSI-BOVESPA pair at all.

Thus, for the purposes of this research, we conclude that a 1% spread may be adequate to simulate an interactive market pairs trading strategy for both pairs considered. It should be mentioned that no transaction costs were actually calculated in the simulation of the pairs trading technique mentioned above.

At 1% price spread, the SENSEX-RTSI pair averaged returns of about 8.64% per trade and the RTSI-BOVESPA pair showed only a 7.8% return. The Sharpe ratios were calculated as 1.166183 and -0.91 respectively. These results indicated that it is more profitable and less risky to invest in the SENSEX-RTSI stock pair than in the RTSI-BOVESPA pair.

### Recommendations:

In our endeavor, we found very little research that explicitly utilizes market capitalization as a normalization tool in regression analysis. For this reason, we realize that the arbitrary nature of the date chosen to select the market capitalization figures may very well be almost completely based on the researcher's own reasons. Consequently, such an act opens the research up for argument in this respect. We hypothesize that if different capitalization figures are selected on a different date, then the resulting choice of explanatory and dependent variables may be severely influenced otherwise. As a result, the estimated regression will be interpreted differently based on the new estimates and may strongly change the results of the ensuing analysis.



## Bibliography

- Aggarwal, R., Inclan, C., & Leal, R. (1999). Volatility in emerging stock markets. *Journal of Financial and Quantitative Analysis*, 34(1), 1-31.
- Ahmed, M. (2008). *Rupee depreciated by 27 percent during 2008*. Retrieved January 2012, from Daily Times:  
[http://www.dailytimes.com.pk/default.asp?page=2008%5C12%5C28%5Cstory\\_28-12-2008\\_pg5\\_5](http://www.dailytimes.com.pk/default.asp?page=2008%5C12%5C28%5Cstory_28-12-2008_pg5_5)
- Ahn, S. K., & Reinsel, G. C. (1990). Estimation for partially nonstationary multivariate autoregressive models. *Journal of the American Statistical Association*, 85(411), 813-823.
- Akaike, H. (1970). Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, 22(1), 203-217.
- Alagidede, P., & Panagiotidis, T. (2009). Modelling stock returns in Africa's emerging equity markets. *International Review of Financial Analysis*, 18(1), 1-11.
- Alexakis, C. (2010). Long-run relations among equity indices under different market conditions: Implications on the implementation of statistical arbitrage strategies. *Journal of International Financial Markets, Institutions & Money*(20), 389-403.
- Alexakis, C. A., & Mavrakis, E. K. (2008). Stock markets' linkages: An empirical investigation for long-term international diversification benefits. *International Research Journal of Finance and Economics*(21), 1450-2887.
- Alexander, C. (1999). Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society, Series A*(357), 2039-2058.
- Alexander, C. (2008). *Market risk analysis, practical financial econometrics* (Vol. 2). West Sussex: John Wiley & Sons Ltd.
- Alexander, C., & Dimitriu, A. (2002). The cointegration alpha: Enhanced index tracking and long-short equity market neutral strategies. *ISMA Discussion Papers in Finance*, 1-54.

- Allaro, H. B., Kassa, B., & Hundie, B. (2010). A time series analysis of structural break time in the macroeconomic variables in Ethiopia. *African Journal of Agricultural Research*, 6(2), 392-400.
- Andrade, S., di Pietro, V., & Seasholes, M. (2005). Understanding the profitability of pairs trading. *Unpublished working paper, UC Berkeley, Northwestern University*.
- Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821-856.
- Andrews, D. W., & Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62(6), 1383-1414.
- Andrews, D. W., & Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62(6), 1383-1414.
- Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), 47-78.
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1-22.
- Bai, J., & Perron, P. (2003). Critical values for multiple structural change tests. *The Econometrics Journal*, 6(1), 72-78.
- Bajo-Rubio, O., Díaz-Roldán, C., & Esteve, V. (2008). US deficit sustainability revisited: A multiple structural change approach. *Applied Economics*, 40(12), 1609-1613.
- Balarezo, J. (2010, August 24). International diversification using cointegration and morden portfolio theory. (*Masters Thesis*) Copenhagen Business School.
- Banerjee, A., Lumsdaine, R. L., & Stock, J. H. (1992). Recursive and sequential tests of the unit-root and trend-break hypotheses: theory and international evidence. *Journal of Business & Economic Statistics*, 10(3), 271-287.
- "Brasilia Declaration". (2003, June 6). *Brasilia Declaration*. Retrieved February 21, 2013, from IBSA: [http://www.ibsa-trilateral.org/brasil\\_declaration.htm](http://www.ibsa-trilateral.org/brasil_declaration.htm)

- Bauwens, L., Deprins, D., & Vandeuren, J.-P. (1997). *Modelling interest rates with a cointegrated VAR-GARCH model*. Center for Operations Research & Econometrics. Université catholique de Louvain.
- Beck, N. (2001). *Time series*. University of California, Department of Political Science, San Diego.
- Beck, T. (2008, April). The econometrics of finance and growth. *Policy research working paper 4608*.
- Behar, G., & Hest, S. (n.d.). *Emerging markets. Analysis and commentary on the future of emerging markets*. Retrieved May 21, 2012, from Northern Trust:  
<http://www.northerntrust.com/pws/jsp/display2.jsp?XML=pages/nt/0802/search.xml&q=Behar&start=20&dc=20>
- BM&FBOVESPA. (n.d.). *BM&FBOVESPA*. Retrieved June 30, 2012, from  
<http://www.bmfbovespa.com.br/home.aspx?idioma=pt-br>
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
- Boswijk, P. H., & Zu, Y. (2007). Testing for cointegration with nonstationary volatility.
- Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Francisco Holden-Day.
- Breusch, T. S., & Godfrey, L. G. (1980). A review of recent work on testing for autocorrelation in dynamic economic models. *University of Southampton*.
- Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B (Methodological)*, 37(2), 149-192.
- Burdekin, R. C., & Silkos, P. L. (1999). Exchange rate regimes and shifts in inflation persistence: does nothing else matter? *Journal of Money, Credit and Banking*, 31(2), 235-247.

- Campbell, J. Y., & Shiller, R. J. (1987, October). Cointegration and tests of present value models. *Journal of political economy*, 95(5), 1062-88.
- Campbell, J. Y., & Shiller, R. J. (1987, October). Cointegration and tests of present value models. *Journal of political economy*, 95(5), 1062-88.
- Caner, M., & Kilian, L. (2001). Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate. *Journal of International Money and Finance*, 20(5), 639-657.
- Chatfield, C. (1980). *The analysis of time series: An introduction* (2 ed.). London: Chapman & Hall.
- Cheng, X. (2008). Mixture time series models and their applications in volatility estimation and statistical arbitrage trading. *Unpublished manuscript Thesis (Masters) University of Hong Kong*, 1-108.
- Cheung, Y.-W., & Lai, K. S. (1993). Finite-sample sizes of Johansen's likelihood ratio tests for cointegration. *Oxford Bulletin of Economics and statistics*, 55(3), 313-328.
- Chow, D. (2010). China's response to the global financial crisis: Implications for US-China economic relations. *I*(47).
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591-605.
- Christiano, L. J. (1992). Searching for a break in GNP. *Journal of Business and Economic Statistics*, 10(3), 237-250.
- Christopoulos, D. K., & Tsionas, E. G. (2005). Productivity growth and inflation in Europe: Evidence from panel cointegration tests. *Empirical Economics*(30), 137-150.
- Cochrane, J. H. (1991). A critique of the application of unit root tests. *Journal of Economic Dynamics and control*, 275-284.
- Crowder, W. J. (1994). Foreign exchange market efficiency and common stochastic trends. *Journal of International Money and Finance*, 13(5), 551-564.

- Davidson, J. E., Hendry, D. F., Srba, F., & Yeo, S. (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *The Economic Journal*, 661-692.
- Dawson, C. (2005, December 25). *Emerging markets: Beyond the big four*. Retrieved 07 28, 2012, from BloombergBusinessweek Magazine:  
<http://www.businessweek.com/stories/2005-12-25/emerging-markets-beyond-the-big-four>
- DeJong, D. N., & Charles, W. H. (1991). Reconsidering trends and random walks in macroeconomic time series'. *Journal of Monetary Economics*, 28(2), 221-254.
- Dickey, D. A. (2005). Stationarity issues in time series models. *North Carolina State University, paper*, 192-30.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431.
- Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4), 1057-1072.
- Diebold, F. X., & Kilian, L. (2001). Unit-root tests are useful for selecting forecasting models. *International journal of forecasting*, 17(2), 323-25.
- Dubois, M., & Bacmann, J.-F. (2002). Volatility in emerging stock markets revisited. *Working paper, University of Neuchâtel*.
- Dwyer, G. P., & Wallace, M. S. (1992). Cointegration and market efficiency. *Journal of International Money and Finance*, 11(4), 318-327.
- Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4), 813-836.
- Engel, C. (1996). A note on cointegration and international capital market efficiency. *Journal of International Money and Finance*, 15(4), 657-660.

- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica. Journal of the Econometric Society*, 987-1007.
- Engle, R. F., & Yoo, B. S. (1987). Forecasting and testing in co-integrated systems. *Journal of econometrics*, 35(1), 143-159.
- Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2), 251-276.
- Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251-276.
- Engle, R. F., & Granger, C. W. (2003). Time-series econometrics: cointegration and autoregressive conditional heteroskedasticity. *Advanced information on the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel*, 1-30.
- Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5), 1749-1778.
- Engle, Robert. (2001). GARCH 101: An introduction to the use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4), 157-168.
- Exley, J., Mehta, S., & Smith, A. (2004). Mean reversion. *Faculty & Institute of Actuaries Finance and Investment Conference*. Brussels.
- Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*(25), 383-417.
- Febrian, E., & Herwany, A. (2007, October 15). Co-integration and causality among Jakarta stock exchange, Singapore stock exchange, and Kuala Lumpur stock exchange. *MPRA*, 4-12.
- Franses, P. H., Kofman, P., & Moser, J. (1994). GARCH effects on a test of cointegration. *Review of Quantitative Finance and Accounting*, 4, 19-26.
- Gatev, E., Goetzmann, W. N., & Rouwenhorst, G. K. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3), 797-827.

- Gebre-Mariam, Y. K. (2011). Testing for unit roots, causality, cointegration, and efficiency: The case of the northwest US natural gas market. *Energy*, 36(5), 3489–3500.
- Glynn, J., Perera, N., & Verma, R. (2007). Unit root tests and structural breaks: a survey with applications. *Faculty of Commerce-Papers*, 455.
- Goldfield, S. M., & Quant, R. E. (1965). Some tests for homoscedasticity. *Journal of the American Statistical Association*, 60, 539–547.
- Go currency.com. (n.d.). Retrieved August 7, 2012, from Go currency.com:  
<http://www.gocurrency.com/>
- Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 121-130.
- Granger, C. W., & Hallman, J. (1991, May). Nonlinear transformations of integrated time series. *Journal of time series analysis*, 12(3), 207-224.
- Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, 2, 111-20.
- Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, 2, 111-20.
- Granger, C. W., & Weiss, A. A. (1983). Time series analysis of error-correction models. *Studies in Econometrics, Time Series, and Multivariate Statistics*, 255-278.
- Greene, W. H. (1993). *Econometric analysis*. Toronto, New York: Maxwell Macmillan International.
- Gregory, A. W., & Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, 70(1), 99-126.
- Guidi, F. (2010). Cointegration relationship and time varying co-movements among Indian and Asian developed stock markets. *MPRA Paper No. 19853*.
- Gujarati, D. N. (1995). *Basic econometrics*. New York: McGraw-Hill.

- Hakkio, C. S., & Rush, M. (1991). Is the budget deficit "too large?". *Economic Inquiry*, 29(3), 429-45.
- Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton university press.
- Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 190-195.
- Hansen, B. E. (1992). Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics*, 10(3), 321-335.
- Hayashi, N. (2005). Structural changes and unit roots in Japan's macroeconomic time series: is real business cycle theory supported? *Japan and the World Economy*, 17(2), 239–259.
- Hendry, D. F. (2010). Professor Sir Clive W.J. Granger and cointegration. *Journal of Financial Econometrics*, 8(2), 162-168.
- Hendry, D. F., & Juselius, K. (2000). Explaining cointegration analysis: Part 1. *The Energy Journal*, 0(1), 1-42.
- Hendry, D. F., & Juselius, K. (2000). Explaining cointegration analysis: Part 1. *The Energy Journal*, 0(1), 1-42.
- Hendry, D. F., & Juselius, K. (2001). Explaining cointegration analysis: Part II. *The Energy Journal*, 75-120.
- Herlemont, D. (2004). Pairs trading, convergence trading, cointegration. *YATS Finances & Technologies*, 5, 1-31.
- Hirotsugu, A. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Hjalmarsson, E., & Österholm, P. (2007). Testing for cointegration using the Johansen methodology when variables are near-integrated. *International Finance Discussion Papers*, 7(915).

- Hogan, S., Jarrow, R., Teo, M., & Warachka, M. (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics*, 73, 525-565.
- Infantino, L. R., & Itzhaki, S. (2010). Developing high-frequency equities trading models. *Doctoral dissertation, Massachusetts Institute of Technology*.
- Islam, A. M., & Ahmed, S. M. (1999). The purchasing power parity relationship: causality and cointegration tests using Korea-U.S. exchange rate and prices. *JOURNAL OF ECONOMIC DEVELOPMENT*, 24(2), 1-111.
- Islam, A. M., & Ahmed, S. M. (1999). The purchasing power parity relationship: causality and cointegration tests using Korea-U.S. exchange rate and prices. *JOURNAL OF ECONOMIC DEVELOPMENT*, 24(2), 1-111.
- Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* 6.3, 255-259.
- Jayanthakumaran, Kankesu, Pahlavani, Mosayeb, & Neri, Frank. (2007). Trade reforms and reekpoints in Australia's manufactured trade: An application of the Zivot and Andrews model. *International Journal of Applied Econometrics and Quantitative Studies*, 4(2).
- Jevons, W. S. (1884). Investigations in currency and finance.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*(12), 231-254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society*, 1551-1580.
- Johansen, S. (1995). Likelihood-based inference in cointegrated vector autoregressive models. *New York*.

- Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration - with application to demand for money. *Oxford bulletin of economics and statistics*, 169-210.
- Johansen, Søren. (1992). Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics*, 52(3), 389-402.
- Juselius, K., & Hendry, D. F. (2001). Explaining Cointegration Analysis: Part II. *The Energy Journal*, 0(1), 75-120.
- Karvinen, M. (2012). *Statistical pairs trading and analyst recommendations*. Master's Thesis, Aalto University, Department of Finance.
- Khan, S. (2010). Crude oil price shocks to emerging markets: evaluating the BRICs case. Available at SSRN 1617762.
- Khandani, A., & Lo, A. (2007). What happened to the quants in August 2007? Available at SSRN 1015987 .
- Kotzeva, M., & Bianka, P. (2006). Labour markets dynamics and characteristics in Bulgaria and Romania-Challenges for a successful integration in the European union. *The Vienna Institute for International Economic Studies*, 30. New Delhi.
- Kuan, C.-M., & Hornik, K. (1995). The generalized fluctuation test: A unifying view. *Econometric Reviews*, 14(2), 135-161.
- Kwitkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992, October–December). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1–3), 159–178.
- Lee, J., & Strazicich, M. C. (2003). Minimum LM unit root test with two structural breaks. *The Review of Economics and Statistics*, 85(4), 1082-1089.
- Lee, T.-H., & Tse, Y. (1996). Cointegration tests with conditional heteroskedasticity. *Journal of Econometrics*, 73(2), 401-410.

- Leykam, K. (2008). Cointegration and volatility in the European natural gas spot markets. *Unpublished manuscript Thesis (Masters) University of St. Gallen*, 1-55.
- List of stock exchanges. (n.d.). *List of stock exchanges*. Retrieved June 30, 2012, from wikinvest: [http://www.wikinvest.com/wiki/List\\_of\\_Stock\\_Exchanges](http://www.wikinvest.com/wiki/List_of_Stock_Exchanges)
- Liu, J., Wu, S., & Zidek, J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, 7, 497-525.
- Liu, M.-H., & Shrestha, K. M. (2008). Analysis of the long-term relationship between macroeconomic variables and the Chinese stock market using heteroscedastic cointegration. *Managerial Finance*, 34(11), 744-755.
- Lo, A. W., & MacKinlay, C. A. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies*, 1(1), 41-66.
- Lucas, A. (1997). Strategic and tactical asset allocation and the effect of long-run equilibrium relations. *Research Memorandum*, 42.
- Lucas, R. E. (1973). Wage inflation and the structure of regional unemployment: Comment. *Journal of money, credit and banking*, 5(1), 382-84.
- Lumsdaine, R. L., & Papell, D. H. (1997). Multiple trend breaks and the unit-root hypothesis. *The Review of Economics and Statistics*, 79(2), 212-218.
- Lütkepohl, H. (1993). *Introduction to multiple time series analysis*. Berlin Heidelberg New York Tokyo: Springer-Verlag.
- Lütkepohl, H. (2004). Recent advances in cointegration analysis. *Economics Working Papers*.
- Lütkepohl, H., Saikkone, P., & Trenkler, C. (2004). Testing for the cointegrating rank of a VAR process with level shift at unknown time. *Econometrica*, 72(2), 647-662.
- Mackinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle, & C. W. Granger, *Long-run economic relationships: Readings in cointegration* (pp. 266-76). New York.

- MarketsWiki. (2012, October 19). *MarketsWiki*. Retrieved July 5, 2012, from MarketsWiki: [http://www.marketswiki.com/mwiki/Main\\_Page](http://www.marketswiki.com/mwiki/Main_Page)
- Matthieu, S. (n.d.). *Estimation of vector error correction model (VECM) by EG or MLE*. Retrieved December 11, 2011, from R Graphical Manual: [http://rgm2.lab.nig.ac.jp/RGM2/func.php?rd\\_id=tsDyn:VECM](http://rgm2.lab.nig.ac.jp/RGM2/func.php?rd_id=tsDyn:VECM)
- Mavrakis, E. K., & Alexakis, C. A. (2008). Stock markets' linkages: An empirical investigation for long-term international diversification benefits. *International Research Journal of Finance and Economics*(21), 163-178.
- Maysami, C. R., Howe, C. L., & Hamzah, A. M. (2004). Relationship between macroeconomic variables and stock market indices: Cointegration evidence from stock exchange of Singapore's All-S sector indices. *Jurnal Pengurusan*(24), 47-77.
- Ministry of Finance. (2012). *The BRICS report: A study of Brazil, Russia, India, China and South Africa with special focus on synergies and complementarities*. New Delhi, India: Oxford University Press.
- Msimanga, N. L. (2010, December). Cointegration, causality and international portfolio deversification: Investigating potential benefits to a South African investor. *Thesis (Master)*, Rhodes University, 1-85.
- Nelson, C. R., & Plosser, C. I. (1982). Trends and random walks in microeconomic time series. *Journal of Monetary Economics*, 10, 139-62.
- Nelson, C. R., & Plosser, C. R. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10(2), 139-162.
- Nelson, C. R., & Plosser, C. R. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10(2), 139-162.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347-370.
- Ng, S., & Perron, P. (1995). Unit root tests in ARMA Models with data dependent methods for the selection of the truncation lag. *Journal of the american statistical association*, 90, 268-81.

- Ng, S., & Perron, P. (2001, November). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69(6), 1519-1554.
- Nielsen, H. B. (2005). *Non-stationary time series, cointegration and spurious regression*. Retrieved October 2011, from [http://www.econ.ku.dk/metrics/Econometrics2\\_05\\_II/Slides/10\\_cointegration\\_2pp.pdf](http://www.econ.ku.dk/metrics/Econometrics2_05_II/Slides/10_cointegration_2pp.pdf)
- Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, 84(405), 223-230.
- Omran, M. (2003). Time series analysis of the impact of real interest rates on stock market activity and liquidity in Egypt: Co-integration and error correction model approach. *International Journal of Business*, 8(3), 1083-4346.
- Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, 57(6), 1361-1401.
- Perron, P. (2006). Dealing with structural breaks. In K. Patterson, & T. C. Mills (Eds.), *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory* (Vol. 1, pp. 278–352). Basingstote, U.K.
- Perron, P., & Yabu, T. (2009). Testing for shifts in trend with an integrated or stationary noise component. *Journal of Business & Economic Statistics*, 27(3), 369-396.
- Perron, P., & Zhou, J. (2008). Testing jointly for structural changes in the error variance and coefficients of a linear regression model. *Boston University - Department of Economics - Working Papers Series*.
- Pesavento, Elena. (2007). Residuals-based tests for the null of no-cointegration: An analytical comparison. *Journal of Time Series Analysis*, 28(1), 111-137.
- Pfaff, B. (2008). *Analysis of integrated and cointegrated time series with R*. New york: Springer Science+Business Media.
- Phillips, P. C., & Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica: Journal of the Econometric Society*, 165-193.

- Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-46.
- Ploberger, W., & Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica*, 60(2), 271-285.
- Pokorny, M. (1987). *An introduction to econometrics*. Oxford: Blackwell.
- Rahbek, A., Hansen, E., & Dennis, J. G. (2002). ARCH innovations and their impact on cointegration rank testing. *Centre for Analytical Finance*(Working Paper 22).
- Rahman, S. (2005). An alternative estimation to spurious regression model. *Nanyang Technological University, School of Humanities and Social Sciences, Economic Growth centre*.
- Ramirez-Beltran, N., & Olivares, L. (1999). Statistical analysis of drug stability. *Computer and Industrial Engineering*, 37(1-2), 351-354.
- Ramos, F. F. (1996). VAR priors : success or lack of a decent macroeconomic theory? *EconWPA*, 18.
- Reade, J. (2005). The cointegration VAR methodology. 1-72.
- Royal society publishing. (n.d.). *Royalsocietypublishing*. Retrieved October 2, 2012, from Royal Society: <http://rspa.royalsocietypublishing.org>
- RTS Index (RTSI). (n.d.). *RTS Index (RTSI)*. Retrieved July 4, 2012, from RTS Exchange: <http://www.rts.ru/en/index/rtsi/>
- Rycx, F. (2012). *Méthodologie quantitative*. Retrieved October 2012, from Applied Econometrics Seminar: <http://homepages.ulb.ac.be/~frycx/>
- Sah, N. A. (2009). Stock market seasonality: A study of the Indian stock market. Retrieved 2011, from <http://ssrn.com/abstract>
- Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3), 599-607.

- Saikkonen, P., & Lütkepohl, H. (2000). Testing for the cointegrating rank of a VAR process with structural shifts. *Journal of Business & Economic Statistics*, 18(4), 451-464.
- Sargan, J. D. (1964). Wages and prices in the United Kingdom: A study in econometric methodology (with discussion). *Econometric Analysis for National Economic Planning*, 16, 25-63.
- Sargent, T. J. (1978, January). A note on maximum likelihood estimation of the rational expectations model of the term structure. *Journal of monetary economics*, 5(1), 133-43.
- Schmidt, A. D. (2008). Pairs trading: a cointegration approach. *Thesis (Honours), University of Sydney*, 1-130.
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464.
- Schwert, W. G. (1989, April). Test for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics*, 7(2), 147-59.
- Sentana, E. (1995). Quadratic ARCH models. *The Review of Economic Studies*, 62(4), 639-661.
- Sharpe, W. F. (1966). Mutual fund performance. *Journal of business*, 119-138.
- Sharpe, W. F. (1975). Adjusting for risk in portfolio performance measurement. *The Journal of Portfolio Management*, 29-34.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1-48.
- Song, N., & Chang, S. J. (2003). Nonstationarity and its consequences in modeling the southern timber market. *Bugs, budgets, mergers, and fire: disturbance economics*, pp. 241-249.
- SouthAfrica.info. (n.d.). *SouthAfrica.info*. Retrieved July 3, 2012, from SouthAfrica.info: <http://www.southafrica.info/>

- Steenbarger, B. N. (2003). *The psychology of trading: Tools and techniques for minding the markets*. New Jersey: John Wiley & sons.
- Stock, J. H., & Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica: Journal of the Econometric Society*, 783-820.
- Stock, J. H., & Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61(4), 783-820.
- Stock, J. H., & Watson, M. W. (1998). Diffusion indexes. *National Bureau of Economic Research*(w6702).
- Subramanian, U. (2008). Cointegration of stock markets in East Asia. *European Journal of Economics*, 14, 450-2275.
- Sudak, D., & Suslova, O. (2009). Behavioral statistical arbitrage. *Unpublished manuscript Thesis(Masters) University of Lausanne*, 1-77.
- Suwardi, E. (2004, February). Exploring the relationship between market values and accounting numbers of firms listed in an emerging market. *Dissertation (PhD), Queensland University of Technology*, 1-235.
- Suwardi, E. (2004, February). Exploring the relationship between market values and accounting numbers of firms listed in an emerging market. *Doctoral Dissertation, Queensland University of Technology*, 1-235.
- Takayanagi, S., & Ishikawa, K. (2012). classical pair trading based on cointegration in finance. *R package*.
- Thomas, D. E. (2006). International diversification and firm performance in Mexican firms: A curvilinear relationship? *Journal of business research*, 59(4), 501-507.
- Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 225-250.
- Triantafyllopoulos, K., & Montana, G. (2011). Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science*, 8(1), 23-49.

- Uhlenbeck, G. E., & Salomon, L. (1930). On the theory of the Brownian motion. *Physical Review*, 36(5), 823.
- West, G. (2008). Interest rate derivatives in the South African market based on the prime rate. *Journal for studies in economics and econometrics*, 32(1), 1-11.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838.
- Wilson, E. J., & Marashdeh, H. A. (2007). Are co-integrated stock prices consistent with the efficient market hypothesis? *Economic Record*(83 ), S87-S93.
- Yahoo finance. (n.d.). *Yahoo finance*. Retrieved June 30, 2012, from Yahoo: <http://finance.yahoo.com/>
- Yamamoto, T. (1996). A simple approach to the statistical inference in linear time series models which may have some unit roots. *Hitotsubashi Journal of Economics*, 37 (2), 87–100.
- Yao, S., Luo, D., & Morgan, S. (2008). Impact of the US credit crunch and housing market crisis on China. *China Policy Institute & Centre for Global Finance* , 1-36.
- Yule, U. G. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *Journal of the royal statistical society*, 89(1), 1-63.
- Zeileis, A. (2001). strucchange: Testing for structural change in linear regression relationships. *R News*, 1(3), 8-11.
- Zeileis, A. (2006). Implementing a class of structural change tests: An econometric computing approach. *Computational Statistics and Data Analysis*, 50(11), 2987 – 3008.
- Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2001). strucchange: Testing for structural change. *Journal of Statistical Software*, 7(2), 1-38.

- Zeileis, A., Kleiber, C., & Kramer, W. (2003). Testing and dating of structural changes in practice. *Elsevier B.V.*, 109–123.
- Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2001). Strucchange: An R package for testing for structural change in linear regression models. *Journal of statistical software*, 7(2), 1-38.
- Zhang, D., Dickinson, D. G., & Barasi, M. (2006). Structural breaks, Cointegration and B Share Discount in Chinese Stock Market. *Ecomod.net*.
- Zhou, Q., & Zhou, Z. (2005). Stock returns, volatility, and cointegration among Chinese. *China and World Economy*, 13, 106-22.
- Zivot, E., & Andrews, D. W. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics*, 10(3), 251-270.



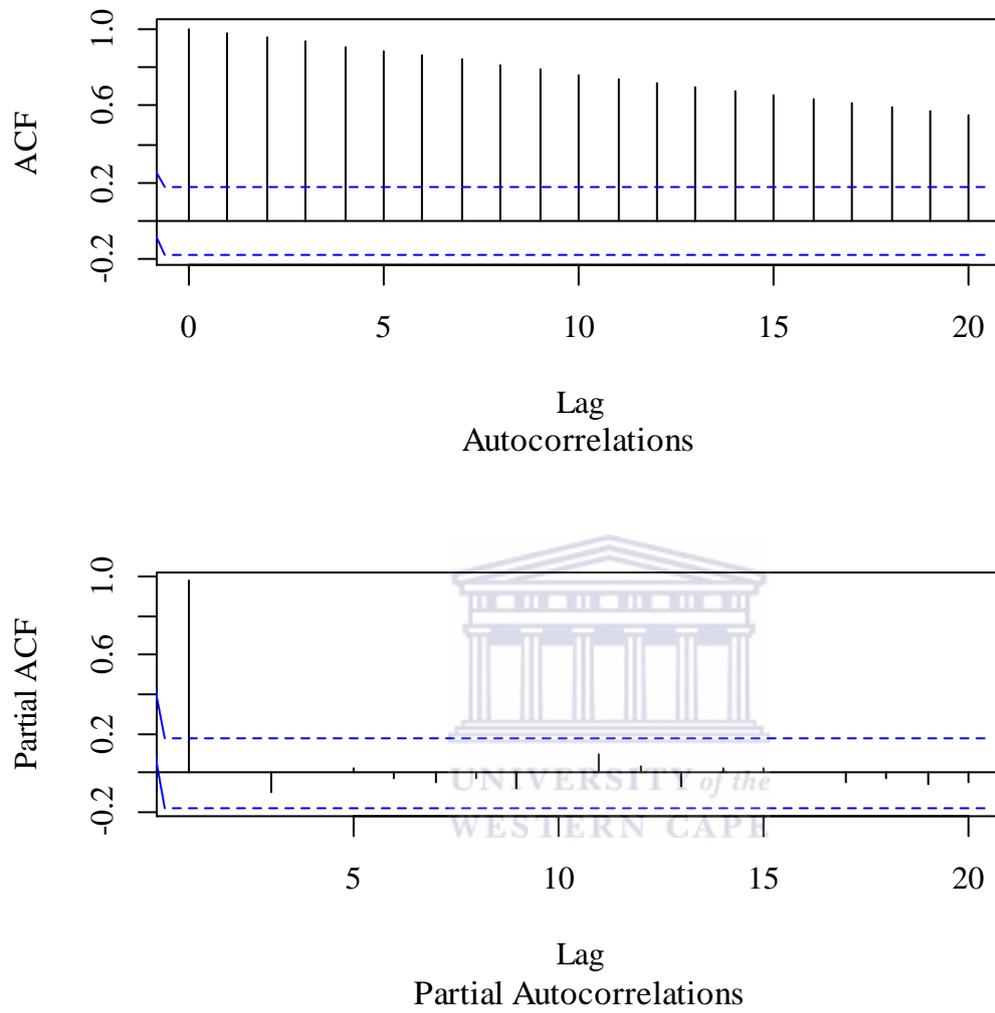
APPENDIX A

Figure A1: JSE ACF and PACF plots

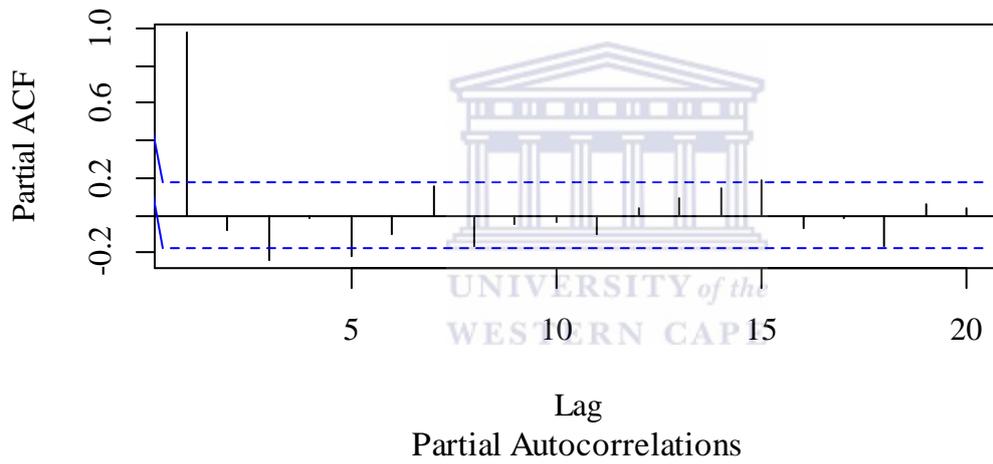
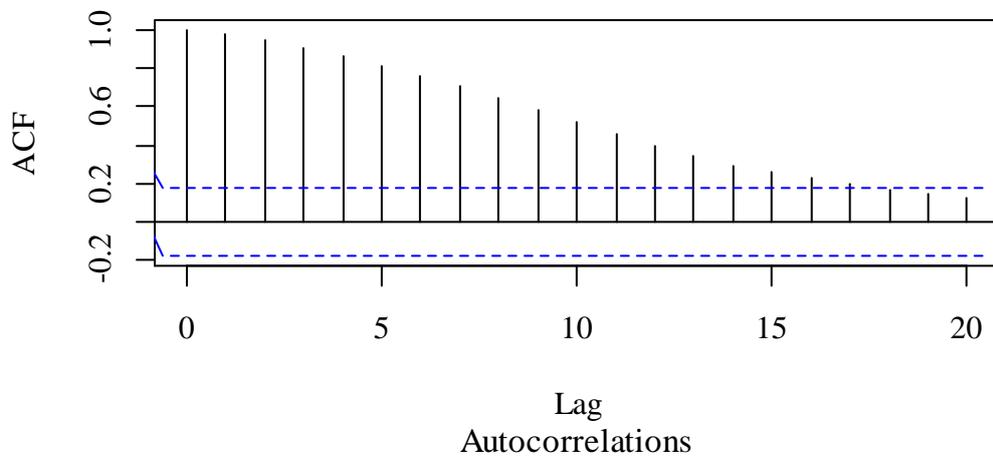


Figure A2: SSE ACF and PACF plots

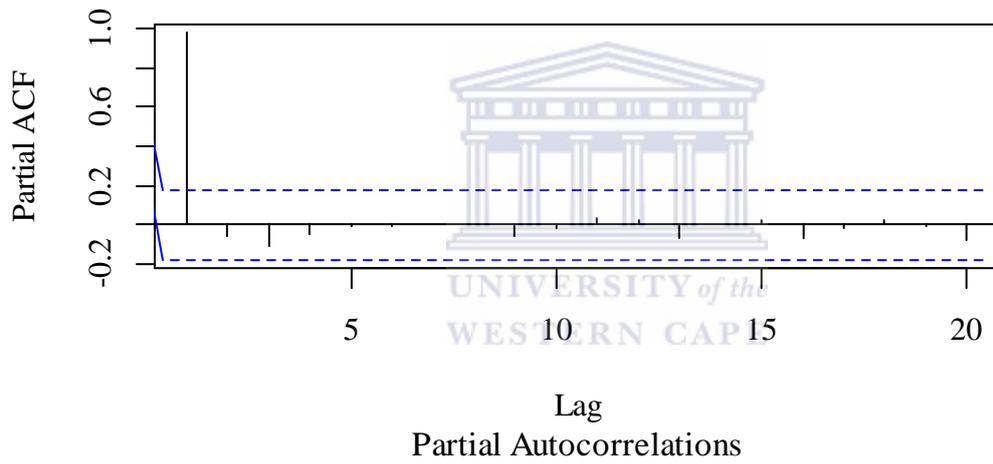
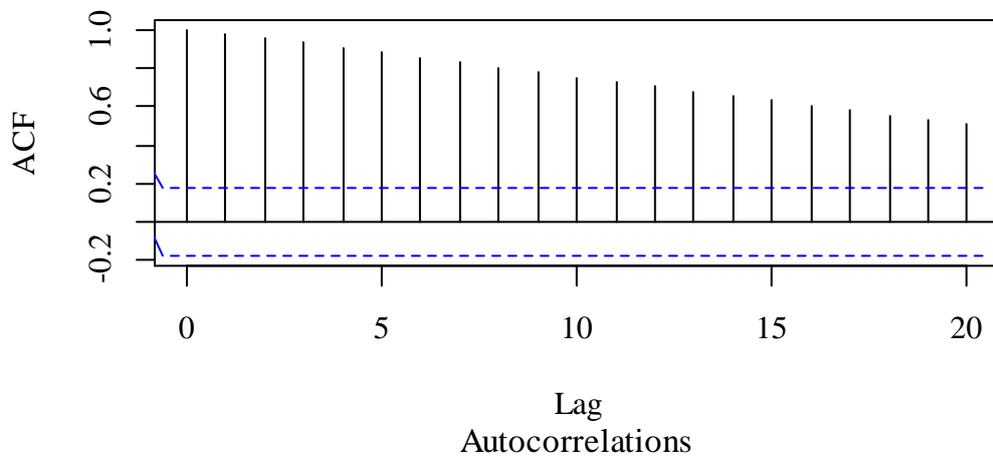


Figure A3: SENSEX ACF and PACF plots

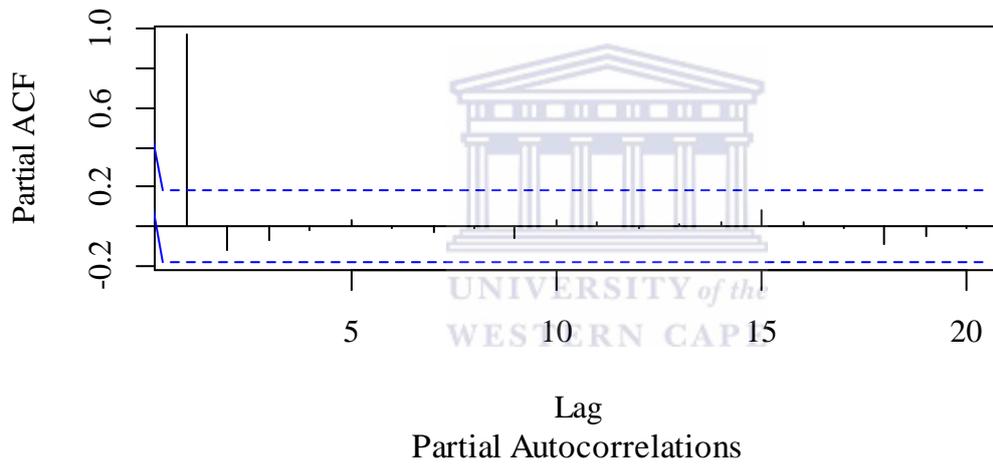
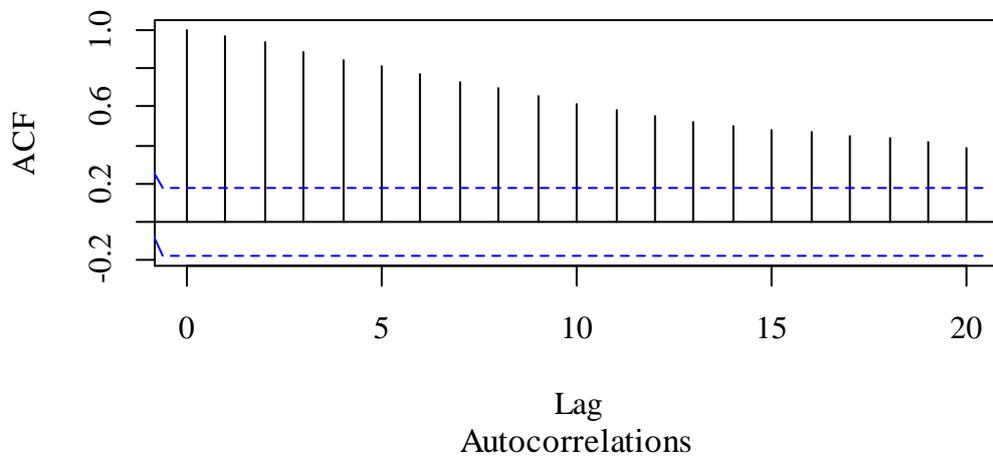


Figure A4: RTSI ACF and PACF plots

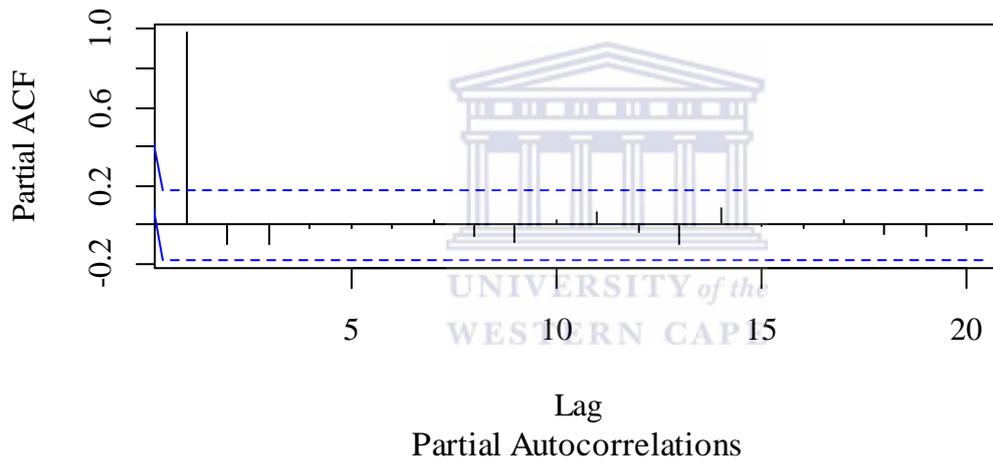
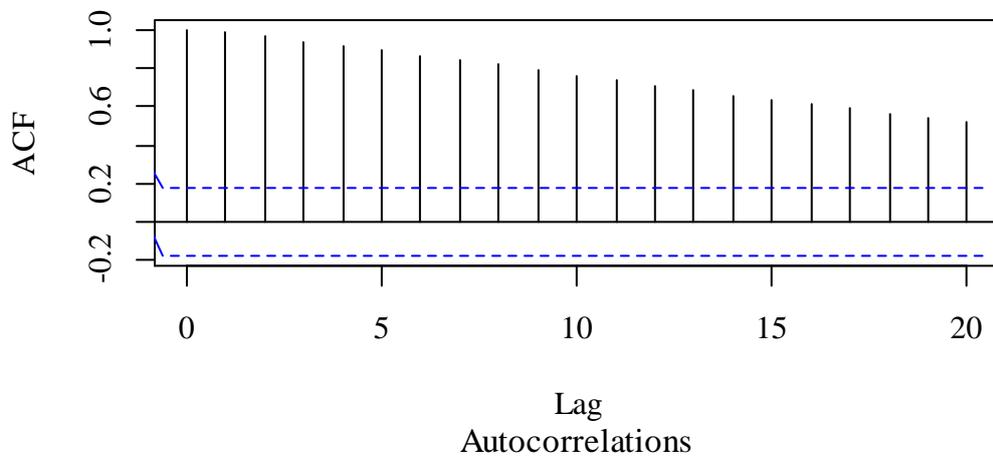


Figure A5: BOVESPA ACF and PACF plots

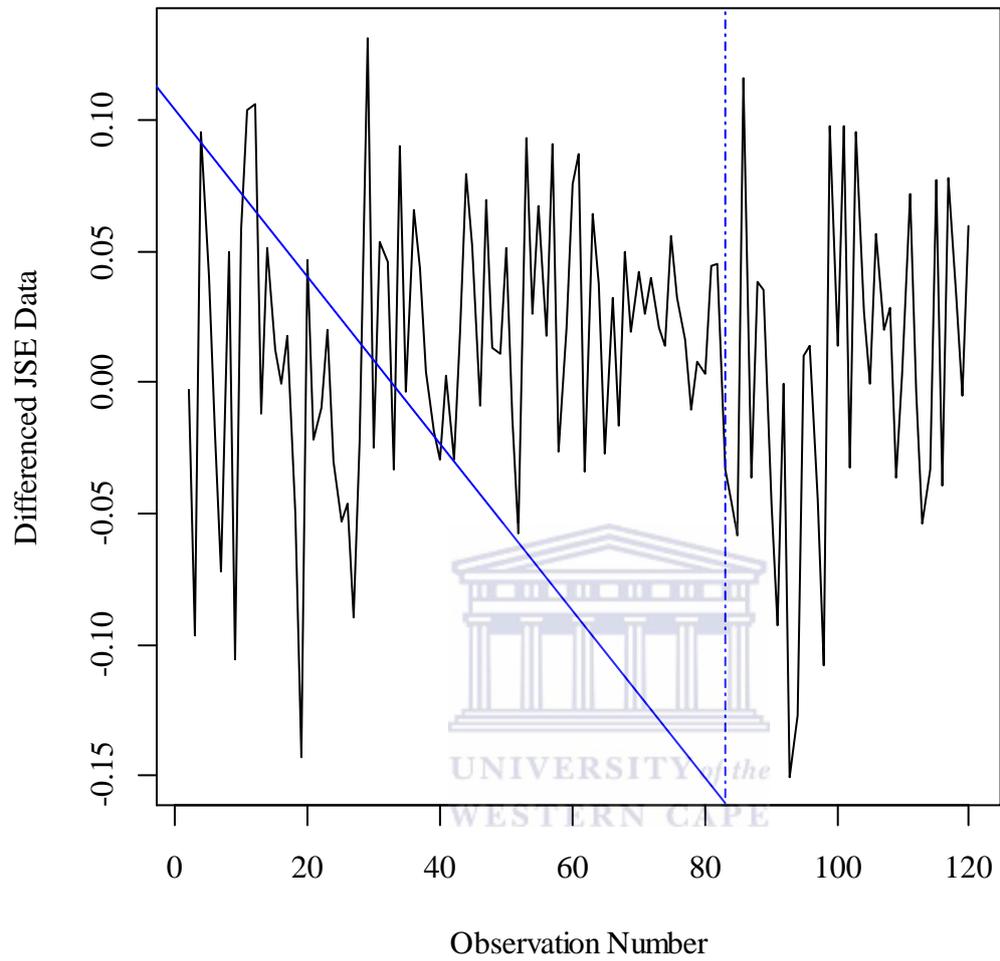
APPENDIX B

Figure B1: Plot of differenced JSE data

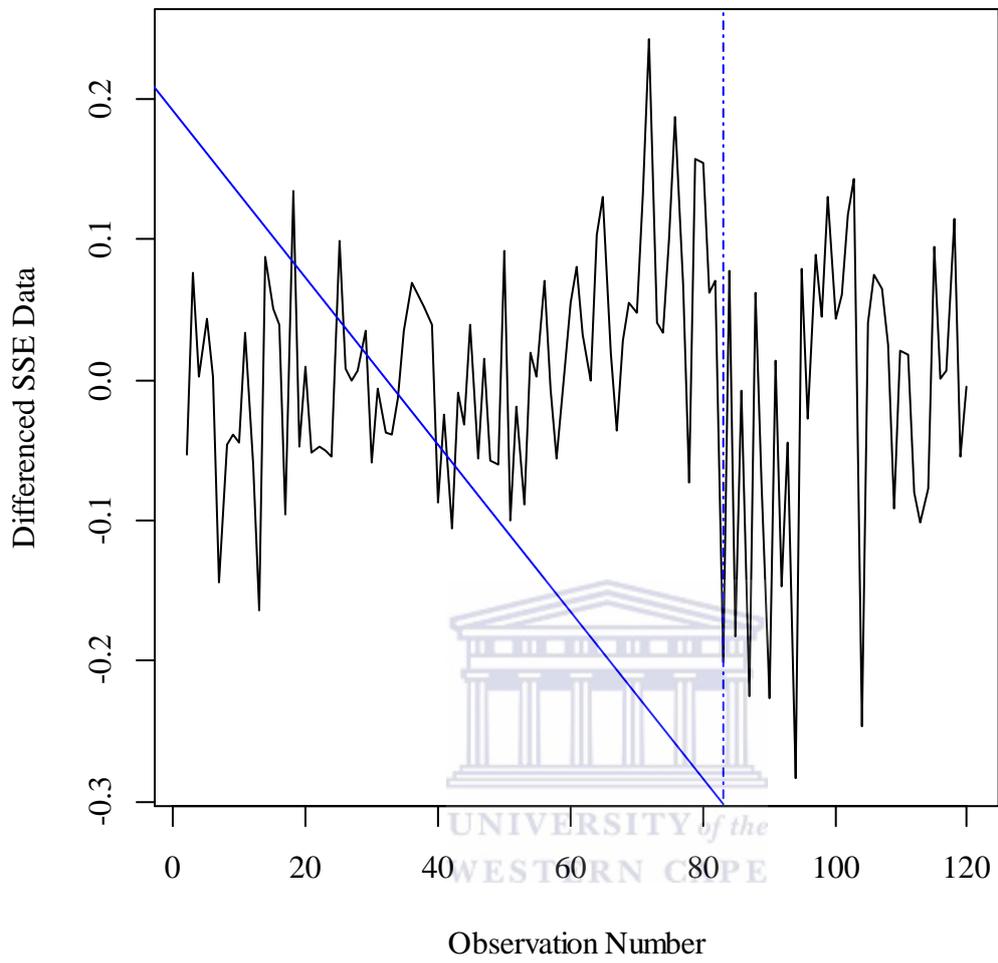


Figure B2: Plot of differenced SSE data

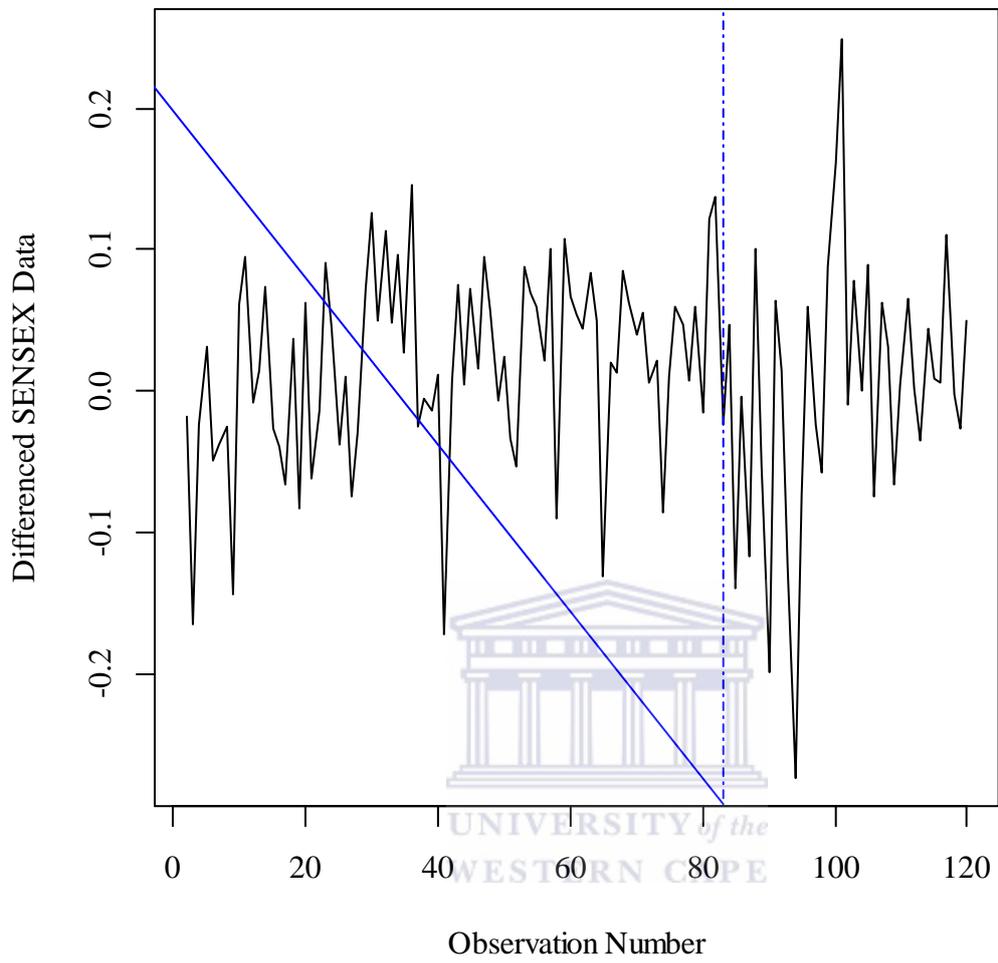


Figure B3: Plot of differenced SENSEX data

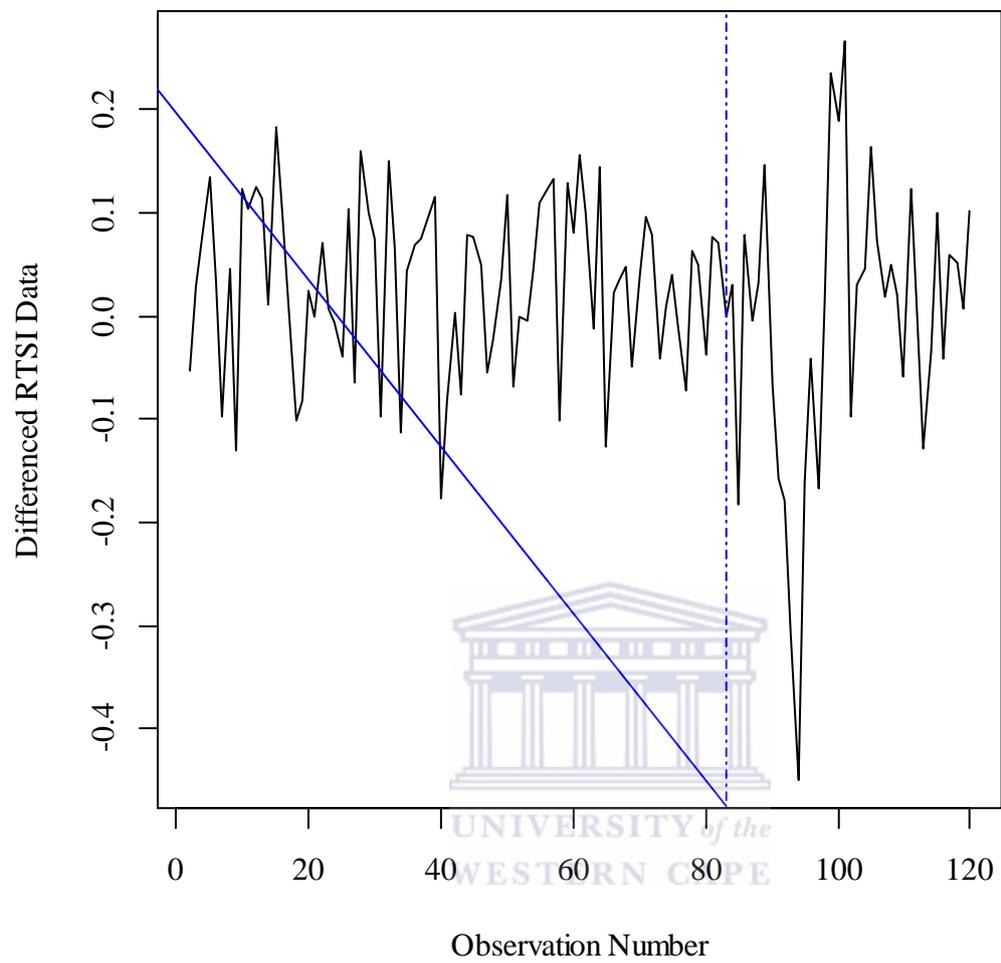


Figure B4: Plot of differenced RTSI data

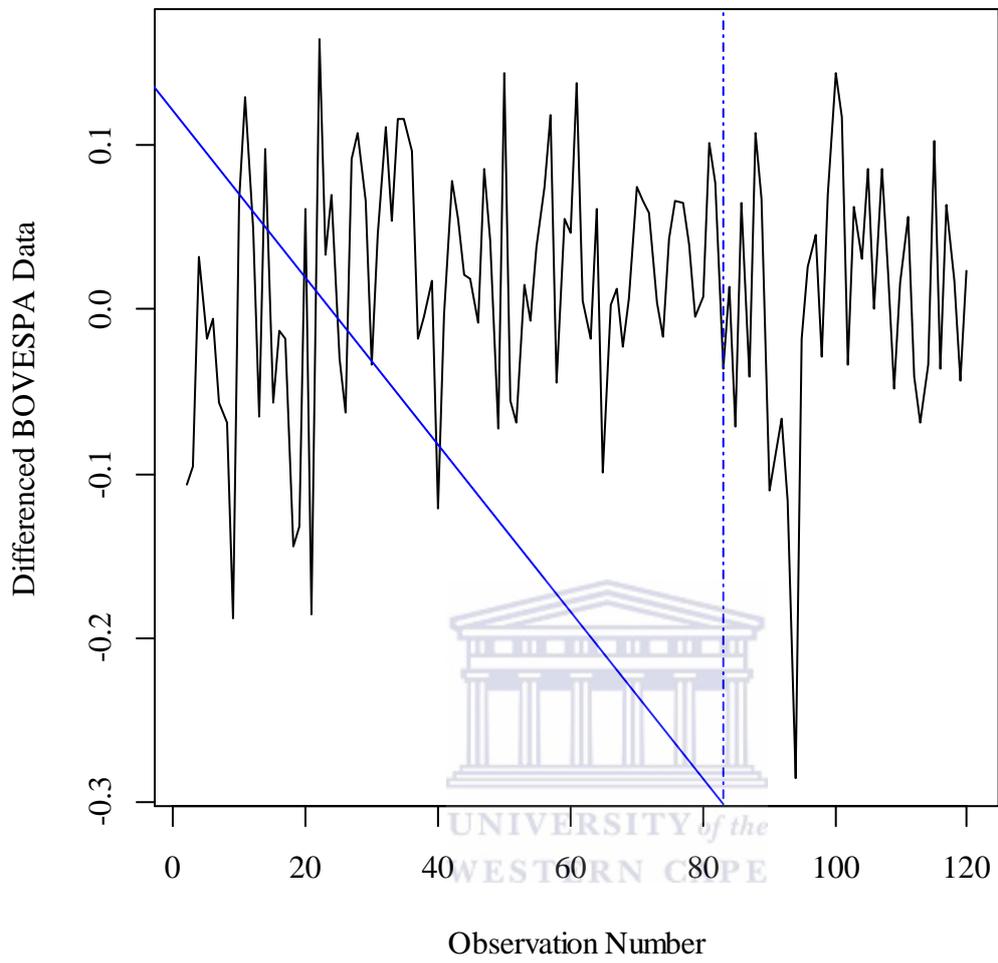


Figure B5: Plot of differenced BOVESPA data

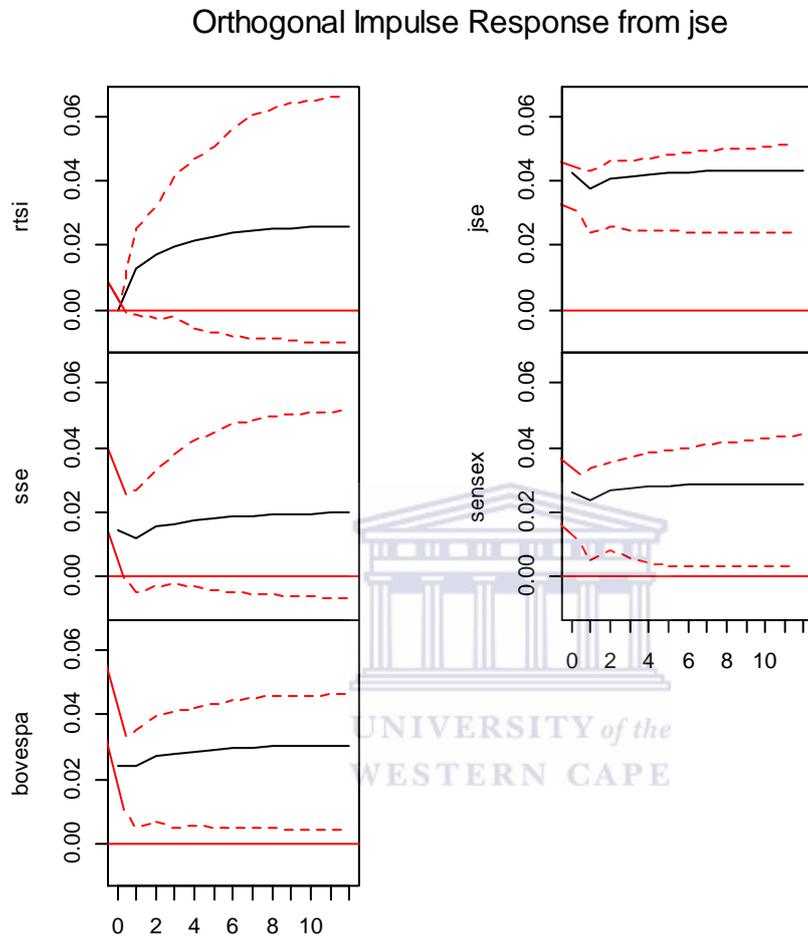
APPENDIX C

Figure C1: Responses to a stock price shock in the JSE with 95% bootstrap confidence interval and runs limited to 12 periods.

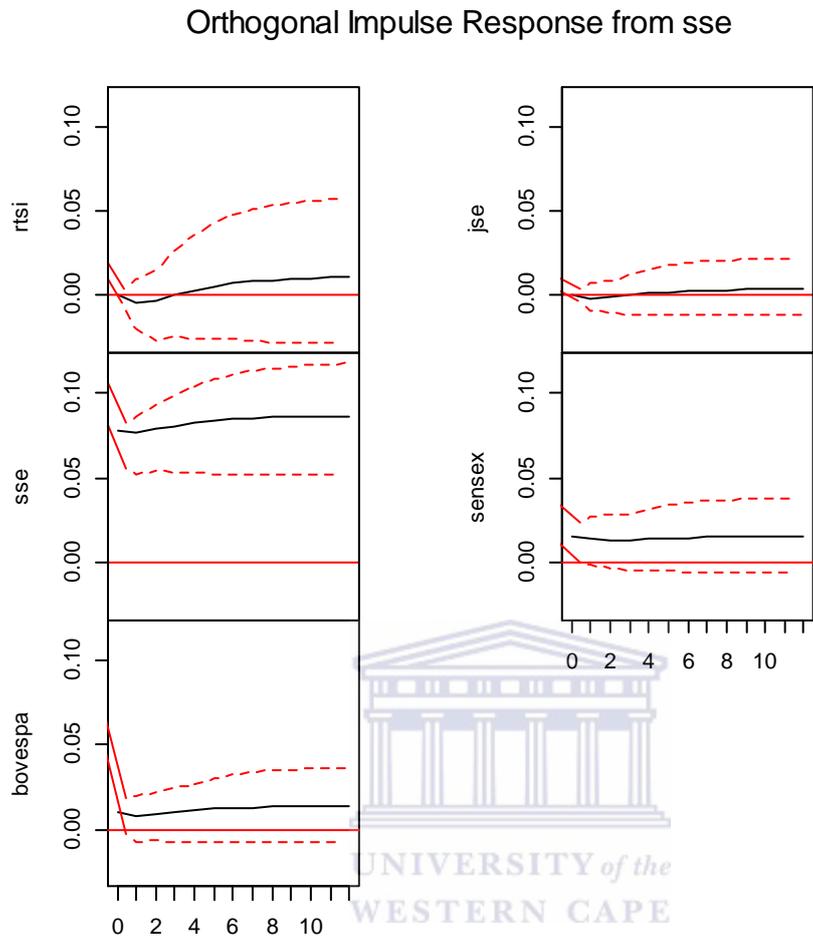


Figure C2: Responses to a stock price shock in the SSE with 95% bootstrap confidence interval and runs limited to 12 periods.

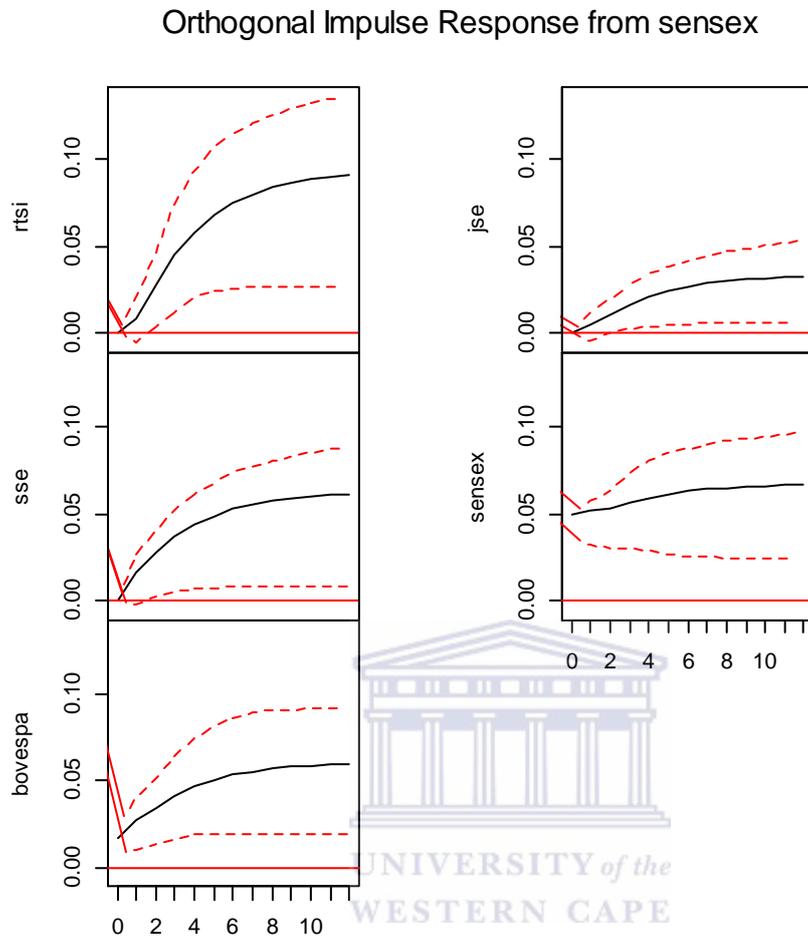


Figure C3: Responses to a stock price shock in the SENSEX with 95% bootstrap confidence interval and runs limited to 12 periods.

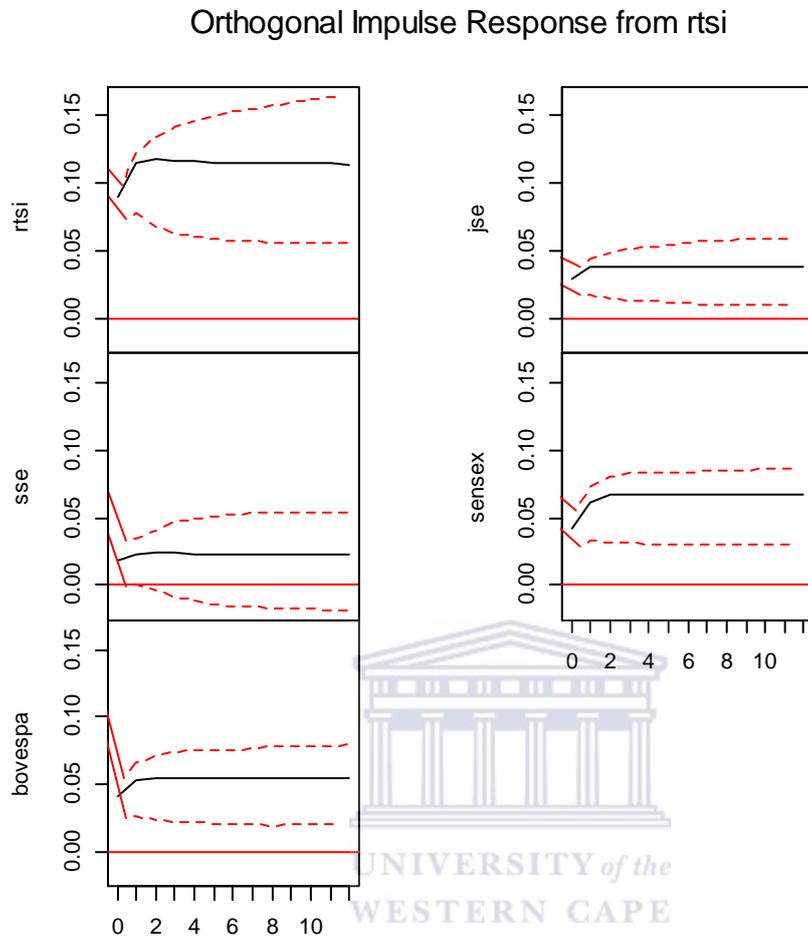


Figure C4: Responses to a stock price shock in the RTSI with 95% bootstrap confidence interval and runs limited to 12 periods.

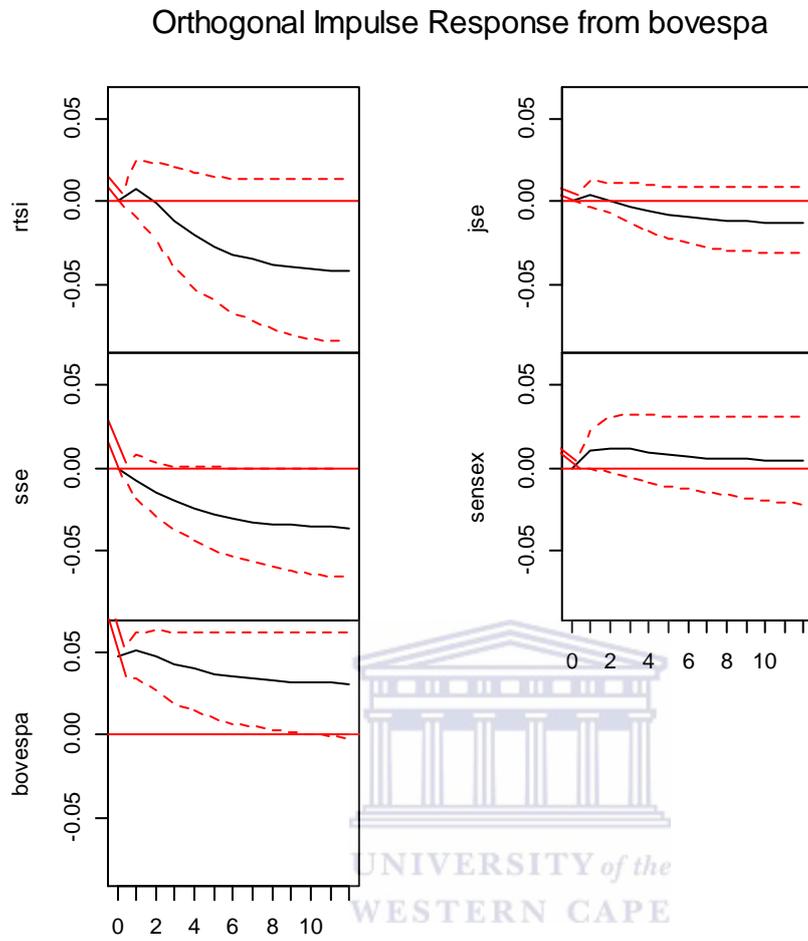
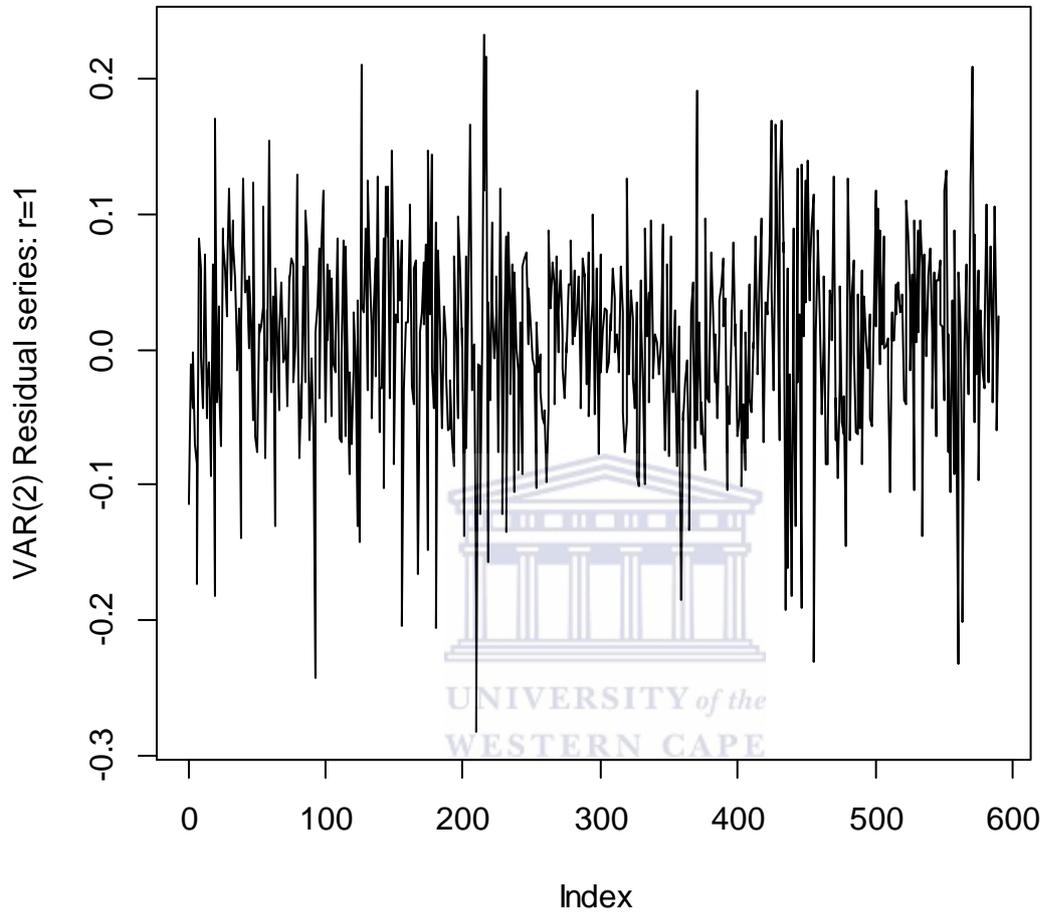


Figure C5: Responses to a stock price shock in the BOVESPA with 95% bootstrap confidence interval and runs limited to 12 periods.

APPENDIX DFigure D1: Residual series of a VAR(2) model with  $r=1$

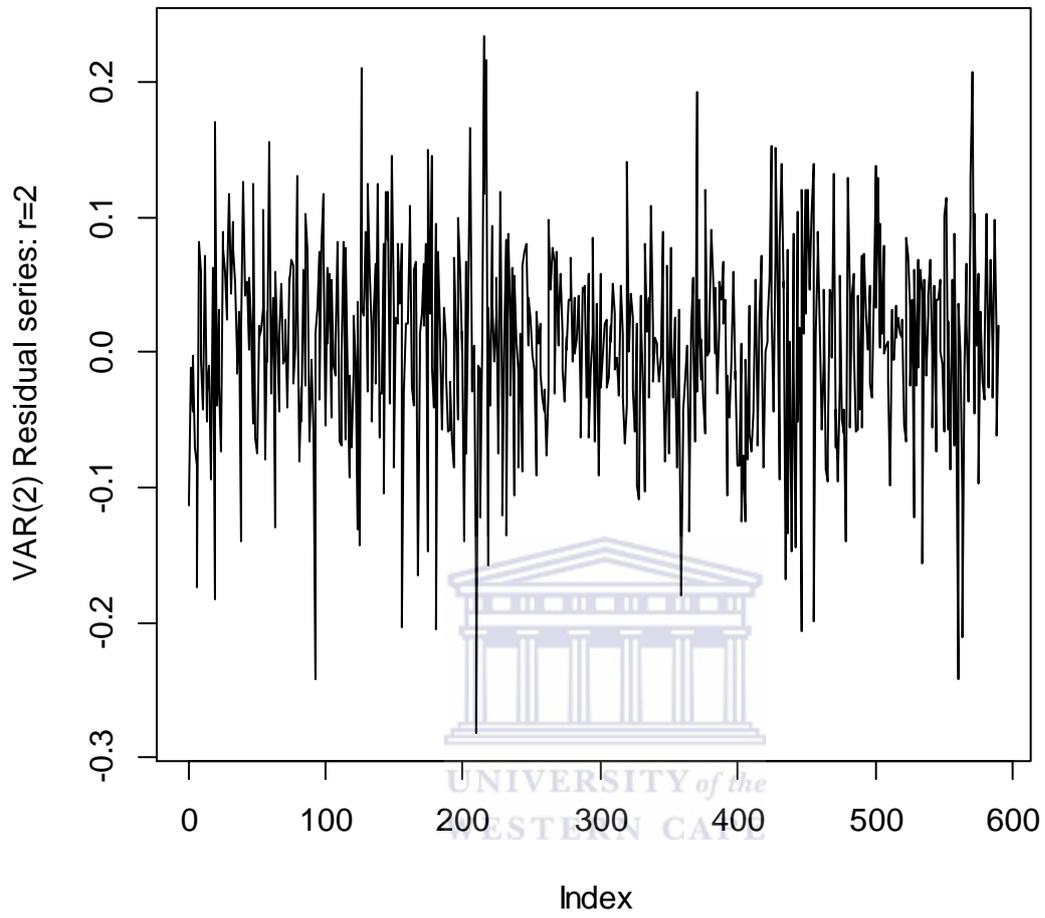


Figure D2: Residual series of a VAR(2) model with  $r=2$

## APPENDIX E

R code used for the analysis

*#Load Relevant Libraries*

```
library(urca)
```

```
library(vars)
```

```
library(fGarch)
```

```
library(PairTrading)
```

```
library(fTrading)
```

```
library(Hmisc)
```

*#Read in the data*

```
mydata=read.csv(file.choose(),header=T)
```

```
attach(mydata)
```

*# verify data is numeric*

```
data.entry(jse,sse,sensex,rtsi,bovespa)
```

*#Identify data as time series*

```
jse=ts(jse)
```

```
sse=ts(sse)
```

```
sensex=ts(sensex)
```

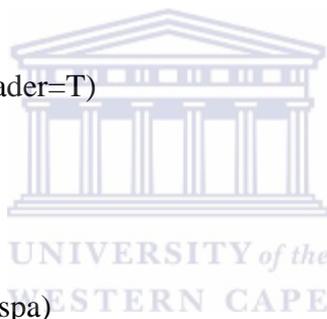
```
rtsi=ts(rtsi)
```

```
bovespa=ts(bovespa)
```

*##testing for correlation between indices*

```
rcorr(as.matrix(data.frame(jse,sse,sensex,rtsi,bovespa), type="pearson"))
```

*##unit root tests: we use illustrate the testing processes using jse as an example otherwise each code is repeated 4 other times to accommodate each stock variable in levels with its own lag structure*



*##Testing without accounting for breaks: these tests are repeated to for the log-transformed prices as well*

*##ADF test*

```
summary(ur.df(jse,lags=1,type="drift"))
```

*##KPSS test*

```
ur.kpss(jse, type = "mu",
        use.lag = 9)
```

*##PP test*

```
pp.test(jse)
```

*##Tests that account for breaks*

*##Andrew-Zivot test*

*##supF test*

```
supF.jse=Fstats(jse~1) ## regressed on the mean
sctest(supF.jse)
```

*##Rec-CUSUM test*

```
sc.jse=efp(jse~1) ## regressed on the mean
```

```
sctest(sc.jse)
```

*#next create a 'Date' variable to index the zoo object to be created*

```
Date= as.Date(a$Date)
```

*#creating objects for plots in raw, log pricing and in differences respectively*

```
rawdata = zoo(data.frame(jse, sensex, sse, rtsi, bovespa),Date)
```

```
log_data = zoo(data.frame(ljse, lsse, lsensex, lrtsi, lbovespa), Date)
```

```
# lstock=log(stock)
```

```
differenced_log_data = zoo(data.frame(diff(ljse), diff(lsse), diff(lsensex), diff(lrtsi),
diff(lbovespa)),Date)
```



*#creating a data frame of the stock variables and choosing the first variable in the data frame to normalize on when estimating the vecm.*

```
d=data.frame(rtsi, bovespa, jse, sse, sensex)
```

*# let rtsi, bovespa, jse, sse and sensex denote log-transformed data henceforth*

*# determining the optimal number of lags for the cointegration rank test*

```
lag.num=VARselect(na.omit(d),lag.max=10,"const")
```

```
lag.num
```

*#then create the zoo object of the 5 stock variables indexed on the Date variable*

```
dzoo=zoo(d,Date)
```

*#create frame with the Structural Shift and transitory dummy variable*

```
dummy=data.frame(sdummy,dum1)
```

*#Estimating the VECM, cointegrating and loading vectors with shift, transitory and 11 seasonal dummies.*

```
vecm<- ca.jo(d, type = "trace",ecdet = "const", K = 2, spec = "longrun",dumvar=dummiz,
season=12)
```

```
vecmE<- ca.jo(d, type = "eigen",ecdet = "const", K = 2, spec = "longrun",dumvar=dummiz,
season=12)
```

*#OLS estimation of the regression estimates*

```
vecjorls<- cajorls(vecm, r = 2)
```

```
vecjorlsE<- cajorls(vecmE, r = 1)
```

*#Transform VECM to VAR(2)*

```
v2v=vec2var(vecm, r = 2)
```

```
v2vE=vec2var(vecmE, r = 1)
```

*#Calculating the t-values*

```
alpha<- coef(vecjorls$rlm)[1, ]
```

```
beta<- vecjorls$beta
resids<- resid(vecjorls$rlm)
N <- nrow(resids)
sigma<- crossprod(resids) / N
```

```
## t-stats for alpha
```

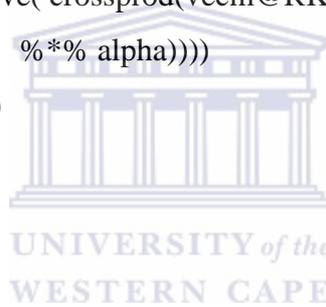
```
alpha.se <- sqrt(solve(crossprod(cbind(vecm@ZK %*% beta, vecm@Z1)))
[1, 1]* diag(sigma))
alpha.t <- alpha / alpha.se
```

```
## t-stats for beta
```

```
beta.se <- sqrt(diag(kronecker(solve( crossprod(vecm@RK[, -1])),
solve(t(alpha) %*% solve(sigma) %*% alpha))))
beta.t <- c(NA, beta[-1] / beta.se)
```

```
#Output
```

```
alpha.t
beta.t
```



### Code used for the Appendices

```
par(mar=c(5.1,4.1,2,2)) # determine the plot margins
plot3 = ts(rawdata, start=2001,frequency=12)
```

```
#plot raw data (figure 3)
```

```
ts.plot(plot3,main="",xlab="")
```

```
#use Times New Roman font in all text on the graph
```

```
windowsFonts(A=windowsFont("Times New Roman"))
```

```
## Figure 3: Time series plots of the raw monthly prices
```

```
title("",sub="          ",font.sub=1,cex.sub=1.1,ylab="Index          monthly          prices
(raw)",xlab="Time",cex.lab=0.9, family="A")
```

```

lines(plot3[,1],col="red")

lines(plot3[,2],col="green")

lines(plot3[,3],col="blue")

lines(plot3[,4],col="orange")

lines(plot3[,5],col="black")

legend("topleft",c("JSE","SSE","SENSEX","RTSI","BOVESPA"),cex=0.65,col=c("red","green",
"blue","orange","black"),lty=1,title="Stock Exchanges")

plot4 = ts(log_data, start=2001,frequency=12)

#plot log transformed data (figure 4)

ts.plot(plot4,main="",xlab="")

windowsFonts(A=windowsFont("Times New Roman"))

##Figure 4: Time series plots of logged monthly prices

title("",sub="Index monthly prices",font.sub=1,cex.sub=1.1,ylab="Index monthly prices
(logs)",xlab="Time",cex.lab=0.9, family="A")

lines(plot4[,1],col="red")

lines(plot4[,2],col="green")

lines(plot4[,3],col="blue")

lines(plot4[,4],col="orange")

lines(plot4[,5],col="black")

legend("bottomright",c("JSE","SSE","SENSEX","RTSI","BOVESPA"),cex=0.65,col=c("red",
"green","blue","orange","black"),lty=1,title="Stock Exchanges")

plot5 = ts(differenced_log_data, start=2001, frequency=12)

```

```
#plot the differenced log data (figure 5)
```

```
ts.plot(plot5,main="",xlab="")
```

```
windowsFonts(A=windowsFont("Times New Roman"))
```

```
## Figure 5: Time series plots of log-differenced monthly prices
```

```
title("",sub="      ",font.sub=1,cex.sub=1.1,ylab="Index      monthly      prices      (log-  
differenced)",xlab="Time",cex.lab=0.9, family="A")
```

```
lines(plot5[,1],col="red")
```

```
lines(plot5[,2],col="green")
```

```
lines(plot5[,3],col="blue")
```

```
lines(plot5[,4],col="orange")
```

```
lines(plot5[,5],col="black")
```

```
legend("bottomright",c("JSE","SSE","SENSEX","RTTI","BOVESPA"),cex=0.65,col=c("red  
","green","blue","orange","black"),lty=1,title="Stock Exchanges")
```

```
op<- par(usr=c(0,1,0,1), xpd=NA) # Reset the coordinates
```

### APPENDIX A: Partial and Autocorrelation plot codes

```
windowsFonts(A=windowsFont("Times New Roman"))
```

```
op = par(oma=c(4,1,1,1)+.1,mfrow=c(2,1)) # Room for the title and legend
```

```
acf(jse,main="Autocorrelations", cex=.9,cex.main=0.9,family="A")
```

```
pacf(jse,main="Partial Autocorrelations", cex=.9,cex.main=0.9, family="A")
```

```
par(op)
```

```
## Figure A1: JSE ACF and PACF plots
```

```
mtext("", font=1,cex=1.1, side=1,family="A")
```

```
op<- par(usr=c(0,1,0,1), xpd=NA) # Reset the coordinates
```

```
dev.off()
```

```

op = par(oma=c(4,0,0,0)+.1,mfrow=c(2,1))
acf(sse,main="Autocorrelations", cex=.9,cex.main=0.9,family="A")
pacf(sse,main="Partial Autocorrelations", cex=.9,cex.main=0.9,family="A")
par(op)
## Figure A2: SSE ACF and PACF plots
mtext("", font=1,cex=1.1, side=1,family="A")
op<- par(usr=c(0,1,0,1), xpd=NA)
dev.off()

```

```

op = par(oma=c(4,0,0,0)+.1,mfrow=c(2,1))
acf(sensex,main="Autocorrelations", cex=.9,cex.main=0.9,family="A")
pacf(sensex,main="Partial Autocorrelations", cex=.9,cex.main=0.9,family="A")
par(op)
## Figure A3: SENSEX ACF and PACF plots
mtext("", font=1,cex=1.1, side=1,family="A")
op<- par(usr=c(0,1,0,1), xpd=NA)
dev.off()

```



```

op = par(oma=c(4,0,0,0)+.1,mfrow=c(2,1))
acf(rtsi,main="Autocorrelations", cex=.9,cex.main=0.9,family="A")
pacf(rtsi,main="Partial Autocorrelations", cex=.9,cex.main=0.9,family="A")
par(op)
##Figure A4: RTSI ACF and PACF plots
mtext("", font=1,cex=1.1, side=1,family="A")
op<- par(usr=c(0,1,0,1), xpd=NA)
dev.off()

```

```

op = par(oma=c(4,0,0,0)+.1,mfrow=c(2,1))
acf(bovespa,main="Autocorrelations", cex=.9,cex.main=0.9,family="A")
pacf(bovespa,main="Partial Autocorrelations", cex=.9,cex.main=0.9,family="A")
par(op)
## Figure A5: BOVESPA ACF and PACF plots
mtext("Figure A5: BOVESPA ACF and PACF plots", font=1,cex=1.1, side=1,family="A")

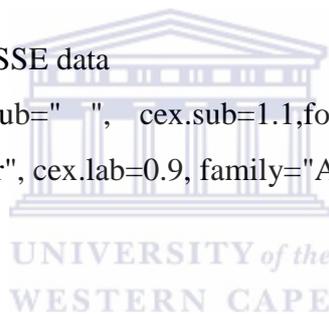
```

```
op<- par(usr=c(0,1,0,1), xpd=NA)
dev.off()
```

APPENDIX B: Individual time series plots of the differenced data with identification of the "blips" for structural shifts

```
windowsFonts(A=windowsFont("Times New Roman"))
par(oma=c(2,0,0,0))
## Figure B1: Plot of differenced JSE data
plot(diff(ts(jse)) , main=" ",sub=" ", cex.sub=1.1,font.sub=1,ylab="Differenced JSE
Data",xlab="Observation Number", cex.lab=0.9, family="A")
abline(v=83,lty=4,col="blue")
```

```
##Figure B2: Plot of differenced SSE data
plot(diff(ts(sse)) , main=" ",sub=" ", cex.sub=1.1,font.sub=1,ylab="Differenced SSE
Data",xlab="Observation Number", cex.lab=0.9, family="A")
abline(v=83,lty=4,col="blue")
```



```
## Figure B3: Plot of differenced SENSEX data
plot(diff(ts(sensex)) , main=" ",sub=" ", cex.sub=1.1,font.sub=1,ylab="Differenced SENSEX
Data",xlab="Observation Number", cex.lab=0.9, family="A")
abline(v=83,lty=4,col="blue")
```

```
## Figure B4: Plot of differenced RTSI data
plot(diff(ts(rtsi)) , main=" ",sub=" ", cex.sub=1.1,font.sub=1,ylab="Differenced RTSI
Data",xlab="Observation Number", cex.lab=0.9, family="A")
abline(v=83,lty=4,col="blue")
```

```
## Figure B5: Plot of differenced BOVESPA data
plot(diff(ts(bovespa)) , main=" ",sub=" ", cex.sub=1.1,font.sub=1,ylab="Differenced
BOVESPA Data",xlab="Observation Number", cex.lab=0.9, family="A")
abline(v=83,lty=4,col="blue")
```

APPENDIX C: Impulse response functions

##the code below is repeated 5 times to accommodate impulses introduced to all 5 variables.

```
irf(v2v, impulse = "stock_variable_name")
```

APPENDIX D: VAR(2) residual plots

```
plot(as.vector(v2v$resid), ylab="VAR(2) residual plot: r=2")
```

```
plot(as.vector(v2vE$resid), ylab="VAR(2) residual plot: r=1")
```

StatArb R Code: Pairs trading

```
p1=data.frame(sensex,rtsi)
```

```
p2=data.frame(rtsi,bovespa)
```

```
plot1=ts(p1,start=2001, end=2010,frequency=12)
```

```
plot2=ts(p2,start=2001, end=2010,frequency=12)
```

```
par(mar=c(5,2,1,1))
```

```
ts.plot(plot1,main="",xlab="")
```

```
## Figure 6: SENSEX and RTSI monthly stock prices
```

```
title("",sub="          ",font.sub=1,cex.sub=1.1,ylab="          Monthly          stock  
prices",xlab="Time",cex.lab=0.9, family="A")
```

```
lines(plot1[,1],col="blue")
```

```
lines(plot1[,2],col="orange")
```

```
legend("topleft", c("SENSEX","RTSI"), cex=0.65, col=c("blue","orange"),lty=1,title="KEY:  
SENSEX-RTSI pair")
```

```
## Figure 7: RTSI and BOVESPA monthly stock prices
```

```
ts.plot(plot2,main="",xlab="")
```

```
title("",sub="          ",font.sub=1,cex.sub=1.1,ylab="Monthly stock prices", xlab="Time",  
family="A")
```

```

lines(plot2[,1],col="orange")
lines(plot2[,2],col="black")
legend("topleft",          c("RTSI","BOVESPA"),          cex=0.65,
col=c("orange","black"),lty=1,title="KEY: RTSI-BOVESPA pair")

#par(op)

```

### 1) RTSI vs SENSEX

RCode:

```

dz=zoo(data.frame(sensex,rtsi), Date)
regdz = EstimateParameters(dz, method = lm)
str(regdz)
#par(oma=c(2,0,0,0))
## Figure 8: Price spread between the RTSI and SENSEX stocks
plot(regdz$spread,main="",sub=" ", xlab="Time", ylab="Spread",cex.sub=1.1,font.sub=1,
cex.lab=0.9, family="A")

#Is spread stationary or not
IsStationary(regdz$spread, 0.1)

#Analyze the risk adjusted performance of the portfolio
sharpeRatio(regdz$spread)

#estimate parameters for back test and chose 30 month period arbitrarily
params1 = EstimateParametersHistorically(dz, period = 30)

#create and plot trading signals
signal11<- Simple(params1$spread, 0.05)
barplot(signal11,col="blue",space = 0, border = "blue",xaxt="n",yaxt="n",xlab="",ylab="")
par(new=TRUE)
## Figure 10a: Trading signals at 5% spread (S-R)

```

```
plot(params1$spread,main="",sub=" ", xlab="Time",cex.sub=1.1,font.sub=1, cex.lab=0.9,
family="A")
```

```
signal12<- Simple(params1$spread, 0.01)
```

```
barplot(signal12,col="blue",space = 0, border = "blue",xaxt="n",yaxt="n",xlab="",ylab="")
```

```
par(new=TRUE)
```

```
## Figure 10c: Trading signals at 1% spread (S-R)
```

```
plot(params1$spread,main="",sub=" ", xlab="Time",cex.sub=1.1,font.sub=1, cex.lab=0.9,
family="A")
```

```
signal13<- Simple(params1$spread, 0.1)
```

```
barplot(signal13,col="blue",space = 0, border = "blue",xaxt="n",yaxt="n",xlab="",ylab="")
```

```
par(new=TRUE)
```

```
plot(params1$spread,main="",sub="Figure 10e: Trading signals at 10% spread (S-R)",
xlab="Time",cex.sub=1.1,font.sub=1, cex.lab=0.9, family="A")
```

*##Observe performance of pair trading (Sensex-RTSI) over specified time period*

```
## Figure 11a: Performance of the SENSEX-RTSI pairs trading at 5% spread
```

```
return.pairtrading <- Return(dz, lag(signal11), lag(params1$hedge.ratio))
```

```
if(!all(is.na(return.pairtrading))){ plot(100 * cumprod(1 + return.pairtrading),main="",sub=""
",xlab="Time",ylab="Returns",cex.sub=1.1,font.sub=1, cex.lab=0.9, family="A") }
```

```
## Figure 11b: Performance of the SENSEX-RTSI pairs trading at 1% spread
```

```
return.pairtrading <- Return(dz, lag(signal12), lag(params1$hedge.ratio))
```

```
if(!all(is.na(return.pairtrading))){ plot(100 * cumprod(1 + return.pairtrading),main="",sub=""
",xlab="Time",ylab="Returns",cex.sub=1.1,font.sub=1, cex.lab=0.9, family="A") }
```

## 2) RTSI vs BOVESPA

RCode:

```
dz2=zoo(data.frame(rtsi,bovespa), Date)
```

```
regdz2 = EstimateParameters(dz2, method = lm)
```

```
str(regdz2)
#par(oma=c(4,0,0,0))
##Figure 9: Price spread between the RTSI and BOVESPA stocks
plot(regdz2$spread,main="",sub=" ", xlab="Time", ylab="Spread",cex.sub=1.1,font.sub=1,
cex.lab=0.9, family="A")
```

```
#Is spread stationary or not
IsStationary(regdz2$spread, 0.1)
```

```
#Analyze the risk adjusted performance of the portfolio
sharpeRatio(regdz2$spread)
```

```
#estimate parameters for back test and chose 30 month period arbitrarily
params2= EstimateParametersHistorically(dz2, period = 30)
```

```
#create and plot trading signals
```

```
signal21<- Simple(params2$spread, 0.05)
barplot(signal21,col="blue",space = 0, border = "blue",xaxt="n",yaxt="n",xlab="",ylab="")
par(new=TRUE)
## Figure 10b: Trading signals at 5% spread (R-B)
plot(params2$spread,main="",sub=" ", xlab="Time",cex.sub=1.1,font.sub=1, cex.lab=0.9,
family="A")
```

```
signal22<- Simple(params2$spread, 0.01)
barplot(signal22,col="blue",space = 0, border = "blue",xaxt="n",yaxt="n",xlab="",ylab="")
par(new=TRUE)
## Figure 10d: Trading signals at 1% spread (R-B)
plot(params2$spread,main="",sub=" ", xlab="Time",cex.sub=1.1,font.sub=1, cex.lab=0.9,
family="A")
```

```
signal23<- Simple(params2$spread, 0.1)
```

```

barplot(signal23,col="blue",space = 0, border = "blue",xaxt="n",yaxt="n",xlab="",ylab="")
par(new=TRUE)
## Figure 10f: Trading signals at 10% spread (R-B)
plot(params2$spread,main="",sub=" ", xlab="Time",cex.sub=1.1,font.sub=1, cex.lab=0.9,
family="A")

```

*#Observe performance of pair trading (Sensex-RTSI) over specified time period*

Figure 11c: Performance of the RTSI-BOVESPA pairs trading at 1% spread

```

return.pairtrading <- Return(dz2, lag(signal22), lag(params2$hedge.ratio))
if(!all(is.na(return.pairtrading))){ plot(100 * cumprod(1 + return.pairtrading),main="",sub="
",xlab="Time",ylab="Returns",cex.sub=1.1,font.sub=1, cex.lab=0.9, family="A") }

```

## Figure 11b: Performance of the RTSI-BOVESPA pairs trading at 10% spread

```

return.pairtrading <- Return(dz2, lag(signal22), lag(params2$hedge.ratio))
if(!all(is.na(return.pairtrading))){ plot(100 * cumprod(1 + return.pairtrading),main="",sub="
",xlab="Time",ylab="Returns",cex.sub=1.1,font.sub=1, cex.lab=0.9, family="A") }

```

APPENDIX FStatistical tablesTable F1: Dickey-Fuller critical values

AR Model	Critical level			
T=Sample size	0.10	0.05	0.025	0.01
T = 25	-2.62	-3.00	-3.33	-3.75
T = 50	-2.60	-2.93	-3.22	-3.58
T = 100	-2.58	-2.89	-3.17	-3.51
T = 250	-2.57	-2.88	-3.14	-3.46
T = 500	-2.57	-2.87	-3.13	-3.44

Notes: The ADF critical values are taken from <http://rspa.royalsocietypublishing.org>.

Table F2: KP SS test critical values

p-value	Critical Value ( $\eta_{\mu}$ )	Critical Value ( $\eta_{\tau}$ )
0.10	0.347	0.119
0.05	0.463	0.146
0.025	0.574	0.176
0.01	0.739	0.216

Notes: The ADF critical values are taken from <http://rspa.royalsocietypublishing.org>.