A Case Study of Arbitrage Opportunities and Efficiency of the JSE

by

Nthabiseng Rapoeea

(u27368557)

Submitted in partial fulfillment of the requirements for the degree of

Magister Scientiae (Financial Engineering)

in the Department of Mathematics and Applied Mathematics

in the Faculty of Natural and Agricultural Science

University of Pretoria

Pretoria

October 2018

DECLARATION

I, the undersigned declare that the dissertation, which I hereby submit for the degree of *Magister Scientiae in Financial Engineering* at the University of Pretoria, is my own independent work and has not previously been submitted by me or any other person for any degree at this or other tertiary institution.

Nthabiseng Rapoeea

October 2018

DEDICATIONS

Dedicated to my late mother, Mme Mmamosebi Rapoeea, for being supportive and always believing in me. Also not forgetting her unconditional love towards me.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Eben Maré, for his time and guidance.

ABSTRACT

This dissertation examines the market efficiency and arbitrage opportunities between 04 January 2000 and 31 December 2015 on selected JSE-listed stocks and equity indices. To assess market efficiency, four tests were performed namely: structural breaks, stationarity, independence and normality. Lastly, the Pairs trading strategy was implemented to examine arbitrage opportunities profitability, after considering trading costs. The results showed that most stocks and indices are in support of the Adaptive Market Hypothesis (AMH) theory. Arbitrage opportunities do appear and disappear over time, and the Pairs trading strategy performance varies with time but overall profitable.

CONTENTS

List of Symbol

cdf	Cumulative density function
mgf	Moment-generating function
pdf	Probability density function
SE	Standard Error
N	Number of observations in a financial time series
x_t	Financial time series
ε_t	White noise
IID	Independent and Identically Distributed
$\boldsymbol{N}$	Normal Distribution
χ^2	Chi-squared Distribution
$\mathbb{N}^+$	Positive integers
P	Probability function or Probability measure
AR (p)	Autoregressive model of order p
Ω	Denotes a set of variables
$\mathcal{F}$	Denotes a family of subsets of Ω
$\mathbb{F}$	Information filtration
$\mathcal{M}$	Martingale

List of Figures

List of Tables

Glossary

Normal distribution is a distribution with a mean of zero and constant volatility of one. A time series X is normally distributed if X~ $\boldsymbol{N}$ (0;1).

Correlation refers to the relationship between the returns of two (or more) stocks over some time horizon. Correlation does not indicate anything about the long-term behaviour of the stock prices or price time series.

Independent and Identically Distributed ($\boldsymbol{IID}$) means that autocorrelation exists in the process, and identically distributed implies that all random variables have the same distribution parameters.

Efficient Market Hypothesis refers to a market in which prices always "fully reflect" available information (Fama, 1970). The efficiency (or lack of efficiency) of a market can either be a Weak Form, Semi-strong Form or Strong Form.

Excess kurtosis is kurtosis minus 3.

Heterogeneity refers to excessive randomness relative to what is expected by chance only.

Heteroskedasticity refers to an independent variable x with an unequal range of values for its dependent variable y.

Kurtosis is the fourth statistical moment which measures the magnitude of "peakedness" of a probability distribution. A normally distributed time series X~ $\boldsymbol{N}$ (0;1) has a skewness of zero and kurtosis of 3.

Large caps (Top 40) are the forty largest companies that are constituents of the FTSE/JSE Africa All Share Index (J203), ranked by full market capitalization.

Leptokurtic refers to kurtosis > 3; A (fat-tailed distribution) with higher "peakedness" around the mean than the normal distribution. A positive excess kurtosis (kurtosis – 3 > 0) implies a higher probability of outliers (extreme positive or negative returns deviations) than that of a normal distribution.

Mesokurtic refers to kurtosis = 3 (normal distribution).

Mid caps are the next sixty largest companies after the selection of the large caps (Top 40), that are constituents of the FTSE/JSE Africa All Share Index (J203) ranked by full market capitalization.

Platykurtic refers to kurtosis < 3; A (thin-tailed distribution) with a lower peakedness around the mean than the normal distribution.

Small caps are the remaining stocks from the FTSE/JSE Africa All Share Index (J203) after the selection of the large and mid caps.

Skewness is the third statistical moment that measures the degree of asymmetry of a probability distribution. A right-skewed distribution has a positive skewness, while a left-skewed distribution has a negative skewness. A normal distribution has skewness of zero. Skewness > 0 signifies many small negative returns and a few extreme positive returns. There is a low probability of extreme positive returns above the mean. Skewness = 0 means that negative and positive returns have an equal probability. Skewness < 0 signifies many small positive returns and a few extreme negative returns. There is a low probability of extreme negative returns below the mean.

Chapter 1

1.1 Introduction

Stock markets have evolved, and continue to do so with time. More than sixty years ago, stock markets were not exactly what they look like today and how they are currently operated. However, the core objectives for market participants are still more or less the same. With the rise of high frequency trading, big data and digitalization, stock markets are also likely to be different in ten or more years from now. The thought of robotic traders running trading floors by then should not be seen as farfetched. While it is evident that the market is continuously changing, investment strategies have seasons of outperformance and underperformance, and anomalies appear and disappear. Trends, panics, crises, recessions, bubbles, booms and crashes come and go. Markets go through different stages of "boom and bust", just like an economy or a business goes through different life cycles. Before the dawn of market efficiency and more stringent regulations, some stock market activities which are considered illegal today were previously not classified as illegal in most countries. But some activities became prohibited after major stock market shocks or crashes or asset destabilization. For instance, bear raids were common before the Great Crash of 1929, however post that crash, the United States Securities Exchange Commission (SEC) imposed stringent regulations on short selling. During the 2007-2009 global financial crisis, the SEC and regulators in other countries banned short selling, hence modifying the way trading was done during that period. Currency pegging also used to be common until speculators/arbitrageurs who had found loopholes to exploit caused major price destabilizations. Although these speculators destabilize asset prices or markets, they somehow encourage better market environments post the events. In order to remain profitable in the financial markets, participants have to continually adapt to the ever-changing market environment and come up with more innovative ways to search for profits.

Stock markets are pivotal in any economy. They are an intermediary for publicly traded companies to raise capital, and they act as an investment platform for institutional and retail investors or other participants in search of profits and for risk management purposes. Stock markets are often expected to allocate resources fairly and efficiently according to the Efficient Market Hypothesis (EMH).

Founded in1887, the Johannesburg Stock Exchange (JSE) is the largest stock market in Africa, and according to the JSE website, was ranked as the 17th largest stock exchange in the world in terms of market capitalization, as of 30 April 2018. The JSE is the most developed stock market in Africa and a global player; it is both order-driven and quote-driven. It uses an electronic trading system with more than three hundred equity securities, and a variety of derivatives and interest rates instruments traded on the exchange.

1.1.1 Objective of the dissertation

The phenomenon of the financial markets is one of the most exciting and is followed by many in this day and age. The importance of financial markets in the global economy has drawn a lot of interest from different spheres such as academia and financial markets participants or non-financial market participants. These people have spent ample time trying to understand the dynamics and behavioural aspects of financial markets. This quest for profiting and understanding financial markets has given birth to many theories and hypotheses. One of the most researched and talked about hypothesis is the Efficient Market Hypothesis (EMH). Although the EMH is one of the most researched topics over the past forty years, it has been disappointing that a consensus about its effectiveness or its practicality has yet to be reached. The inconclusiveness to this topic therefore suggests the complexity of the financial markets.

In an effort to remediate some of the EMH deficiencies, address some challenges the Efficient Market Hypothesis has been unable to solve, and expand the market hypothesis, Andrew W. Lo proposed what he termed the Adaptive Market Hypothesis (AMH) (Lo, 2004). Lo notes that markets are not always efficient. This means that usually competitive and adaptive investment strategies undergo cycles of profitability and unprofitability in response to changing business conditions. Also, arbitrage opportunities do exist from time to time, but disappear as they are discovered and exploited (Lo, 2004).

In this dissertation the main objective is to marry market efficiency theory's two schools of thought, namely, the Efficient Market Hypothesis (EMH) and the Adaptive Market Hypothesis (AMH). Firstly, the dissertation tackles Weak Form efficiency as a barometer for the Efficient Market Hypothesis. Three important and linked financial time series analysis tests namely, stationarity, independence and normality are performed in order to conclude on this subject. The aim is to demonstrate the Weak Form efficiency or inefficiency of JSE-listed stocks and equity indices forming part of the sample used for this dissertation. This will be achieved by testing whether the daily returns distribution profile of the stocks and indices is normal, and whether there exists a level of predictability and stationarity in the daily returns.

In addressing the Weak Form efficiency question. In Figure 1.1, the following are examined:

i. Stationarity in stocks and indices daily prices by deploying the Phillips-Perron test.
ii. Independence in stocks and indices daily returns by deploying the Ljung-Box test and Autocorrelation test.
iii. Normality in stocks and indices daily returns by deploying the Jarque-Bera test and the Shapiro-Wilk test.

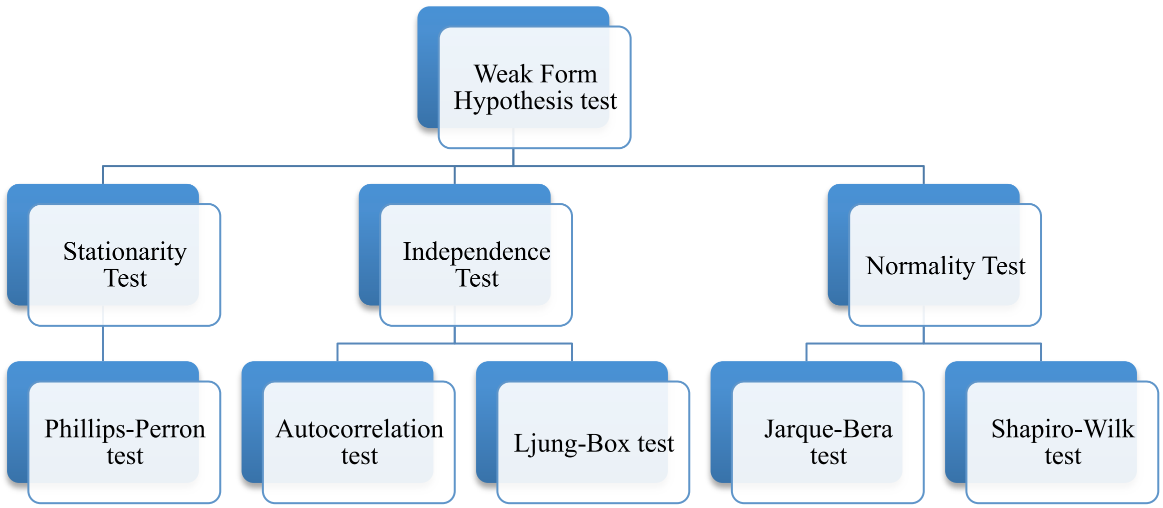


Figure 1.1: The Weak Form efficiency data analysis strategy

Secondly, the Adaptive Market Hypothesis (AMH) concept is tackled by testing daily returns of stocks and indices for potential structural breaks and arbitrage opportunities. Potential drivers of structural breaks are studied and arbitrage opportunities are explored on a selected sub-sample by testing whether arbitrage opportunities do exist and whether they are profitable, after considering trading costs. In addition, arbitrage opportunities profits patterns are studied in order to conclude whether these opportunities go through different cycles of profitability and unprofitability.

In testing for the Adaptive Market Hypothesis, the following concepts are examined, outlined in Figure 1.2:

i. Structural breaks by implementing the Bai-Perron test.
ii. Arbitrage opportunities through the implementation of the Pairs trading strategy.

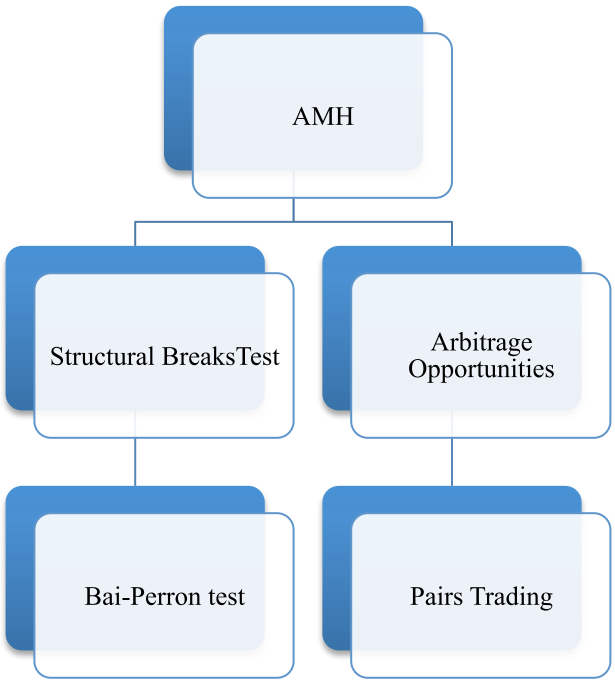


Figure 1.2: The Adaptive Market Hypothesis data analysis strategy

1.1.2 The dissertation structure

The remainder of the dissertation will proceed as follows. The final section of this chapter reviews key considerations in the literature. In Chapter 2, market efficiency is discussed, and particularly the Weak Form hypothesis in more detail. The data, timeframe, how the stocks were selected are described and stock descriptions are also provided.

In Chapter 3, stocks and indices prices are tested for stationarity through the implementation of the Phillips-Perron test. The chapter starts by describing the test and the phenomenon of stationarity and concludes with a summary based on the results.

In Chapter 4, stocks and indices returns are tested for independence. The chapter starts by formulating the concept of independence. Then the tests to be implemented are descibed, namely the Autocorrelation and Ljung-Box test, and why they were chosen. The data is described, the results are analysed and the chapter is summarised.

In Chapter 5, the Jacque-Bera and Shapiro-Wilk tests are implemented to examine the distribution profile of stocks and indices returns. The data and methodology are described, the results are presented and the chapter is summarised.

In Chapter 6, stocks and indices returns are tested for structural breaks. The Bai-Perron test is implemented to search for structural breaks. Structural breaks are tested by allowing and disallowing heterogeneity and autocorrelation in the residuals. And then events and news are searched for on Bloomberg and companies' websites around the break dates. The results are presented and the chapter is summarised.

In Chapter 7, the Pairs trading strategy is deployed to search for arbitrage opportunities. The pairs selection and methodology are described, results are presented and the chapter is summarised.

Chapter 8 concludes the dissertation.

1.2 Literature review

The market efficiency framework has been one of the most talked about topics by academics and professionals in the field of finance and econometrics. Even though it is so well researched, a consensus has yet to be reached as to whether the markets are generally efficient. The history of the efficient markets framework formulation has been a collective effort with its origins dating as far back as the late 1800s. Some key contributors to the market efficiency framework were Jules Regnault (Regnault, 1863), Louis Bachelier (Bachelier, 1900), Paul A. Samuelson (Samuelson, 1965) and Eugene F. Fama (Fama, 1970). Regnault used a binomial model concept by likening the markets to a two-sided dice, with a 50/50 probability of price increase or decrease (Jovanovic and Le Gall, 2001). He added that, in the absence of new information, stock prices are driven by investors' expectations. Bachelier initiated the field of mathematical finance through his *Théorie de la Spéculation* Ph.D dissertation. Paul A. Samuelson extended Bachelier's work of *Théorie de la Spéculation,* formulating the Brownian motion process and later publishing a paper titled "Proof That Properly Anticipated Prices Fluctuate Randomly". Samuelson addressed the martingale property of an unbiased random walk, further formulating an axiom of mathematically expected Futures prices. Samuelson argued that it is impossible to make profit by using past Futures price changes, as the market quotation is already priced in all information known about the Future and discounted future contingencies (Samuelson, 1965). Integrating all the work done by his colleagues, Fama formalized the Efficient Market Hypothesis (EMH) in a paper titled "Efficient Capital Markets: A Review of Theory and Empirical Work" (Fama, 1970). Fama's paper laid a good foundation for financial markets studies, but it became controversial as a consensus has never been reached regarding market efficiency.

Fama (1970) noted that a market in which prices always "fully reflect" available information is called efficient. He formulated three tests for market efficiency: Weak Form, Semi-strong Form and Strong Form.

i. Weak Form tests whether historical prices can be used to predict future prices.
ii. Semi-strong Form tests whether prices adjust to new publicly available information in a timely and instantaneously manner.
iii. Strong Form tests whether there is a group of investors who have monopolistic access to private information relevant to price formation (Fama, 1970).

To a certain extent, the EMH obviates the need for active asset management or arbitrageurs and advocates for passive asset management. Its fundamental principles imply that it is impossible for anyone to consistently outperform an efficient market because prices "fully reflect" all available information. However, as more studies were done, it became clearer that the theory of market efficiency faces some fundamental structural deficiencies. For instance, contradictory results are observed at times if different models are applied to the same dataset. This phenomenon has been observed in different stock markets regardless of their developmental state or size. For instance, Poterba and Summers (1987) carried out a study on U.S listed companies using a dataset between 1926 and 1985. Their study concluded that short term horizon stock returns exhibit positive serial correlations, while longer term horizon stock returns exhibit negative serial correlations, hence they rejected the Random Walk Hypothesis. Their findings were also consistent with Fama and French (1986b) findings. Lo and MacKinlay (1988) carried out a similar study with a dataset between 1962 and 1985 using the volatility-based specification test on weekly returns. They similarly rejected the Random Walk Hypothesis, but noted that the stationary mean-reverting model of Poterba and Summers (1987) could not account for the departures of weekly returns from Random Walk. They also added that a rejection of the Random Walk model does not necessarily imply inefficiency in asset price formation. Likewise, there are studies in South Africa with inconclusive or opposing findings. For example, Magnusson and Wydick (2002) concluded that the JSE is Weak Form efficient, while Appiah-Kusi and Menyah (2003) concluded that the JSE is Weak Form inefficient. Phiri (2015) further tested the Weak Form efficiency in five major JSE equity indices using linear and nonlinear unit root testing procedures. The Weak Form efficiency hypothesis was not rejected when the linear unit root test procedure was applied. On the contrary, the Weak Form efficiency hypothesis was rejected when the nonlinear unit root test procedure was applied (Phiri, 2015).

Efforts have been made to remediate some of the EMH deficiencies, address some challenges which the Efficient Market Hypothesis has been unable to solve and expand the market hypothesis. Lo (2012) proposed what he termed the Adaptive Market Hypothesis (AMH), which states that markets are not as rational as EMH implies but are rather driven by fear and greed. They are not always efficient but are usually competitive and adaptive, varying in their degree of efficiency as the environment and investors change over time. Under AMH, investment strategies undergo cycles of profitability and unprofitability in response to changing business conditions. Arbitrage opportunities do exist from time to time, but disappear as they are discovered and exploited, and innovation is key to alpha searching (Lo, 2004). Lo does not refute the Efficient Market Hypothesis but tries to marry the two schools of thought, emphasizing that investors learn to adapt to changing economic environment, and markets are consistently changing. The concept of AMH is still relatively new and the literature is limited, however, a few studies have been done in some developed and emerging markets, including South Africa. For example, Borges (2011) tested the Lisbon PSI-20 Index for Weak Form using various Random Walk tests, and the results showed that the degree of Lisbon PSI-20 Index efficiency improved with time. Another study by Jefferis and Smith (2005), analyzed the efficiency of seven African stock markets (South Africa, Egypt, Morocco, Nigeria, Zimbabwe, Mauritius and Kenya) using a GARCH approach with time-varying parameters. Results revealed that three of the seven stock markets became Weak Form efficient towards the end of the assessment period between 1990 and end of June 2001, while South Africa remained Weak Form efficient throughout the assessment period. Likewise, Smith and Dyakova (2014) studied the predictability of eight African stock markets (Egypt, South Africa, Kenya, Nigeria, Morocco, Zambia, Tunisia and Mauritius) using three-sample variance ratio tests on a rolling window over the period between February 1998 to December 2011. Results showed that inefficient periods are often followed by efficient periods. Moreover, they found that the degree of predictability is driven by size, liquidity and market quality. Hudson and Manahov (2014) further investigated stock market dynamics and market efficiency formation by implementing econometric techniques on artificially developed markets. They discovered that heterogeneity leads to higher market efficiency, and that stock market dynamics and nonlinearity are explained by an evolutionary process which is what AMH advocates. Other studies support the concept of AMH (Urqhart and McGroarty, 2014; Rodriguez et al., 2014; Noda, 2016, Ghazani and Araghi, 2014).

Numerous past studies have found it rather challenging to reach a conclusive outcome with regards to the theory of market efficiency. This dissertation however aims to take research a step further and test whether arbitrage opportunities do arise when inefficiencies occur. The dissertation implements a relative value arbitrage strategy namely, Pairs trading, a branch of statistical arbitrage strategy. Relative value strategy aims to exploit any mispricing between two securities with similar economic attributes, while maintaining a diversified risk profile. Since its discovery in the 1980s by the Wall Street quants at Morgan Stanley, Pairs trading strategy has been popular with a lot of hedge fund managers. Numerous studies have been done in both emerging and developed markets. The concept of Pairs trading strategy involves exploitation of any mispricing as a result of the Law of One Price violation. This strategy works well on mean-reverting stocks, but can yield losses when the convergence trade diverges. Implementing Pairs trading strategy on the JSE, Govender (2011), Mashele et al. (2013), Masindi (2014), Augustine (2014) showed significant profitability, after considering trading costs. Mashele et al. (2013) found that trading costs have to be at most 20bps for the Pairs trading strategy to be profitable. Trading costs are particularly important to take into account to reach a conclusive outcome. This is because pairs trading involves trading two transactions twice.

In other markets, Do and Faff (2012) conducted a pairs trading profitability study on U.S equities between 1963 and 2009. They found that net of trading costs, Pairs trading strategy remained profitable yielding a risk-adjusted return of 30bps per month on a well refined inter-industry group. Perlin (2009) found that pairs trading and market neutral strategies were profitable on the Brazilian stock market. Gundersen (2014) also examined twenty-five largest shares listed on the Oslo Stock Exchange for profitability of the pairs trading strategy using three months of high frequency intraday data and found that the strategy yielded positive risk-adjusted returns, net of fees.

Chapter 2: Market Efficiency definition

2.1 Market efficiency

The origins of market efficiency date as far back as the late 1800s, however in a paper titled Efficient Capital Markets: A Review of Theory and Empirical Work, Eugene F. Fama formulated what popularly became known as the Efficient Market Hypothesis (EMH) (Fama, 1970). The EMH concept of market efficiency is an “all-or-nothing” approach. This means that a capital market in which prices always “fully reflect” all available information is called “efficient” (Fama, 1970). The market efficiency test has three forms namely:

i. Weak Form: can historical asset prices be used as the best predictor of future asset prices?
ii. Semi-strong Form: do asset prices efficiently adjust to new publicly available information (such as earnings announcements, dividend announcements, mergers and acquisitions, and more)?
iii. Strong Form: do some individuals have monopolistic access to any relevant, private information and use it for their benefit?

The concept of market efficiency is consistent with the Fair Game Model (Samuelson, 1965). This model implies that, given prices which “fully reflect” all available information, it is impossible to earn greater returns without taking additional risk. Furthermore, historical prices are not useful in predicting future prices. Market efficiency supports a buy-and-hold strategy or index tracking strategy and advocates for the impossibilities of profitable arbitrage strategies. An efficient market is completely characterized by the absence of arbitrage opportunities and dominated securities with respect to information set (see Jarrow and Larsson, 2012).

Definition 2.2.1. (Jarrow and Larsson, 2012) A market $(\mathbb{F}, S)$ is called efficient on [0, T] with respect to $\mathbb{F}$ if there exists a consumption good price index ψ and economy $(\text{P}, \mathbb{F}, \{\varepsilon_i\}_{i=1}^{n}, \{U_i\}_{i=1}^{n})$ which (ψ, S) is an equilibrium price process S on [0, T].

If this holds for every T$< \infty$, the market is called efficient w.r.t $\mathbb{F}$.
where

$\mathbb{F}$ is the information filtration.
S is a stochastic process.

Lemma 2.2.1. (Jarrow and Larsson, 2012) Suppose an equilibrium is given. Then holding the market portfolio is a maximal strategy, i.e. $H = (H^1, \dots, H^d)$ given by
$H^i(t) \equiv \alpha^i, \quad i = 1, \dots, d$ is maximal.

Lemma 2.2.2. (Jarrow and Larsson, 2012) Suppose an equilibrium is given, then for each fixed $i \in \{0,1,\dots,d\}$, the strategy $H = (H^0, \dots, H^d)$ given by

$$\begin{cases} H^i \equiv 1, & j \neq i \\ H^j \equiv 0, & j \neq i \end{cases}$$

is maximal, i.e. no arbitrage holds.

Theorem 2.2.1. (Jarrow and Larsson, 2012) Let $(\mathbb{F}, S)$ be a market. The following statements are equivalent:

a) $(\mathbb{F}, S)$ is efficient on [0,T],
b) $(\mathbb{F}, S)$ satisfies no arbitrage on [0,T],
c) There exists a probability Q, equivalent to P, such that S is an $(\mathbb{F}, Q)$ martingale on [0, T]. That is $\mathcal{M}(\mathbb{F}, S, T) \neq \emptyset$.

Proof

a) If $(\mathbb{F}, S)$ is efficient on [0,T], there is a consumption good price index ψ in an economy $(\mathrm{P}, \mathbb{F}, \{\varepsilon_i\}_{i=1}^n, \{U_i\}_{i=1}^n)$ such that (ψ, S) is an equilibrium price process.
b) If $(\mathbb{F}, S)$ satisfies no arbitrage then all the strategies $H^i$are maximal. Finite sums of maximal strategies are again maximal.
c) Assume there exists an equivalent martingale measure Q.

Let all investors in the economy have logarithmic utilities

$$U_i(x) = \begin{cases} \ln x, & x > 0 \\ -\infty, & x \leq 0 \end{cases}$$

Suppose $\mu(\{T\}) = 1$. Let $\psi(t) \equiv 1$

Assume ε_i are identically zero, then the investors only receive utility from the liquidation of asset dividends.

Assume investor beliefs are given by

$$\frac{d\mathrm{P}^*}{dQ} = Z(T),$$

where

$Z(t) = \frac{\alpha^1 S^1(t) + \cdots \alpha^d S^d(t)}{\alpha^1 + \cdots + \alpha^d}$ which is a Q −martingale.

The i^{th} investor's optimization problem is then

$$Sup\{E_{\mathrm{P}^*}[U_i(X(T))]: X(T) = x_i + \int_0^T H(S)dS(s), H \text{ is admissible}\}.$$

$U_i(x) = -\infty$ for $x \leq 0$, we restrict attention to strategies for which $X(T) > 0$.

Then due to the supermartingale property of $X = x_i + \int H(t)ds(t)$ under Q, $X(t) \geq E_Q(X(T)|\mathcal{F}_t) \geq 0$ for all $t \leq T$.

Hence we only need to consider x_i −admissible strategies.

If Theorem 2.2.1 does not hold it means

a) The market is not efficient,
b) Arbitrage opportunities do exist, and
c) The market does not follow a martingale, buy-and-hold strategy would not be a maximal strategy and Lemma 2.2.1 would be false.

Let daily stock returns time series be represented as $x_t \dots x_n$. This time series is a martingale if $E[x_{t+1}|\Phi_t\ ] = x_t$
where

Φ_t is the information set at t.

This means that if the information set remains constant at time x_{t+1}, the stock's expected return at time x_{t+1} should equal x_t.

However, should $E[x_{t+1}|\Phi_t\ ] \geq x_t$, then the time series is a submartingale or supermartingale if $E[x_{t+1}|\Phi_t\ ] \leq x_t$. In the case of a submartingale, the use of arbitrage strategies such as Pairs trading would not yield greater returns than that of a buy-and-hold strategy, Fama (1970). Contrarily, in the case of supermartingale, the use of Pairs trading strategy or other technical analysis strategies may yield greater returns than that of a buy-and-hold strategy. A market where the use of technical analysis yields profitable returns is considered Weak Form inefficient.

2.2 Weak Form

The Weak Form tests whether historical asset prices can be used to predict future asset prices or returns. There are many methods available to test for the predictability of asset prices such as traditional statistical tests, technical analysis, and more. The Random Walk Hypothesis is highly pivotal when testing the Weak Form hypothesis. In order to conclude whether a stock is Weak Form efficient or not, three tests are run namely, stationarity, independence, and normality tests (parametric and non-parametric tests were implemented). In addition, stocks and indices were tested for the potential existence of structural changes.

Let's assume a random stock represented by x. It's daily returns time series are represented as $x_t \dots x_n$.

Daily returns are calculated as

$$x_t = \frac{x_t}{x_{t-1}} - 1. \tag{1}$$

A Random Walk process without a drift of risky asset's returns is represent by (Tsay, 2002)

$$x_t = x_{t-1} + \varepsilon_t.$$

$\varepsilon_t \sim IID(0, \sigma^2)$ is white noise.

$\varepsilon_t = 0$ implies fair game.

where

x_t is the time series of asset's daily returns.

ε_t is the white noise process.

Fair game means that, at any given stage, the probability of one player winning or losing equals that of the player's opponent. Alternatively, the probability of an asset price moving up equals that of a downward move. Fair game is critical in the measure of market efficiency, especially Weak Form. If a market or an asset is fair game, then it is impossible for one to profit from it using historical prices to predict future prices.

EMH implies that all publicly available information is "fully reflected" in the asset. Today's asset price is the best predictor of tomorrow's asset price. However, new information arriving tomorrow is independent of today's asset price. Therefore, an asset price with this property is called a Random Walk.

Let x represent the asset price

$$x_t = x_{t-1} + \varepsilon_t.$$

$$\varepsilon_t \sim IID \quad \mathbb{N}(0, \sigma^2).$$

Tsay (2002) represents a Random Walk process with a drift of a risky asset's returns as:

$$x_t = x_{t-1} + \mu + \varepsilon_t.$$

$$\varepsilon_t \sim IID \quad \mathbb{N}(0, \sigma^2).$$

where

x_t is the time series of asset's daily returns.

ε_t is the white noise process.
μ is the drift.

The Random Walk Model is given by Fama (1970) as:

$$f(x_{i,t+1} \mid \Phi_t) = f(x_{i,t+1}). \qquad (2)$$

This implies that the conditional and marginal probability of independent random variables is identical. The density function f must be the same for all t.

A Random Walk hypothesis process with IID is therefore:

$$x_t = x_0 + \sum \varepsilon_t, \quad t = 0,1, \ldots$$

$$\mathrm{E}\,[\varepsilon_t] = 0, \text{ for all t.}$$

Since errors have no speculation, this implies fair game and that the increments are uncorrelated.

Sequence $\{x_{t+1}\}$ is a fair game w.r.t information sequence $\{\Phi_t\}$ which implies that the sequence is a martingale. If the stock price/return does not follow a martingale then the use of technical analysis or arbitrage strategies may yield profitable returns than the use of a buy-and-hold strategy. The martingale principle was introduced by Bachelier in 1900 through his Theory of Speculation Ph.D dissertation in which he noted that the expected profits of a speculator are zero.

The **Fair Game Model** is given by Fama (1970) as:

$$\phi[x_{t+1} \mid \phi_t] = \phi_t x_{t+1} = 0. \qquad (3)$$

For $\phi[x_{t+1} \mid \phi_t] = x_t$, is a martingale w.r.t ϕ_t .

2.3 Data

This dissertation analyzed ninety stocks and four major equity indices. The data sample was collected between 4 January 2000 and 31 December 2015. In order to have an inclusive and comprehensive study, four major equity indices were selected and all stocks that were listed on the Johannesburg Stock Exchange and met the following requirements:

- Listed and traded on the Johannesburg Stock Exchange (JSE) as of 04 January 2000.
- No trading disruptions during the assessment period, such as stock suspension except for market wide suspensions.
- Companies did not merge or were not acquired by another during the assessment period.

Thirty-one stocks were constituents of the FTSE/JSE Africa Top 40 Index. Forty stocks were constituents of the FTSE/JSE Africa Mid Cap Index. Nineteen stocks were constituents of the FTSE/JSE Small Cap Index. In addition, the following corresponding equity indices were included in the analysis: FTSE/JSE Africa All Share Index, FTSE/JSE Africa Top 40 Index, FTSE/JSE Africa Mid Cap Index and the FTSE/JSE Africa Small Cap Index. Stocks and indices descriptions are displayed in Table (2.1) – Table (2.4) below.

The test sample was made up of the daily price time series of ninety stocks and four equity indices. The stocks data sample was collected between 4 January 2000 and 31 December 2015, while the indices data sample started later than 4 January 2000 as the indices were created and listed on the stock exchange later on.

Stocks/Indices	Start Date	End Date	Number of Observations
Large cap stocks	4-Jan-2000	31-Dec-2015	3999
Mid cap stocks	4-Jan-2000	31-Dec-2015	3999
Small cap stocks	4-Jan-2000	31-Dec-2015	3999
FTSE/JSE Africa All Share Index	3-Jun-2002	31-Dec-2015	3397
FTSE/JSE Africa Top 40 Index	21-Jun-2002	31-Dec-2015	3384
FTSE/JSE Africa Mid Cap Index	12-Jun-2007	31-Dec-2015	2141
FTSE/JSE Africa Small Cap Index	31-Dec-2001	31-Dec-2015	3502

Table 2.1: Daily price data sample

JSE Code	Share Name	Sub-sector class	Sector class	Business Operations
AGL	Anglo American Plc	Basic Resources	Basic Materials	Bulk commodities mining including precious metals
AMS	Anglo American Platinum Ltd	Basic Resources	Basic Materials	Mining and production of platinum and platinum group metals
ANG	AngloGold Ashanti Ltd	Basic Resources	Basic Materials	Mining and exploration of gold
APN	Aspen Pharmacare Holdings Ltd	Health Care	Health Care	Supplier of pharmaceuticals
BGA	Barclays Africa Group Ltd	Banks	Financials	Banking and other financial services
BIL	BHP Billiton Plc	Basic Resources	Basic Materials	Mineral exploration and production
BVT	Bidvest Ltd	Industrial Goods & Services	Industrials	Office products, packaging, food and allied products manufacuring and distribution
CFR	Compagnie Financiere Richemont	Personal & Household Goods	Consumer Goods	Luxury goods
DSY	Discovery Ltd	Insurance	Financials	Financial services and health insurance
FSR	FirstRand Ltd	Banks	Financials	Banking and other financial services
GRT	Growthpoint Properties Ltd	Real Estate	Financials	Real estate
IMP	Impala Platinum Holdings Ltd	Basic Resources	Basic Materials	Production and refinery of platinum and platinum group metals
INL	Investec Ltd	Financial Services	Financials	Banking and other financial services
IPL	Imperial Holdings Ltd	Industrial Goods & Services	Industrials	Logistics
ITU	Intu Properties Plc	Real Estate	Financials	Property development
MDC	Mediclinic Internatational Ltd	Health Care	Health Care	Health care facilities
MRP	Mr Price Group Ltd	Retail	Consumer Services	Clothing and general merchandise
MTN	MTN Group Ltd	Telecommunications	Telecommunications	Communications services
NED	Nedbank Group Ltd	Banks	Financials	Banking and other financial services
NPN	Naspers Ltd -N-	Media	Consumer Services	Multinational internet and media group
NTC	Netcare Ltd	Health Care	Health Care	Health care facilities
OML	Old Mutual Plc	Insurance	Financials	Life insurance and other financial services
RMH	RMB Holdings Ltd	Banks	Financials	Investments and other financial services
SAB	SABMiller plc	Food & Beverage	Consumer Goods	Brewery and soft drinks
SBK	Standard Bank Group Ltd	Banks	Financials	Banking and other financial services
SNH	Steinhoff International Holdings N.V.	Personal & Household Goods	Consumer Goods	Household goods and clothing manufacturing and sourcing
SHP	Shoprite Holdings Ltd	Retail	Consumer Services	Food and household goods retailer
SLM	Sanlam Ltd	Insurance	Financials	Life insurance and other financial services
SOL	Sasol Ltd	Chemicals	Basic Materials	Integrated energy and chemicals
TBS	Tiger Brands Ltd	Food & Beverage	Consumer Goods	Branded foods manufacturer
WHL	Woolworths Holdings Ltd	Retail	Consumer Services	Food and general merchandise retailer

Table 2.2: Top 40 stocks and their descriptions

Source: Bloomberg and company website

JSE Code	Share Name	Sub-sector class	Sector class	Business Operations
ACL	ArcelorMittal SA Ltd	Basic Resources	Basic Materials	Steel production
AFE	AECI Ltd	Chemicals	Basic Materials	Commercial explosives and specialty chemical manufacturing
ARI	African Rainbow Minerals Ltd	Basic Resources	Basic Materials	Diversified mining and minerals
ASR	Assore Ltd	Basic Resources	Basic Materials	Base minerals and metals mining
AVI	AVI Ltd	Food & Beverage	Consumer Goods	Packaged foods and personal care brands
BAT	Brait SE	Financial Services	Financials	Investments in privately owned businesses
BAW	Barloworld Ltd	Industrial Goods & Services	Industrials	Integrated logistics and other mobility services
CLS	Clicks Group Ltd	Retail	Consumer Services	Food, drugs and household retailer
DST	Distell Group Ltd	Food & Beverage	Consumer Goods	Distillation and distribution of spirits and wines
DTC	Datatec Ltd	Technology	Technology	ICT supply and other technology solutions
EOH	EOH Holdings Ltd	Technology	Technology	Information Technology
FBR	Famous Brands Ltd	Travel & Leisure	Consumer Services	Packaged foods and beverages
GFI	Gold Fields Ltd	Basic Resources	Basic Materials	Gold production
GND	Grindrod Ltd	Industrial Goods & Services	Industrials	Freight services, shipping and financial services
HAR	Harmony Gold Mining Company Ltd	Basic Resources	Basic Materials	Gold mining and exploration
HYP	Hyprop Investments Ltd	Real Estate	Financials	Real estate investments
ILV	Illovo Sugar Ltd	Food & Beverage	Consumer Goods	Sugar production
ITE	Italtile Ltd	Retail	Consumer Services	Home products retailer
KAP	KAP Industrial Holdings Ltd	Industrial Goods & Services	Industrials	Logistics
LBH	Liberty Holdings Ltd	Insurance	Financials	Insurance and other financial services
LON	Lonmin Plc	Basic Resources	Basic Materials	Production of platinum and platinum group metals
NHM	Northam Platinum Ltd	Basic Resources	Basic Materials	Platinum extraction
NPK	Nampak Ltd	Industrial Goods & Services	Industrials	Packaging
OCE	Oceana Group Ltd	Food & Beverage	Consumer Goods	Fishing, cold storage and shipping
OMN	Omnia Holdings Ltd	Chemicals	Basic Materials	Specialised agricultural, mining and chemical solutions
PIK	Pick n Pay Stores Ltd	Retail	Consumer Services	Food and general merchandise retailer
PPC	PPC Ltd	Construction & Materials	Industrials	Cement and lime manufacturing and distribution
PSG	PSG Group Ltd	Financial Services	Financials	Diversified financial services
RCL	RCL Foods Ltd	Food & Beverage	Consumer Goods	Food production
RLO	Reunert Ltd	Industrial Goods & Services	Industrials	ICT, electrical engineering and applied electronics
SAC	SA Corporate Real Estate Ltd	Real Estate	Financials	Diversified real estate investment trust
SAP	Sappi Ltd	Basic Resources	Basic Materials	Coated wood-free paper production
SNT	Santam Ltd	Insurance	Financials	Short-term insurance
SPG	Super Group Ltd	Industrial Goods & Services	Industrials	Logistics and supply chain management
SUI	Sun International Ltd	Travel & Leisure	Consumer Services	Entertainment and leisure
TFG	The Foschini Group Ltd	Retail	Consumer Services	Clothing retailer
TON	Tongaat Hulett Ltd	Food & Beverage	Consumer Goods	Agriculture and agri-processing
TRE	Trencor Ltd	Industrial Goods & Services	Industrials	Marine cargo containers leasing, managing and trading
TRU	Truworths International Ltd	Retail	Consumer Services	Clothing retailer
TSH	Tsogo Sun Holdings Ltd	Travel & Leisure	Consumer Services	Entertainment and leisure

Table 2.3: Mid caps stocks and their descriptions

Source: Bloomberg and company website

JSE Code	Share Name	Sub-sector class	Sector class	Business Operations
ACT	AfroCentric Investment Corporation Ltd	Health Care	Health Care	Health administration and health risk management solutions
ADH	ADvTECH Ltd	Retail	Consumer Services	Formal education services
ADR	Adcorp Holdings Ltd	Industrial Goods & Services	Industrials	Human capital management and business process outsorcing services
AEG	Aveng Group Ltd	Construction & Materials	Industrials	Construction, mining, engineering, manufacturing and processing
AEL	Allied Electronics Corporation (A Share) Ltd	Industrial Goods & Services	Industrials	Information Technology, telecommunication and multi-media provider
AEN	Allied Electronics Corporation (N Share) Ltd	Industrial Goods & Services	Industrials	Information Technology, telecommunication and multi-media provider
AFX	African Oxygen Ltd	Chemicals	Basic Materials	Welding and gases manufacturer
BRN	Brimstone Investment Corporation (N Share) Ltd	Investment Instruments	Financials	Financial services
CAT	Caxton and CTP Publishers and Printers Ltd	Media	Consumer Services	Publishing, digital, packaging and printing services provider
CLH	City Lodge Hotels Ltd	Travel & Leisure	Consumer Services	Lodging
CSB	Cashbuild Ltd	Retail	Consumer Services	Bulding materials and associated products retailer
GRF	Group Five Ltd	Construction & Materials	Industrials	Construction, engineering, manufacturing and properties industries
HDC	Hudaco Industries Ltd	Industrial Goods & Services	Industrials	Industrial and automobile products retailer
MRF	Merafe Resources Ltd	Basic Resources	Basic Materials	Mining and exploration of ferrochrome
MTA	Metair Investments Ltd	Automobiles & Parts	Consumer Goods	Automotive industry products manufacturer and distributor
MUR	Murray & Roberts Holdings	Construction & Materials	Industrials	Engineering and construction services provider
OCT	Octodec Invest Ltd	Real Estate	Financials	Real estate owners and developers
PGR	Peregrine Holdings Ltd	Financial Services	Financials	Investment management
WBO	Wilson Bayly Holmes-Ovcon Ltd	Construction & Materials	Industrials	Building construction, civil engineering, roads and earthworks

Table 2.4: Small caps stocks and their descriptions

Source: Bloomberg and company website

JSE Code	Bloomberg Code	Index Name	Index Description
J203	JALSH Index	FTSE/JSE Africa All Share Index	Represents 99% of the full market capitalisation of all eligible equities listed on the Main Board of the JSE.
J200	Top 40 Index	FTSE/JSE Africa Top 40 Index	Represents the forty largest companies which are constituents of the All Share Index (J203), ranked by full market capitalisation.
J201	MIDCAP Index	FTSE/JSE Africa Mid Cap Index	Represents the next sixty largest companies which are constituents of the All Share Index (J203), ranked by full market capitalisation, after the selection of the Top 40 constituents (J200).
J202	JSMLC Index	FTSE/JSE Africa Small Cap Index	Represents the remaining stocks from the All Share Index (J203) after the selection of the Top 40 Index (J200) and MIDCAP Index (J201) constituents.

Table 2.5: Indices and their descriptions

Source: www.jse.co.za

Chapter 3: Test for Stationarity

3.1 Introduction

Stationarity is one of the most important factors in financial time series analysis. A financial time series that follows a Random Walk process is known to be non-stationary or at least weakly stationary (Tsay, 2002). The presence of stationarity in a financial time series indicates mean-reversion which leads to violation of the Random Walk Hypothesis. Random Walk has an infinite variance which simply implies that the best prediction of tomorrow's asset price is today's asset price. On the other hand a stationary process is predictable, based on its mean-reverting behaviour.

Let $\{x_{t1},\ldots,x_{tn}\}$ and $\{x_{tn+1},\ldots,x_{tn+N}\}$ be joint distributions of a financial time series.

If $\{x_{t1},\ldots,x_{tn}\}$ is identical to $\{x_{tn+1},\ldots,x_{tn+N}\}$ such that the joint distribution of

$\{x_{t1},\ldots,x_{tn}\}$ is time-invariant to $\{x_{tn+1},\ldots,x_{tn+N}\}$ then a financial time series $\{x_t\}$ is said to be strictly stationary

$$n, N \in \mathbb{N}^{+}.$$

If the mean and covariance or variance of a financial time series $\{x_t\}$ are identical, implying time-invariance, then such financial time series is said to be weakly stationary, e.g. $E(x_t)$ and $Var(x_t)$ are both time-invariant or constant (Tsay, 2002).

A financial time series $\{x_t\}$ is said to be non-stationary if it has a unit-root, or its mean and variance are time varying. A unit-root non-stationary time series follows Random Walk.

Let's assume a Random Walk process (Tsay, 2002):

$$x_t = \alpha + \beta_1 t + \beta_2 x_{t-1} + \varepsilon_t. \quad (8)$$

where

$\varepsilon_t \sim IID(0, \sigma^2)$ is white noise.

If $\alpha = 0, \beta_1 = 0$ and $\beta_2 = 1$ then $x_t = x_{t-1} + \varepsilon_t$ is a Random Walk process with no drift. Therefore x_t is non-stationary.

$$\mathrm{E}(x_t) = E(x_{t-1} + \varepsilon_t),$$

$$= E\,(x_{t-1}) + E(\varepsilon_t),$$
$$= E\,(x_{t-1}).$$

Since $E(\varepsilon_t) = 0$.

$$\text{Var}(x_t) = Var\,(x_{t-1} + \varepsilon_t),$$
$$= Var\,(x_{t-1}) + Var(\varepsilon_t),$$
$$= Var\,(x_{t-1}) + \sigma^2.$$

If $\alpha \neq 0, \beta_1 = 0$ and $\beta_2 = 1$ then $x_t = \alpha + x_{t-1} + \varepsilon_t$ is a Random Walk with a drift. Therefore x_t is non-stationary.

$$\text{E}(x_t) = E(\alpha + x_{t-1} + \varepsilon_t),$$
$$= E\,(\alpha) + E\,(x_{t-1}) + E(\varepsilon_t),$$
$$= \alpha + E\,(x_{t-1}).$$

Since $E(\varepsilon_t) = 0$.

$$\text{Var}(x_t) = Var\,(\,\alpha + x_{t-1} + \varepsilon_t),$$
$$= Var(\alpha\,) + Var\,(x_{t-1}) + Var(\varepsilon_t),$$
$$= \alpha + Var\,(x_{t-1}) + \sigma^2.$$

If $\alpha \neq 0, \beta_1 \neq 0$ and $\beta_2 = 0$ then $x_t = \alpha + \beta_1 t + \varepsilon_t$ is a Random Walk with a deterministic trend.

Therefore x_t is non-stationary.

$$\text{E}(x_t) = E\,(\alpha + \beta_1 t + \varepsilon_t),$$
$$= E\,(\alpha) + E\,(\beta_1 t) + E(\varepsilon_t),$$
$$= \alpha + \beta_1 t.$$

Since $E(\varepsilon_t) = 0$.

$$\text{Var}(x_t) = Var\,(\alpha + \beta_1 t + \varepsilon_t),$$
$$= Var(\alpha\,) + Var\,(\beta_1 t) + Var(\varepsilon_t),$$
$$= \alpha + \beta_1 t + \sigma^2.$$

If $\alpha \neq 0, \beta_1 \neq 0$ and $\beta_2 = 1$ then $x_t = \alpha + \beta_1 t + x_{t-1} + \varepsilon_t$ is a Random Walk with a drift and deterministic trend.
Therefore x_t is non-stationary.

$$\begin{aligned} \mathrm{E}(x_t) &= E\,(\alpha + \beta_1 t + x_{t-1} + \varepsilon_t\,), \\ &= E\,(\alpha) + E(\beta_1 t) + E\,(x_{t-1}) + E(\varepsilon_t), \\ &= \alpha + \beta_1 t + E\,(x_{t-1}). \end{aligned}$$

Since $E(\varepsilon_t) = 0$.

$$\begin{aligned} \mathrm{Var}(x_t) &= Var\,(\alpha + \beta_1 t + x_{t-1} + \varepsilon_t), \\ &= Var(\alpha\,) + Var(\beta_1 t) + Var\,(x_{t-1}) + Var(\varepsilon_t), \\ &= \alpha + \beta_1 t + Var\,(x_{t-1}) + \sigma^2. \end{aligned}$$

A discrete time stochastic process $x_t]_{t=1}^{T}$ is stationary if both $E(x_t)$ and $Var(x_t)$ are finite constants.

When $Cov(x_t, x_s)$ depends only on t−s then the process is weakly stationary or covariance stationary (Tsay, 2002). The joint distribution of (x_t, x_s) depends only on t−s.

Assume an Autoregressive model of order 1 [AR(1)]

$$x_t = \alpha + \rho x_{t-1} + \varepsilon_t.$$

$\varepsilon_t \sim IID \;\; \mathbb{N}\,(0, \sigma^2)$ and $\rho < 1$. Then ρ is called the first order autocorrelation coefficient. If $\rho = 0$ then the process is mostly stationary, hence mean-reverting.

For an Autoregressive model of order p, AR (p)

$$x_t = \alpha + \rho_1 x_{t-1} + \rho_2 x_{t-2} + \cdots + \rho_p x_{t-p} + \varepsilon_t.$$

For a stationary process $E(x_t) = E(x_{t-1})$ and $Var(x_t) = Var(x_{t-1})$.

$$E(x_t) = \frac{\alpha}{1-\rho}.$$

$$Var(x_{t-1}) = \frac{\sigma^2}{1-\rho^2}.$$

If $|\rho| < 1$ then $E(x_t)$ and $Var(x_t)$ are finite.

The autocorrelation coefficient ρ is given by:

$$\rho = \frac{cov(x_t, x_{t-1})}{V(x_t)}.$$

3.2 Methodology and data

To test whether daily stock daily prices are non-stationary and hence follow Random Walk, this dissertation deployed the Phillips-Perron (PP) test utilizing the "tseries" RStudio package. The Phillips-Perron test is a nonparametric method for unit-root testing, introduced by Phillips and Perron in 1988. It is an extension of the Dickey-Fuller test with a robust feature to detect any serial correlation and heteroskedasticity in the white noise (error term). The Phillips-Perron test is an improvement of existing unit-root testing methods. Its modified test statistics lead to consistent estimates of variance parameters (σ^2 and λ^2), given below under the hypothesis test. The Phillips-Perron test is also more relevant if the residuals of a unit process are heterogeneous or weakly dependent (Enders, 2004).

The main advantages the Phillips-Perron test has over the Dickey-Fuller (DF) or Augmented Dickey-Fuller (ADF) tests are its robustness to detect serial correlation and heteroskedacity in the error term. The Phillips-Perron test also does not require the user to specify the lag length of a regression. However, its main disadvantages are its low power in detecting stationarity when a root is close to the non-stationarity boundary and its failure to detect stationarity in the presence of structural change (DeJong et al., 1992). The Dickey-Fuller and Phillips-Perron tests have often been criticized for their low power to reject the null hypothesis in the presence of structural instability. The application of the Phillips-Perron test is limited in the South African literature, instead the DF and ADF tests have been more common. However, a few studies have been carried out in other markets (Higgs & Worthington, 2004). Refer to Appendix III for additional information regarding the Phillips-Perron test.

The test sample was made up of daily price time series as described in Table 2.1 above.

Assume the regression model

$$x_t = \rho_0 x_{t-1} + \varepsilon_t; \qquad \varepsilon_t \sim IID(0, \sigma^2), \qquad (9)$$

For $\rho_0 = 1$

$$N(\rho_N - \rho_0) = \frac{N^{-1}\sum_{t=1}^{N} x_{t-1}\varepsilon_t}{\sum_{t=1}^{N} x_{t-1}\varepsilon_t} \Rightarrow \frac{\int_0^1 B(r)dB(r)}{\int_0^1 B(r)^2\, dr}.$$

where

$B(r)$ is a standard Brownian motion.

Hypothesis test

$H_0 : \delta = 0$ (Existence of unit-root, $\beta_2 = 1$).
$H_1 : \delta \neq 0$ (Stationary, $\beta_2 \neq 1$).

The Z -test statistic is given by (Phillips and Perron, 1988)

$$z_t = \sqrt{\frac{\sigma^2}{\lambda^2}}t_\delta - \frac{1}{2}\left(\frac{\lambda^2 - \sigma^2}{\lambda^2}\right)\left(\frac{N\big(SE(\delta)\big)}{\sigma^2}\right).$$

with

$$Z_\delta = N\delta - \frac{1}{2}\frac{N^2\big(SE(\delta)\big)}{\sigma^2}(\lambda^2 - \sigma^2).$$

$$\sigma^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}E({\varepsilon_t}^2).$$

$$\lambda^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}E(\frac{1}{N}\sum_{t=1}^{N}{\varepsilon_t}^2).$$

where

z_t is the t ratio of t.

The Phillips-Perron test is a negatively skewed one-tailed test. The null hypothesis of non-stationarity (existence of unit-root) is rejected if $Z_t < Z_\alpha$ and p-value $\lessdot$ 0.05.

3.3 Results

At a 95% confidence level, only three stocks (Datatec Ltd, KAP Industrial Holdings Ltd and Pick n Pay Stores Ltd) rejected the null hypothesis of a unit root. Datatec Ltd (DTC) was the only stock that rejected the null hypothesis of a unit root at a 99% confidence level. In the study of stationarity based on weekly returns collected from 23 February 1973 and 5 April 2002 by Mangani (2007), Pick n Pay Stores Ltd (PIK) rejected the null hypothesis of non-stationarity. This outcome suggests that PIK has not evolved with time from stationarity to non-stationarity.

The empirical results from the analysis are consistent with other previously documented JSE related studies such as that of Mangani (2007) and Van Rensburg (1999). The failure to reject the null hypothesis at a 95% confidence level indicates the existence of a unit root at both the 95% and 99% confidence levels. These results imply the following:

1. Existence of non-stationarity in the stock price.
2. Less likelihood of mean-reversion in the stock price, and hence the existence of Random Walk in stock prices.

In general, most of the stocks and equity indices included in this study exhibit non-stationarity. A complete set of results is in Table (3.2) – Table (3.5) under Appendix I. Table (3.1) displays stocks that rejected the null hypothesis of non-stationarity.

		Critical Value		**-3.41**	**-3.96**
		Confidence Level		**95%**	**99%**
Share code	tau-statistic	Lag	p-value	H0	H0
DTC	-4.7955	10	0.0100	Reject	Reject
KAP	-3.4265	10	0.0493	Reject	Accept
PIK	-3.5000	10	0.0423	Reject	Accept

Table 3.1: Phillips-Perron test results of stocks that rejected the null hypothesis of non-stationarity at a 95% confidence level.

Figure (3.1) – Figure (3.3) graphical display daily price charts of two stocks that rejected the null hypothesis of non-stationarity either at a 99% or 95% confidence level. The stock that failed to reject the null hypothesis of non-stationarity is also included. The existence of a meandering trend is visible for stocks that reject the null hypothesis of non-stationarity (Datatec Ltd and KAP Industrial Holdings Ltd).

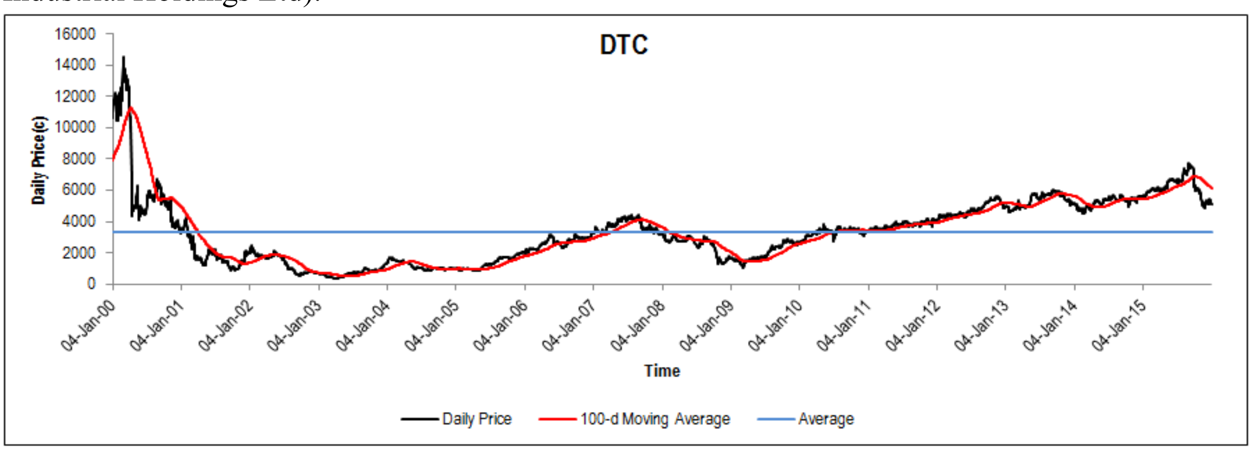


Figure 3.1: Datatec Ltd (DTC) daily price movement from 04 January 2000 to 31 December 2015.

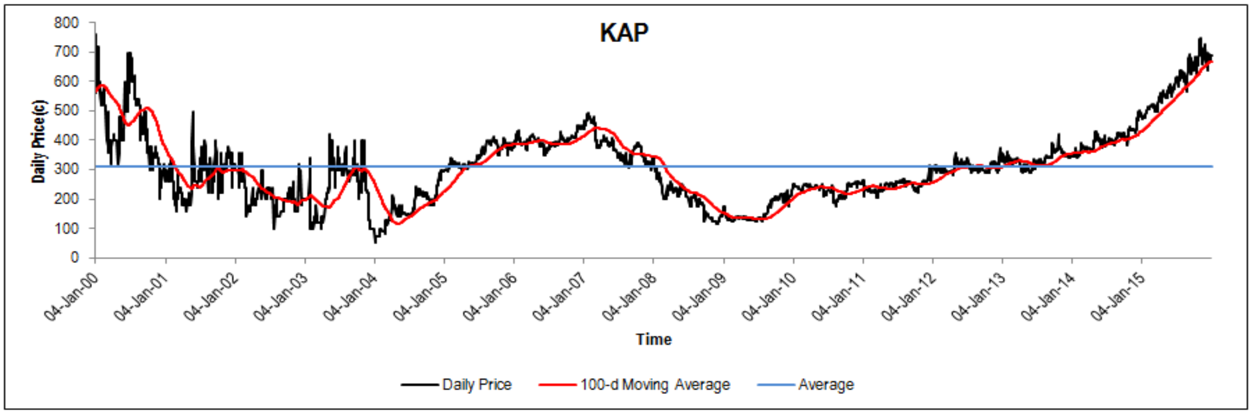


Figure 3.2: KAP Industrial Holdings Ltd (KAP) daily price movement from 04 January 2000 to 31 December 2015.

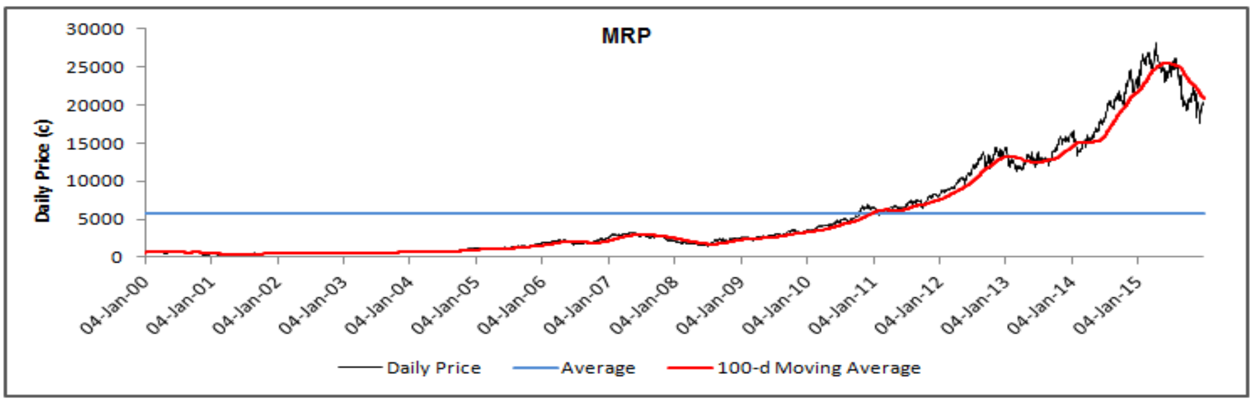


Figure 3.3: Mr Price Group Ltd (MRP) daily price movement from 04 January 2000 to 31 December 2015.

Figure (3.4) – Figure (3.7) display a stock (Nedbank Group Ltd) that failed to reject the null hypothesis of non-stationarity, but graphically displays the phenomenon of meanderness. The reason that this stock failed to reject the null hypothesis could be due to the low power of the Phillips-Perron Test to detect stationarity in the presence of structural breaks.

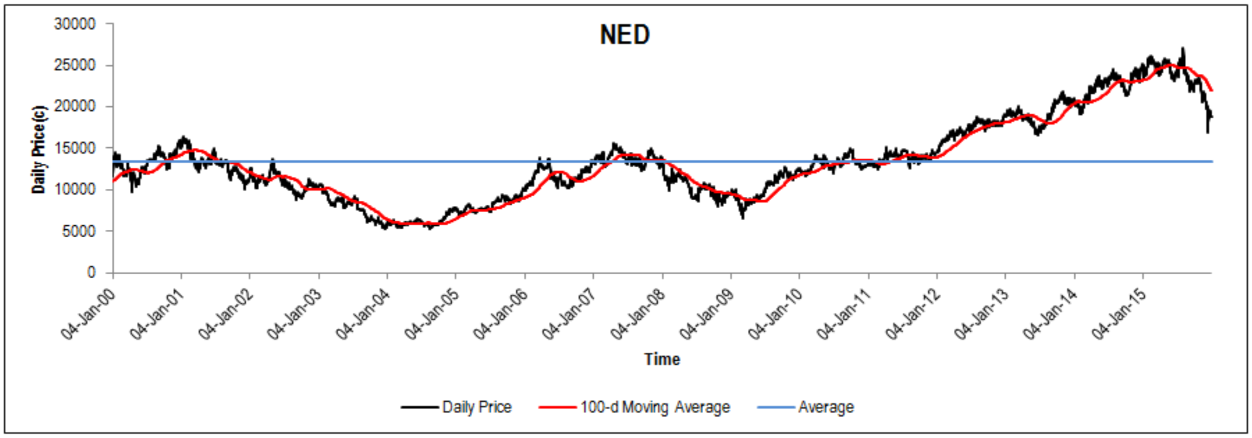


Figure 3.4: Nedbank Group Ltd (NED) daily price movement from 04 January 2000 to 31 December 2015.

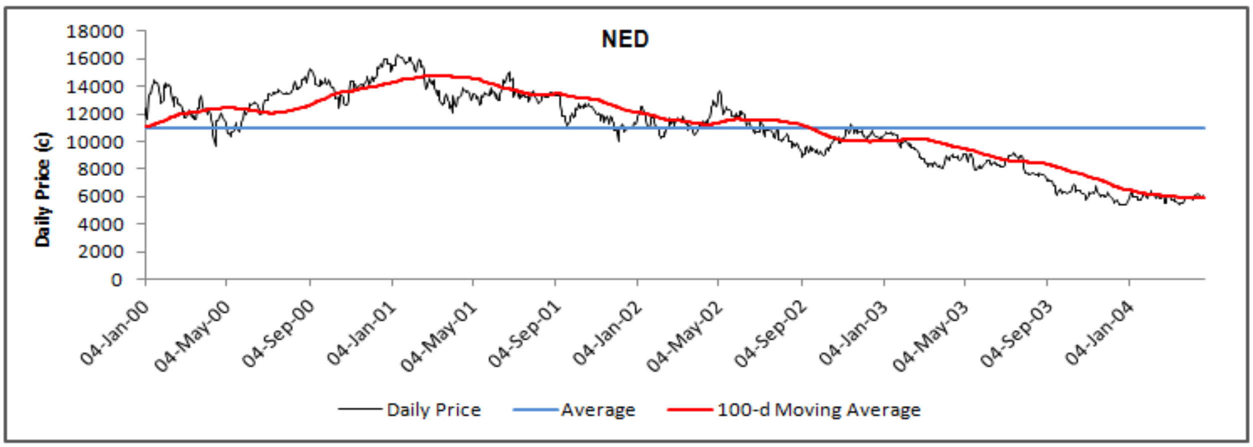


Figure 3.5: Nedbank Group Ltd (NED) daily price movement from 04 January 2000 to 22 April 2004.

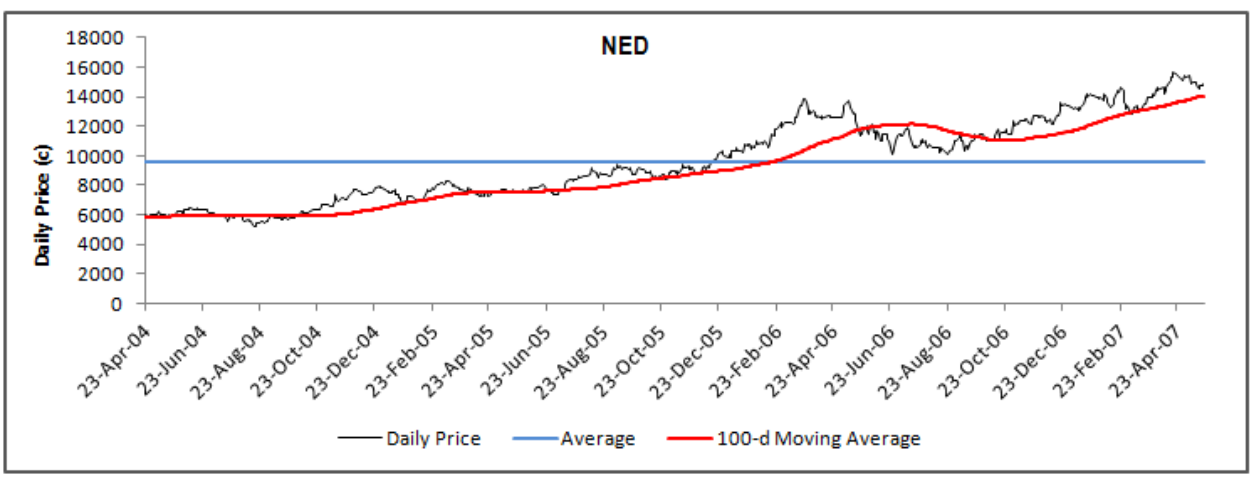


Figure 3.6: Nedbank Group Ltd (NED) daily price movement from 23 April 2004 to 22 May 2007.

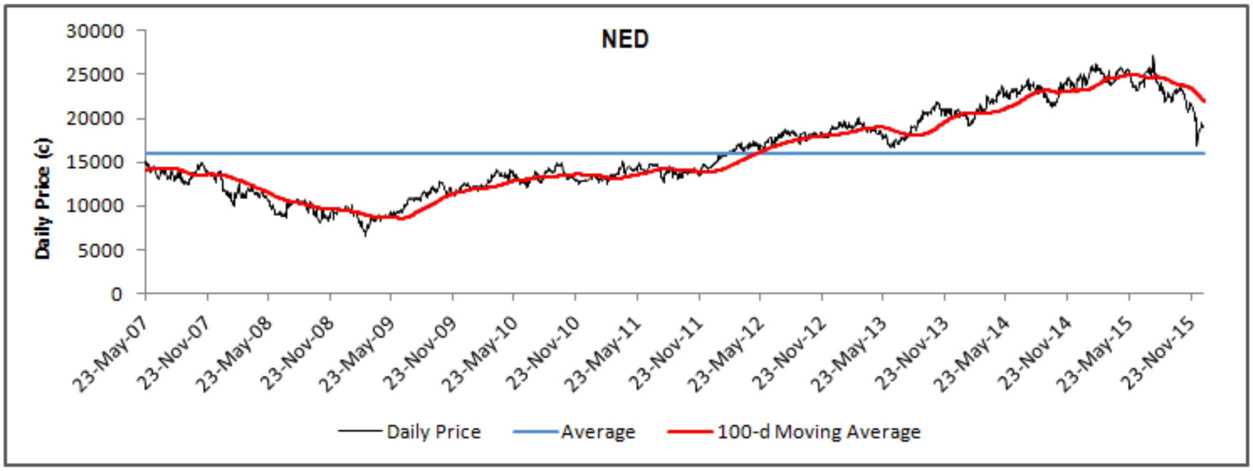


Figure 3.7: Nedbank Group Ltd (NED) daily price movement from 23 May 2007 to 31 December 2015.

3.4 Summary

This chapter looked at one of the most important principles in financial time series, namely, stationarity. A financial time series that exhibits stationarity violates the Random Walk Hypothesis. A stationary process implies existence of mean-reversion and a higher likelihood of predictability of price/returns. A predictable time series violates the Weak Form test hypothesis, which states that if a market is Weak Form efficient it is impossible to use historical prices to predict future prices.

Three stocks rejected the null hypothesis of non-stationarity at the 95% confidence level. Only one stock rejected the null hypothesis of non-stationarity at the 99% confidence level. These results imply that most of the stocks and indices included in this study are non-stationary, and these results are consistent with that of Mangani (2007) and Van Rensburg (1999). However caution has to be exercised because of the Phillips-Perron test low power to detect stationarity in the presence of structural breaks.

Chapter 4: Test for Independence

4.1 Introduction

This chapter tested the critical problem of dependency in daily returns of stocks and equity indices. The study of independence is crucial for the hypothesis of Market Efficiency and the Random Walk Model. If stock returns are dependent, there will exist a level of predictability in the direction of the underlying asset price and mean-reversion. It is possible to create a deterministic model from a dependent time series. A dependent time series is not likely to follow the Random Walk process and would violate the Weak Form test of the Efficient Market Hypothesis. Also, the use of technical analysis may yield profits when applied on a time series that is predictable or dependent. Two tests of independency were used namely, Autocorrelation (serial correlation) and the Ljung-Box test.

4.2 Methodology and data

Two random variables (X and Y) are said to be independent if they are uncorrelated, that is

$\rho_{x,y} = 0.$

where

$\rho_{x,y}$ represents the correlation between variable X and Y.

Hence if

$\mathrm{P}(X, Y) = \mathrm{P}(X)\mathrm{P}(Y)$

and

$X \in A, Y \in B$ then $\mathrm{P}(X, Y) \in A \times B$.

If $f(x, y)$ is a joint pdf of (X and Y)

$f(x, y) = g(x)h(y)$.

where

g is the pdf of X.

and

h is the pdf of Y.

Then X and Y are statistically independent.

The data sample used was made up of daily returns calculated from daily prices as described in Table 2.1 above.

4.2.1 Autocorrelation (Serial correlation) test

Correlation is a measure of independence between two or more random variables (X and Y). If the correlation coefficient $\rho = 0$ then X and Y are independent, i.e. X and Y are uncorrelated, given that X and Y have a bivariate normal distribution.

The correlation between random variable X and Y is formulated as follows:

$$\rho_{x,y} = \frac{Cov\ (X,Y)}{\sqrt{Var(X)}\sqrt{Var\ (Y)}}. \quad (10)$$

The correlation coefficient is between -1 and 1, $-1 \leq \rho \leq 1$.

where

$\rho = 0$ two variables are uncorrelated.

$\rho \geq 0$ two variables exhibit positive correlation.

$\rho \leq 0$ two variables exhibit negative correlation.

$\rho_{xy} = \rho_{yx}$.

Alternatively

If $Cov\ (X, Y) = 0$.

Then the random variables X and Y are mutually independent.

When one random variable X is studied and correlation is used as a measure of linear dependence between the values of the same time series, this is called autocorrelation or serial correlation. The Autocorrelation test, which has been frequently used in the past to test independency in time series, was deployed. It is a powerful test and is useful to detect stationarity in the time series. The Autocorrelation Function (ACF) converges to zero geometrically if the time series is stationary (Enders, 2004). Mlambo and Biekpe (2007) used the Autocorrelation Function (ACF) to test for serial correlation in ten African stock markets returns.

The correlation coefficient between x_t and x_{t-l} is called the lag_l autocorrelation of the financial time series formulated (Campbell, Lo and MacKinlay, 1997).

$$\rho_l = \frac{cov(x_t, x_{t-l})}{\sqrt{Var\ (x_t)Var(x_{t-l})}},$$

$$= \frac{cov(x_t, x_{t-l})}{\sqrt{Var\ (x_t)}}.$$

Since $Var\ (x_t) = Var(x_{t-l})$.

The autocorrelation coefficient is between -1 and 1, meaning, $1 \leq \rho \leq 1$.
$\rho_0 = 1$ and $\rho_l = \rho_{-l}$.

When the autocorrelation coefficients of a financial time series returns is close to zero at all lags, such a series is considered to be independent. However, the major drawback lies with the number of lags used. If too few lags are used, the presence of a higher-order autocorrelation could be missed, and if too many lags are used, the test may not have much power due to insignificant higher-order autocorrelations (Campbell et al. (1997)).

$$S.E = \frac{1}{\sqrt{n}}.$$

<u>Hypothesis test</u>

H_0 : No first order autocorrelation ($\rho_1 < abs(2\ x\ S.E)$).
H_1 :Existance of first order autocorrelation ($\rho_1 > abs(2\ x\ S.E)$).

The null hypothesis is rejected if

$$|\ t\ | > |\ t_{\frac{1-\alpha}{2}}\ |.$$

$$t = \frac{\sqrt{N - \rho_1}}{\sqrt{1 - R^2}}.$$

where

ρ_1 is first order autocorrelation.

R^2 is the r-squared.

4.2.2 The Ljung-Box test

The Ljung-Box test is an asymptotically distributed, two-sided hypothesis test of Random Walk that was introduced by Greta M. Ljung and George E. P. Box in 1978 (Box and Ljung, 1978). This test has a robust feature to detect departures from zero serial correlation in either direction and at all lags. However, the number of lags included in the test is important. If too few are used, this could result in higher-order serial correlation distortion. On the other hand, if too many lags are used it can lead to lower test power due to an insignificant higher-order autocorrelation (Campbell et al., 1997). Mlambo and Biekpe (2007) applied the Ljung-Box test to test the independency of ten African stock markets returns. This test was run for one degree of freedom and eighteen degrees of freedom, both at a 95% confidence level. (Refer to Appendix III for the derivation of the Ljung-Box test statistic.)

The Ljung-Box test statistic is given by Box and Ljung (1978):

$$Q_l = N(N+2)\sum_{i=1}^{l}\frac{{\rho_i}^2}{N-i}. \qquad (11)$$

with

$$\rho_i = \frac{\sum_{t=i+1}^{N} x_t x_{t-i}}{\sum_{t=1}^{N} {x^2}_t}.$$

where

N is the time series size.
ρ_i is the autocorrelation at lag i.
l is the number of lags being tested.

<u>Hypothesis test</u>

$H_0 : \mathrm{x_t}$ does not exhibit lack of fit, $\rho_i = 0$.
$H_1 : \mathrm{x_t}$ does exhibit lack of fit, $\rho_i \neq 0$.

The null hypothesis is rejected if

$Q = {\chi^2}_{1-\alpha,h}$ with h degrees of freedom and significance level α.
$h = l - 2$, And p-value $\lessdot$ 0.05.

4.3 Results

4.3.1 Autocorrelation (Serial correlation) test results

The independence tests assess whether past stock returns have some form of correlation to future stock returns, or whether current stock returns can be used to predict future stock returns. On average, more large caps did not reject the null hypothesis of no first order serial correlation. However, as market capitalization decreased, more stocks failed to reject the null hypothesis of no first order serial correlation. This outcome suggests a positive relationship between market capitalization and independence. As size increased, so did the level of independence. This could be as a result of thin-trading in smaller caps.

Except for the FTSE/JSE Africa Small Cap Index, Datatec Ltd (DTC) and Wilson Bayly Holmes-Ovcon Ltd (WBO), there was prevalence of negative serial correlation between lag 2 and 5 in other stocks and indices. This was despite the fact that some stocks did not breach the lower boundary of the standard error. The presence of negative serial correlation in a stock was often associated with mean-reversion which is a violation of the Random Walk model (Leung and Chen 2010). Table (4.1) displays stocks and indices that failed to reject the null hypothesis of no first order serial correlation. A complete set of results is in Table (4.2) – Table (4.5) under Appendix I.

	2 x Std error = 0.0316307				p-value =	0.05
	Share Code	**1st Order**	**1st Order p-value**	**1st lag with AC**	**1st lag with AC p-value**	**H0**
Top 40	AGL	0.025	0.111	3	-0.070	Accept
	ANG	0.009	0.568	50	0.040	Accept
	BGA	0.031	0.051	2	-0.064	Accept
	BIL	0.002	0.922	3	-0.066	Accept
	CFR	-0.007	0.668	3	-0.048	Accept
	FSR	0.011	0.496	2	-0.076	Accept
	GRT	-0.003	0.868	13	-0.043	Accept
	INL	0.003	0.852	2	-0.036	Accept
	IPL	0.016	0.310	2	-0.042	Accept
	ITU	-0.013	0.396	6	-0.034	Accept
	NED	0.023	0.138	2	-0.061	Accept
	SAB	-0.006	0.693	2	-0.049	Accept
	SBK	-0.005	0.730	2	-0.092	Accept
	SHP	-0.014	0.368	2	-0.063	Accept
	SLM	-0.016	0.325	2	-0.075	Accept
	SNH	0.005	0.736	2	-0.074	Accept
	TBS	-0.012	0.451	2	-0.033	Accept
	WHL	-0.026	0.106	19	-0.041	Accept
Mid caps	AFE	-0.005	0.751	6	-0.042	Accept
	ASR	-0.019	0.236	17	0.040	Accept
	AVI	0.010	0.525	2	-0.040	Accept
	GFI	0.023	0.150	2	-0.047	Accept
	GND	0.024	0.123	3	-0.039	Accept
	HYP	-0.018	0.243	7	0.036	Accept
	PIK	-0.027	0.091	2	-0.059	Accept
	PPC	-0.028	0.143	2	-0.050	Accept
	RLO	0.002	0.910	4	-0.073	Accept
	SPG	-0.010	0.546	8	0.040	Accept
	TON	-0.009	0.566	10	0.040	Accept
	TRU	0.021	0.188	6	-0.056	Accept
Small caps	CLH	-0.007	0.643	8	-0.041	Accept
	HDC	-0.023	0.148	9	-0.034	Accept
	MRF	-0.010	0.524	10	-0.046	Accept
	PGR	-0.029	0.064	50	-0.034	Accept
	WBO	0.016	0.306	10	0.033	Accept

Table 4.1: Autocorrelation test results of stocks with daily returns that failed to reject the null hypothesis of no first order serial correlation.

AngloGold Ashanti Ltd (ANG) breaches the standard error threshold only at lag 50.

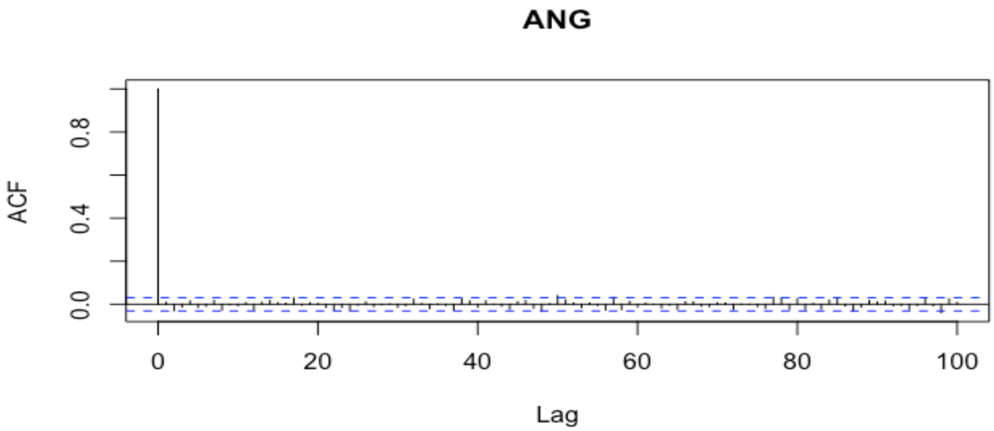


Figure 4.1: AngloGold Ashanti Ltd (ANG) Autocorrelation Function daily returns from 04 January 2000 to 31 December 2015.

Reunert Ltd (RLO) breaches the standard error threshold at lag 8 and it exhibits a negative correlation.

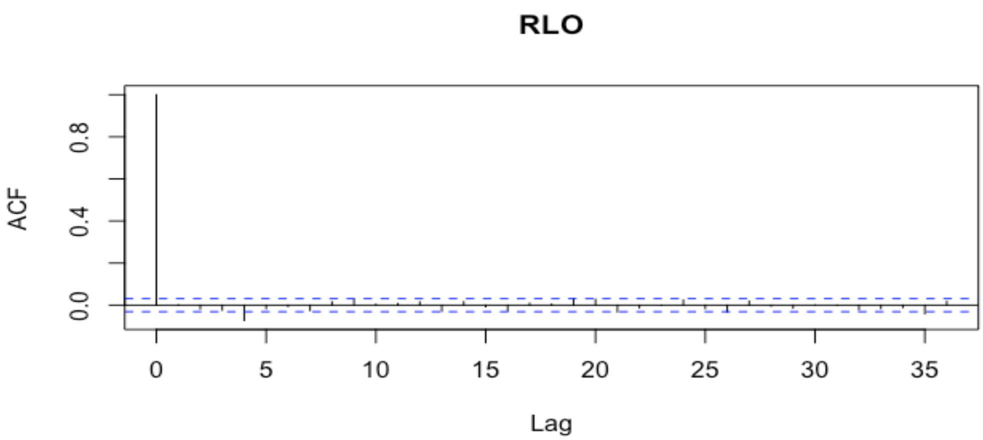


Figure 4.2: Reunert Ltd (RLO) Autocorrelation Function daily returns from 04 January 2000 to 31 December 2015.

4.3.2 The Ljung-Box test results

The Ljung-Box test with one degree of freedom mirrored the results of the Autocorrelation test. However, as the degrees of freedom were increased to eighteen, only a few stocks failed to reject the null hypothesis of independence (does not show a lack of fit). These results indicated the likelihood of serial correlation presence in almost all stocks at a higher number of lags. Table (4.6) and Table (4.7) display stocks/indices that failed to reject the null hypothesis of independence (does not lack of fit). The rest of the results are in Table (4.8) – Table (4.11) under Appendix I.

	df = 1			Chi^2 = 3,841	
		Confidence level		95%	99%
	Share/Index Code	X-Squared	p-value	H0	H0
Top 40	AGL	2.543	0.111	Accept	Accept
	ANG	0.327	0.568	Accept	Accept
	BGA	3.801	0.051	Accept	Accept
	BIL	0.010	0.922	Accept	Accept
	CFR	0.184	0.668	Accept	Accept
	FSR	0.464	0.496	Accept	Accept
	GRT	0.028	0.868	Accept	Accept
	INL	0.035	0.852	Accept	Accept
	IPL	1.032	0.310	Accept	Accept
	ITU	0.721	0.396	Accept	Accept
	NED	2.204	0.138	Accept	Accept
	SAB	0.156	0.693	Accept	Accept
	SBK	0.119	0.730	Accept	Accept
	SHP	0.811	0.368	Accept	Accept
	SLM	0.969	0.325	Accept	Accept
	SNH	0.113	0.736	Accept	Accept
	TBS	0.569	0.451	Accept	Accept
	WHL	2.610	0.106	Accept	Accept
Mid caps	AFE	0.101	0.751	Accept	Accept
	ASR	1.403	0.236	Accept	Accept
	AVI	0.404	0.525	Accept	Accept
	GFI	2.076	0.150	Accept	Accept
	GND	2.379	0.123	Accept	Accept
	HYP	1.364	0.243	Accept	Accept
	PIK	2.862	0.091	Accept	Accept
	PPC	2.147	0.143	Accept	Accept
	RLO	0.013	0.910	Accept	Accept
	SPG	0.364	0.546	Accept	Accept
	TON	0.330	0.566	Accept	Accept
	TRU	1.731	0.188	Accept	Accept
Small caps	CLH	0.215	0.643	Accept	Accept
	HDC	2.091	0.148	Accept	Accept
	MRF	0.406	0.524	Accept	Accept
	PGR	3.418	0.064	Accept	Accept
	WBO	1.048	0.306	Accept	Accept
Indices	Top 40	0.177	0.674	Accept	Accept
	JALSH	0.028	0.867	Accept	Accept
	JSMLC	1.809	0.179	Accept	Accept

Table 4.6: Ljung-Box test results of daily returns that failed to reject the null hypothesis of independence with one degree of freedom at 95% and 99% Confidence Level.

	df = 18			Chi^2 = 28.87	
	Confidence level			95%	99%
	Share/Index Code	X-Squared	p-value	H0	H0
Top 40	ANG	19.407	0.367	Accept	Accept
	IPL	24.153	0.150	Accept	Accept
	ITU	16.401	0.565	Accept	Accept
	MRP	25.119	0.122	Accept	Accept
	TBS	26.026	0.099	Accept	Accept
	WHL	11.623	0.866	Accept	Accept
Mid caps	AFE	27.493	0.070	Accept	Accept
	GFI	26.436	0.090	Accept	Accept
	HAR	23.485	0.173	Accept	Accept
	SNT	26.992	0.079	Accept	Accept
	TON	26.054	0.099	Accept	Accept
	TRE	27.214	0.075	Accept	Accept
Small caps	CLH	22.191	0.224	Accept	Accept
	HDC	21.755	0.243	Accept	Accept
	PGR	19.992	0.333	Accept	Accept
	WBO	11.433	0.875	Accept	Accept
Index	JSMLC	13.523	0.760	Accept	Accept

Table 4.7: Ljung-Box test results of daily returns that failed to reject the null hypothesis of independence with eighteen degrees of freedom 95% and 99% Confidence Level.

4.4 Summary

Another important test in financial time series is the test of independency. Two parametric tests were used: Autocorrelation and Ljung-Box. With one degree of freedom, the Ljung-Box test and Autocorrelation test results aligned for all stocks and indices. However, as the degrees of freedom were increased, more stocks failed to reject the null hypothesis of independence under the Ljung-Box test.

Overall, the empirical results showed that more stocks rejected the null hypothesis of independence (does not lack of fit). In terms of size, more large caps daily returns were tilted towards independence, relative to smaller caps. This observation revealed a positive relationship between market capitalization and independence. As market capitalization increased so did the likelihood of independence. Cohen et al. (1983) found that a lack of trading activity can induce autocorrelation in returns. Furthermore, Smith and Dyakova (2014) found that the degree of predictability is driven by size, liquidity and market quality.

The Autocorrelation test revealed a prevalence of negative serial correlation in most stocks. The presence of negative serial correlation is more problematic for the Weak Form efficiency or the Random Walk Hypothesis, as it is an indication of mean-reversion. In general all stocks and indices exhibited serial correlation at different lags. Negative serial correlation was found to be more prevalent in mid and small caps at the first lag.

Chapter 5: Test for normality

5.1 Introduction

In financial time series analysis, the goodness-of-fit measure is as important as the other two tests previously discussed. In a risk-neutral world it is often assumed that stock prices returns are normally distributed. However, it is often found that is not the case in the real world. The study of normality is linked to that of independence, and often the two tests are studied together. The study of independence is also linked to that of stationarity. Traditional asset-allocation frameworks such as the mean-variance theory assume that asset price returns are independent and normally distributed. In mathematical finance, pricing models such as the Black-Scholes model make an assumption that log returns of stock prices are independent and normally distributed. However, in reality asset returns often exhibit non-normality which can be detected due to the presence of fat-tails and negative or positive skewness.

Definition 5.1.1. $x_1, x_2, \dots, x_n$ are random variables and i.i.d with mean μ_n and variance σ^2.
$S_n = \sum_{i=1}^{n} x_i$,

$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$,
$n\mu = \sum_{i=1}^{n} x_i$.

$S_n^2 = \sum_{i=1}^{n} x_i^2$,
$\sigma^2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$,
$n\sigma^2 = \sum_{i=1}^{n} x_i^2$.

As $n \longrightarrow \infty$, $S_n \sim \boldsymbol{N}\ (n\mu, n\sigma^2)$
Therefore the sum of identically distributed variables is normally distributed.

A financial time series x_t is regarded as normally distributed if

$$x_t \sim \boldsymbol{N}(0, \sigma^2).$$

5.2 Methodology and data

In this chapter of the dissertation two normality tests were run to assess whether stocks and indices daily returns, forming part of this dissertation, are normally distributed. There are various methods that have been used in the past to test normality in a financial time series. These methods include graphical, numerical or formal parametric tests. In this dissertation, two parametric tests of normality were implemented, namely, the Jarque-Bera test and the Shapiro-Wilk test.

The data sample was made up of daily returns calculated from daily prices as described in Table 2.1 above. The test was done for both Jarque-Bera and Shapiro-Wilk at 95% and 99% confidence levels. The "tseries" RStudio package was used to compute Shapiro-Wilk test statistic, skewness and kurtosis. The final Jarque-Bera calculation was performed in Excel using RStudio skewness and kurtosis outputs.

5.2.1 Jarque-Bera test

The Jarque-Bera test was introduced by Carlos M. Jarque and Anil K. Bera in 1987 as a normality test for observations and regression residuals using the Lagrange multiplier. The Jarque-Bera test incorporates the statistical third and fourth moments as part of normality testing. The third moment (skewness) measures the degree of asymmetry of a probability distribution, while the fourth moment (kurtosis) measures the magnitude of "peakedness" of a probability distribution or the "thickness" of the tails of a probability distribution. The Jarque-Bera test is asymptotically distributed as a chi-squared with 2 degrees of freedom. The test is known to work well in large sample sizes (Jarque and Bera, 1987), however, it has poor power with short-tailed distributions, especially bimodal shaped distributions (Thadewald and Büning, 2007). The Jarque-Bera test is one of the most frequently used normality tests. In the South African context, its application is included in studies such as that of Phiri (2015), Smith and Dyakova (2014), and more.

The Jarque-Bera test statistic is given by (Jarque and Bera, 1987):

$$JB = \frac{N}{6}S^2 + \frac{N}{24}(K-3)^2, \tag{12}$$

$$= \frac{N}{6}\left(S^2 + \frac{(K-3)^2}{4}\right),$$

$$= \frac{N}{6}\left(S^2 + \frac{\gamma^2}{4}\right).$$

$$JB \sim {\chi^2}_2.$$

where

$\gamma = K - 3$ is excess kurtosis.

$S = \frac{\mu^3}{\sigma^3} = \frac{\mu^3}{{\mu_2}^{3/2}} = \frac{E(x-\bar{x})}{E[(x-\bar{x})^2]^{\frac{3}{2}}}$ is the skewness.

$K = \frac{\mu^4}{\sigma^4} = \frac{\mu^4}{{\mu_2}^2} = \frac{E(x-\bar{x})}{E[(x-\bar{x})^2]^2}$ is the kurtosis.

Under normality assumption:

$$S \sim \boldsymbol{N}\left(0, \frac{6}{N}\right).$$

$$K - 3 \sim \boldsymbol{N}\left(0, \frac{24}{N}\right).$$

Let's assume that x represents daily stock returns for all $x \in \mathbb{N}^+$

The third moment of the standardized random variable x is given by

$$S(x) = \frac{E(x-\bar{x})}{E[(x-\bar{x})^2]^{\frac{3}{2}}}.$$

while the fourth moment of the standardized random variable x is given by

$$K(x) = \frac{E(x-\bar{x})}{E[(x-\bar{x})^2]^2}.$$

<u>The third moment hypothesis test</u>

$H_0 : \mathrm{S}(x) = 0.$
$H_1 : S(x) \neq 0.$

H_0 is rejected if $|\,\mathrm{t}\,| > Z_{\frac{\alpha}{2}}$

where

$$t = \frac{S(x)}{\sqrt{\frac{6}{N}}}.$$

<u>The fourth moment hypothesis test</u>

$H_0 : \mathrm{K}(x) = 3.$
$H_1 : \mathrm{K}(x) \neq 3.$

H_0 is rejected if | t | $> Z_{\frac{\alpha}{2}}$.

where

$$t = \frac{K(x)-3}{\sqrt{\frac{24}{N}}}.$$

<u>The Jarque-Bera hypothesis test</u>

H_0 : $\mathrm{K}(x) = 3$ and $\mathrm{S}(x) = 0$ (The stock returns are normally distributed).
H_1 : $\mathrm{K}(x) \neq 3$ and $S(x) \neq 0$ (The stock returns are not normally distributed).

where

K(x) is kurtosis of time series x.
$S(x)$ is skewness of time series x.

The null hypothesis for Jarque-Bera is rejected if

$$JB > 5.991,$$

For 95% confidence level.
or

$$JB > 9.21,$$

For 99% confidence level.

5.2.2 Shapiro-Wilk test

The Shapiro-Wilk (SW) test was introduced by Samuel S. Shapiro and Martin Wilk in 1965, as part of the goodness-of-fit measure. The Shapiro-Wilk test has the ability to detect any departures from normality due to skewness or kurtosis. Its *t*-statistic is based on the regression of the order sample statistics on the expected normal order statistics. The Shapiro-Wilk test is known as one of the most powerful normality tests for all types of distributions (Razali and Wah, 2011) and is a good replacement for the Jarque-Bera test, on short-tailed distributions, especially bimodal distributions (Thadewald and Büning, 2007). However, its power can be low for small sample sizes (Razali and Wah, 2011). Refer to Appendix III for the W test statistic derivation.

Let x_t represent a financial time series, then the Shapiro-Wilk test statistic is given by (Shapiro and Wilk, 1965):

$$W = \frac{(\sum_{i=1}^{n} a_i x_{(i)})^2}{\sum_{i=i}^{n}(x_i - \bar{x}_n)^2}. \quad \in (0,1]. \tag{13}$$

where

$$(a_i, \dots, a_n) = \frac{m^T v^{-1}}{\sqrt{m^T v^{-1} v^{-1}}}.$$

$(m_i, \dots, m_n)$ represents the moments order statistics.

v is the $n \times n$ covariance matrix of $x_{(i)}$.

$$W = \frac{(\sum_{i=1}^{n} a_i x_{(i)})^2}{\sum_{i=i}^{n}(x_i - \bar{x}_n)^2} \approx \frac{\sigma^2(n-1)}{\sigma^2(n-1)} = 1.$$

Hypothesis test

$H_0 : W \approx 1$ (The stock returns are normally distributed).
$H_1 : W < 1$ (The stock returns are not normally distributed).

The null hypothesis is rejected if
$W < 1$.

And $p - value < 0.05$.

5.3 Results

5.3.1 Jarque-Bera test results

Under the Jarque-Bera test, only Barlow World Ltd (BAW) failed to reject the null hypothesis of normality at the 95% and 99% confidence levels. This means that this stock's daily returns were normally distributed at the 95% and 99% confidence levels. Murray & Roberts Holdings (MUR) failed to reject the null hypothesis of normality only at the 99% confidence level. A complete set of results is shown in Table (5.4) – Table (5.7) under Appendix I.

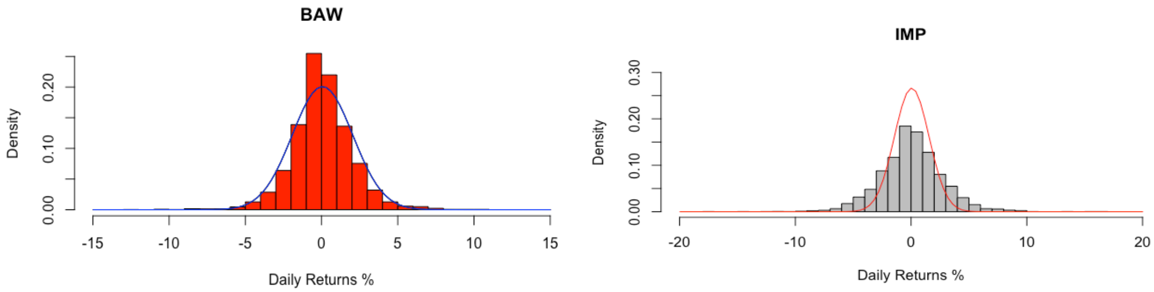


Figure 5.1: Barlow World Ltd (BAW) and Impala Platinum Holdings Ltd daily returns from 04 January 2000 to 31 December 2015.

Jarque-Bera test results demonstrated the presence of non-normality in most daily stocks and indices' prices returns. Except for a few large caps, most stocks and indices' daily prices returns were leptokurtic (kurtosis > 3, fat-tails). Leptokurtic returns imply a higher probability of outliers (extreme positive or negative returns) than that of a normally distributed process. On the other hand, the results also showed the significant exhibition of platykurtic returns in large caps. Table (5.1) displays stocks with playkurtic returns per capitalization category.

	Critical Value =				5.991	9.21
	Confidence Level =				95%	99%
	Share Code	**Skewness**	**Kurtosis**	**Jarque-Bera Statistic**	**H0**	**H0**
Top 40	AMS	0.032	2.447	51.616	Reject	Reject
	IMP	0.072	2.186	113.901	Reject	Reject
	IPL	0.102	2.540	42.262	Reject	Reject
	NED	0.193	2.712	38.752	Reject	Reject
	NTC	0.347	2.899	81.965	Reject	Reject
	RMH	0.244	2.630	62.469	Reject	Reject
	SBK	0.203	2.858	30.701	Reject	Reject
	SHP	0.134	2.297	94.258	Reject	Reject
	SLM	0.196	2.721	38.540	Reject	Reject
	SOL	0.196	2.538	61.333	Reject	Reject
	TBS	0.161	2.281	103.319	Reject	Reject
	WHL	0.208	2.319	106.128	Reject	Reject
Mid caps	CLS	0.027	2.731	12.572	Reject	Reject
	TRU	0.164	2.439	70.337	Reject	Reject
Small caps	MUR	0.110	2.935	8.717	Reject	Accept

Table 5.1: Jarque-Bera test results of platykurtic stock returns.

The FTSE/JSE Africa Top 40 Index was tilted towards the mesokurtic with skewness close to zero. On the other hand, the FTSE/JSE Africa Small Cap Index had an extreme negative skewness and was leptokurtic, even though most of its constituents were positively skewed. This phenomenon was driven by co-skewness and negative skewness in the underlying constituents of the FTSE/JSE Africa Small Cap Index carrying higher weights. As size decreased, more leptokurtosis was observed among the stocks. However, most stocks and indices returns were positively skewed and leptokurtic (fat right tails) in general, which implies log-normal distributions. Log-normally distributed returns suggest many small negative returns and a few extreme positive returns.

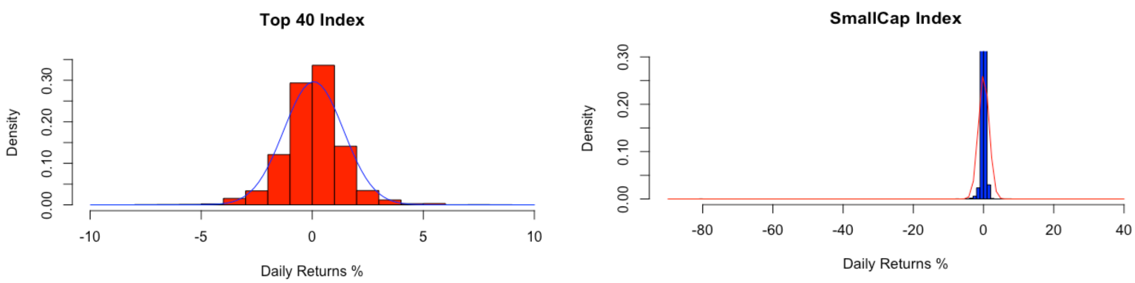


Figure 5.2: FTSE/JSE Africa Top 40 Index and FTSE/JSE Africa Small Cap Index daily returns from 04 January 2000 to 31 December 2015.

				Critical Value =	5.991	9.21
				Confidence Level =	95%	99%
	Share/Index Code	**Skewness**	**Kurtosis**	**Jarque-Bera Statistic**	**H0**	**H0**
Top 40	AMS	0.032	2.447	51.616	Reject	Reject
	BGA	0.046	4.513	382.756	Reject	Reject
Mid caps	ARI	-0.028	4.068	190.651	Reject	Reject
	BAW	-0.018	3.057	0.760	Accept	Accept
	CLS	0.027	2.731	12.572	Reject	Reject
	SAC	0.021	5.727	1239.078	Reject	Reject
Small caps	AEG	-0.043	4.393	324.292	Reject	Reject
Index	Top 40	0.014	3.294	12.296	Reject	Reject

Table 5.2: Jarque-Bera test results of stocks and index with skewness insignificant to zero.

	Share/Index Code	Skewness	Kurtosis	Jarque-Bera Statistic	H0	H0
	Critical Value =				5.991	9.21
	Confidence Level =				95%	99%
Top 40	ITU	-0.203	4.529	416.846	Reject	Reject
	MRP	-0.987	17.077	33661.828	Reject	Reject
Mid caps	DTC	-0.650	15.663	26991.865	Reject	Reject
	HYP	-0.061	10.984	10622.021	Reject	Reject
	NHM	-0.113	10.072	8339.675	Reject	Reject
	NPK	-0.214	3.673	105.788	Reject	Reject
	PSG	-0.320	51.119	385779.217	Reject	Reject
	SPG	-0.228	14.202	20936.476	Reject	Reject
	TFG	-0.228	4.220	282.793	Reject	Reject
	TON	-0.070	4.607	433.634	Reject	Reject
Small caps	AFX	-0.063	3.983	163.671	Reject	Reject
Indices	MidCap	-0.212	5.255	469.516	Reject	Reject
	JSMLC	-45.278	2446.235	871982062.057	Reject	Reject

Table 5.3: Jarque-Bera test results of stocks and indices with negative skewness.

5.3.2 The Shapiro-Wilk test results

The Shapiro-Wilk test rejected the null hypothesis of normality for all four indices and stocks, at 95% and 99% confidence levels. The test statistic for all indices and stocks was less than 1, and the p-value for all was less than $2{,}2 \times 10^{-16}$. The closer the test statistic (greater than 0.99) is to one, the higher the probability of the stock returns process being normally distributed, considering the p-value as well.

Overall, the Shapiro-Wilk test results were consistent with that of the Jarque-Bera test. The two tests have yielded non-normality in most of the stocks and indices' daily returns tested. A complete set of results is in Table (5.8) – Table (5.11) under Appendix I.

5.4 Summary

In this chapter the distribution profile of four indices and ninety JSE-listed stocks was tested. Two normality tests were implemented, namely, the Jarque-Bera and the Shapiro-Wilk. The normality tests results showed that the analyzed sample's return distribution profile was not normal. In most instances the stock returns' skewness was not equal to zero nor was kurtosis equal to three. The results of this chapter are congruent with other past studies that have been done. Bekaert and Harvey (2002) and Mlambo et al. (2003) found that emerging markets stock returns deviate from normality and exhibit asymmetric distribution with heavy tails. This supports Mangani (2007) and Page (1993) who found that the return distribution of JSE-listed stocks is not normally distributed.

The FTSE/JSE Africa Top 40 Index exhibited mesokurtosis, with skewness insignificant to zero, while the FTSE/JSE Africa Small Cap Index was leptokurtic with negative skewness. As size decreased more leptokurtosis among stocks was observable. Inversely, as size increased more platykurtosis was observable. Given the power of the Shapiro-Wilk test over the Jarque-Bera test, it is concluded that all stocks and indices are not normally distributed. Most stocks and indices' daily returns are positively skewed and leptokurtic. This implies asymmetric tails towards positive values, meaning that daily stocks and indices returns analyzed have a log-normal distribution profile.

Chapter 6: Structural breaks

6.1 Introduction

There is a vast amount of well-documented literature involving structural breaks or regime change. Structural breaks are associated with regime change or parameter instability in the regression. These changes are often driven by internal and/or external factors. Internal factors could be due to changes specific to the underlying security, while external factors could be of a systemic nature. Systemic related events include, but are not limited to, political instability, economic policy changes, tax reforms, financial crises, wars, natural disasters, and more. The presence of structural breaks or regime change in a financial regression can lead to distortion in the symmetry and tails of conditional distribution of returns. In this section of the dissertation the presence of structural breaks was tested using the Bai-Perron test of multiple structural breaks.

All major financial market events which could have had an external influence for a structural break, during the sample assessment timeframe (04 January 2000 – 31 December 2015), are listed in Table 6.1 below. Although most of these events were not South African stock market (JSE) specific, JSE-listed stocks were tested to ascertain whether they could have been impacted by the second-round effects. The analysis tested whether there is a link between major financial market events or company-specific events and structural breaks in the stocks.

Table 6.1: Past major financial markets events.

Event	Country of origination	Occurrence start date
Technology stocks bubble	United States	March 2000
911 terror attacks	United States	September 2001
Stock market downturn	US, Asia and Europe	October 2002
Chinese stock market bubble	China	February 2007
U.S bear market	United States	October 2007
Financial crisis	United States	September 2008
Dubai debt crisis	United Arab Emirates	November 2009
Eurozone debt crisis	Several EU members	April 2010
U.S flash crash	United States	May 2010
Global stock markets fall	Italy and Spain	August 2011
Chinese stock market crash	China	June 2015

Earlier studies focused more on the single structural break phenomena; these studies include work done by Chow (1960). Chow's study subdivided the regression into two subsamples then made a comparison between the sum of squared residuals fitting from the entire sample and the subsamples. Post Chow's work, others made significant improvements and contributions. These include the work of Andrews (1993), Andrews et al. (1996) and Bai (1997). Bai and Perron

(1998) utilized previous work done by others as a foundation and modeled a linear regression estimated by least squares for multiple structural breaks of unknown dates. Bai and Perron's work has become a pivotal tool in studies of multiple structural breaks of unknown dates. Other contributors on the topic of multiple structural breaks include the work of Garcia and Perron (1996), Cooper (1998), and more.

6.2 Methodology and data

To study potential structural breaks in the financial time series of ninety individual stocks and four equity indices listed on the JSE, the Bai-Perron test was deployed, utilizing the Bai (1997) intuitive methodology of sequential (one at a time) multiple breaks approach. Bai (1997) makes inferences in the presence of multiple breaks with unknown break dates.

The studies of structural breaks (regime changes) go hand-in-hand with that of unit root (stationarity). In the presence of structural breaks, most econometric tests fail to reject the null hypothesis of non-stationarity due to the low power of the test (Perron, 2005). Perron (2005) found that:

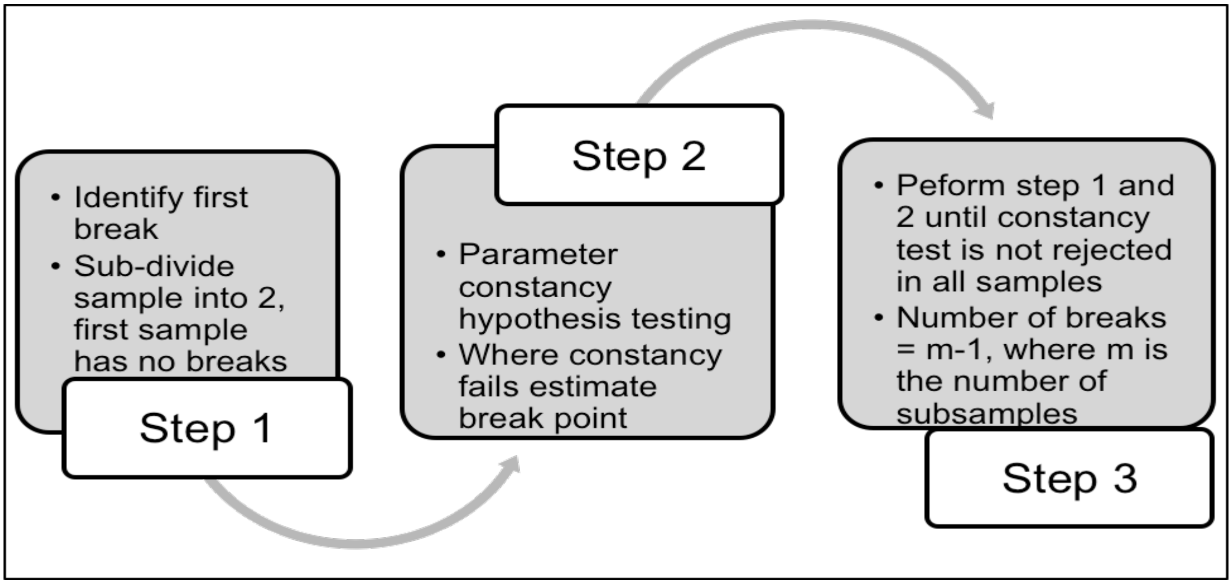
1) Most tests that attempt to distinguish between a unit root and a (trend) stationary process will favor the unit root model, when the true process is subject to structural changes but is otherwise (trend) stationary within regimes specified by the break dates.

2) Most tests trying to assess whether structural change is present will reject the null hypothesis of no structural change when the process has a unit root component but with constant model parameters.

The approach used is of sequential (one by one) estimation of multiple breaks as outlined by Bai (1997). This approach is known for its robustness to misspecification in the number of breaks and computational savings. The number of least-squares regressions required to compute all break points is of order T, the sample size (Bai,1997). Limited literature exists regarding structural breaks in the South African context.

To perform the test of multiple structural breaks, a MATLAB code developed by Yohei Yamamoto, 20 June 2012 was used. Yohei's MATLAB code is a direct and complete translation of the Gauss codes of Bai and Perron (1998), and Bai and Perron (2003). (Refer to Appendix III and Appendix V for the detailed Bai-Perron test description and MATLAB code, respectively). The code was run twice per stock's absolute returns and index's absolute returns. The first run allows for heterogeneity and autocorrelation in the residuals at 90%, 95%, 97.5% and 99%, while the second run disallows heterogeneity and autocorrelation in the residuals at 90%, 95%, 97.5% and 99% confidence levels.

Data analyzed included absolute daily returns calculated from daily prices as presented in Table 2.1 above. The MATLAB code determines the number of unknown breaks, limited to a maximum of five structural breaks. A three-step procedure is followed:



Consider a simple model with mean shifts in a linear process; first begin with the model with two breaks (Bai, 1997):

$$Y_t = \mu_1 + X_t \ ; \quad if\ t \le k_1^0 \tag{4}$$

$$Y_t = \mu_2 + X_t \ ; \quad if\ k_1^0 + 1 \le t \le k_2^0$$

$$Y_t = \mu_3 + X_t \ ; \quad if\ k_2^0 + 1 \le t \le T$$

where

μ_i is mean of regime i (i = 1,2,3).

X_t is linear process of martingale differences such that $X_t = \sum_{j=0}^{\infty} a_j \in_{t-j}$.

$k_1^0$and k_2^0 are unknown break points.

The idea of a sequential approach is to treat the model as if there is only one break at a time, which is obtained by minimizing the Sum of Squared residuals among all possible sample splits.

The Sum of Squared residuals is given by:

$$S_T(k) = \sum_{t=1}^{k}(Y_t - \bar{Y}_k)^2 + \sum_{t=k+1}^{T}(Y_t - \bar{Y}_k{}^*)^2 \tag{5}$$

where

$\bar{Y}_k$ is the mean of the first k observations.

$\bar{Y}_k^{\ *}$ is the mean of the last $T - k$ observations.

Let $\hat{k} = argmin_{1 \le k \le T-1} S_T(k)$ is the break point estimator, I assume that $\hat{k}$ is T-consistent

t = is an estimated break fraction

$$\sum_{t=1}^{T}(Y_t - \bar{Y})^2 \ = S_T(k) + \ TV_T(k)^2 \tag{6}$$

where

$\bar{Y}$ is the overall mean and

$V_T(k) = \left(\bar{Y}_k^{\ *} - \bar{Y}_k\right)\sqrt{\frac{k(T-k)}{T^2}}$.

Hence $\hat{k} = argmin_k S_T(k) = \ argmax_k V_T(k)^2$; $\hat{k} = T\hat{\tau}$

$\hat{k}$ and $\hat{\tau}$ are referred to as estimated break points.

Now consider the model with more than two breaks:

$$Y_t = \ \mu_1 + \ X_t \ \ ; \quad if \ t \le k_1^0$$

$$Y_t = \ \mu_2 + \ X_t \ \ ; \quad if \ k_1^0 + 1 \ \le t \le k_2^0$$

$$Y_t = \ \mu_{m+1} + \ X_t \ \ ; \quad if \ k_m^0 + 1 \ \le t \le T \tag{7}$$

where

$\mu_i \neq \mu_{i+1}$.

$k_i^0 = [T\tau_i^0]$; $\tau_i^0 \in (0,1)$ and $\tau_i^0 < \tau_{i+1}^0$.

For $i = 1, \dots m$ with $\tau_{m+1}^0 = 1$.

6.3 Results

Except for Mr Price Group Ltd (MRP), all stocks had at least one structural break. Firstly, there was a prevalence of between two and four structural breaks in most stocks. When it comes to indices, the FTSE/JSE Africa Mid Cap Index had two structural breaks, while the other indices had three breaks each. There were four stocks that produced five structural breaks. This number increased to five stocks when heterogeneity and autocorrelation in the residuals was disallowed. The stocks with five breaks are displayed in Table (6.2). A complete set of results is in Table (6.5) – Table (6.7) under Appendix I. The results presented are when heterogeneity and autocorrelation in the residuals was allowed at a 90% confidence level unless stated otherwise.

	Share Code	First break	Second break	Third break	Fourth break	Fifth break
Top 40	APN	30/05/2002	18/11/2004	30/11/2007	15/06/2010	11/02/2013
Mid caps	ACL	23/08/2002	29/06/2005	18/01/2008	23/07/2010	01/08/2013
	AVI	02/09/2002	10/02/2005	30/07/2007	22/12/2009	28/06/2012
	BAT*	05/06/2022	11/01/2005	09/01/2008	17/06/2010	16/05/2013
Small caps	AFX	05/08/2002	25/04/2005	10/01/2008	02/07/2010	23/04/2013

Table 6.2: Stocks with five structural breaks based on absolute daily returns between 04 January 2000 and 31 December 2015.

Secondly, the difference in the number of structural breaks is not major when heterogeneity and autocorrelation in the residuals was allowed or disallowed. The results show that most large caps and all indices produce the same number of structural breaks, whether heterogeneity and autocorrelation in the residuals was allowed or disallowed. Table (6.3) displays stocks that had additional structural breaks when heterogeneity and autocorrelation in the residuals was disallowed. The break dates when heterogeneity and autocorrelation in the residuals was allowed and disallowed were the same. The main difference came only with additional break dates, when heterogeneity and autocorrelation in residuals was disallowed, as per stocks in Table (6.3).

	Share Code	First break	Second break	Third break	Fourth break
Top 40	MDC	11/02/2004			
	ITU			30/05/2012	
	IPL				30/04/2013
	SOL				20/05/2013
Mid caps	BAT	05/06/2002	11/01/2005		
	ITE			26/07/2013	
	KAP			26/06/2007	11/12/2009
	LON				08/04/2013
	RCL				24/01/2013
	SNT				06/06/2012
Small caps	AEG	26/03/2003			
	AEN	02/03/2005			
	BRN		27/02/2009		
	CAT				06/08/2013
	HDC		15/07/2009	06/08/2013	
	MTA		15/07/2009	05/08/2013	
	PGR		16/02/2007		
	WBO				03/06/2013

Table 6.3: Stocks with additional structural breaks when heterogeneity and autocorrelation in the residuals is disallowed.

Thirdly, there were more structural breaks at lower confidence levels. This was evident in instances where the number of structural breaks was not the same across all confidence levels tests. Lastly, no direct relationship between major financial events and structural breaks in the stocks and indices was found instead second-round effects. There was also no correlation among stocks within the same sub-sectors. However, it was found that structural breaks in stocks were driven by company-specific events. Events such as dealings in company securities by directors, publication of good or bad news relating to the company, better or worse than expected financial results, planned mergers and acquisitions, and more. However, these kinds of events did not always lead to structural breaks whenever they occurred. It seemed they are influenced more by the general state of the market. A complete list of events around structural breaks dates per stock is in Appendix IV. Stocks which events or news occurring around their break dates could not be found were: RMH, SAB, DST, FBR, EOH, KAP, SNT, TSH, ACT and MUR.

Index Code	First break	Second break	Third break
JALSH	12/10/2005	12/12/2007	06/07/2010
Top 40	04/07/2007	15/07/2009	30/11/2011
MIDCAP	22/09/2009	18/09/2014	
JSMLC	11/05/2006	10/06/2010	28/03/2013

Table 6.4: Indices structural breaks based on 04 January 2000 to 31 December 2015 absolute returns, at 90%, 95%, 97.5% and 99% confidence levels.

6.4 Summary

The results showed that most JSE-listed stocks have structural breaks, and that structural breaks were driven more by company-specific events. Events such as dealing in securities by company directors, announcements about investments, mergers and acquisitions, bad or good company related news or events, and more, can lead to structural breaks. However, occurrence of these events linked to structural breaks was influenced by other external factors such as the global or local markets outlook or the general state of the local or global market. The results also revealed no strong correlation of structural breaks in stocks within the same sub-sector class. Lastly, more structural breaks were produced at lower confidence levels and when heterogeneity in the residuals was disallowed.

Chapter 7: Arbitrage opportunities

7.1 Introduction

Any stock market in the world is made up of at least three types of participants, also known as traders. These are speculators, arbitrageurs and hedgers.

- Speculators can be likened to players at a casino; these participants bet on the future direction of an asset price/market.
- Arbitrageurs search for and exploit mispricing and inefficiencies.
- Hedgers are risk management professionals.

These three above mentioned participants all have a different but important role to play. Arbitrageurs and speculators can cause price destabilization at times, but their role is crucial to maintain and sustain market efficiency and drive prices back to the equilibrium when deviations occur.

The Law of One Price states that financial instruments with identical payoff profiles or economic attributes should have the same price at any given point in time, e.g. the price of gold should be the same in all markets, after considering the exchange rate and cost of carry. Violation of the Law of One Price could lead to arbitrage opportunities. Pure arbitrage is defined as a riskless profit at zero upfront cost. Many strategies labeled as arbitrage are in reality quasi-arbitrage. They are not pure arbitrage, as their profits are not riskless.

Definition 7.1.1 Let V define a portfolio of any financial securities

If

$$V(0) = 0 \quad \text{; at } t = 0$$
$$V(t) \geq 0 \quad \text{; at } t = n$$

Then V is an arbitrage for all t, and has positive probability ($V(t) \geq 0$) for some t.

Definition 7.1.2 (Delbaen and Schacharmayer, 1994) A market ($\mathbb{F}, S$) satisfies no arbitrage on [0,T] if the process S together with the filtration $\mathbb{F}$, satisfies no arbitrage, given Theorem 2.2.1.

7.1.1 The concept of Pairs trading

There is a vast amount of well-documented literature regarding arbitrage. This topic is closely related to market efficiency theory, as it is believed that arbitrage strategies are more profitable in less efficient markets. In this section of the dissertation the focus is on relative value arbitrage strategy namely, Pairs trading, a branch of statistical arbitrage strategy. Since its discovery in the 1980s by the Wall Street quants led by Nunzio Tartaglia at Morgan Stanley, Pairs trading strategy has become popular with a lot of hedge fund managers. Pairs trading strategy is a rule-based strategy that uses statistical methods and requires little to no human intervention for the decision-making process of buying and selling stocks with similar economic attributes. Simply put, it is a market neutral trading strategy that works well in the presence of mean-reversion. The challenge is for the pair to converge, as money can be lost when the convergence trade diverges. This means that the price of a sold stock continues to rise while the price of a purchased stock continues to fall.

A pairs trade is intended to exploit any mispricing or market inefficiencies. A trade is initiated by first identifying two stocks or financial instruments with similar economic attributes and that have historically moved together. If at any stage, the absolute spread between the price of the two stocks or instruments is above a certain threshold, a relatively outperforming (price is higher) stock or instrument gets sold while a relatively underperforming stock or instrument (price is lower) gets bought. The view is that the two stocks or instruments will converge to their historical behaviour of symmetry, meaning that the outperforming stock or instrument price will fall and the underperforming stock or instrument price will rise until the two stocks reach their equilibrium again.

Past JSE-listed stocks pairs trading studies done include those of Govender (2011), Masindi (2014), Augustine (2014) and Mashele et al. (2013). These studies showed significant profitability, after trading costs. Mashele et al. (2013) also found that trading costs have to be at most 20bps for the pairs trading strategy to be profitable.

7.2 Methodology and data

A Pairs trading strategy can be likened to a self-financing trading strategy, initiated when one stock is sold and the other is bought. The proceeds from the sale are used for the purchase of the other stock, however in reality the proceeds from the sale are not always sufficient for the purchase of the other stock.

7.2.1 Self-financing strategy (trade)

Assume $x(t)$ represents stock A.
$y(t)$ represents stock B.

The value of the self-financing trade is denoted by

$$V(t) = x(t) - y(t). \tag{14}$$

where

$x(t)$ is bought and $y(t)$ is sold.

7.2.2 Pairs selection

A simple selection process was followed by identifying stocks based on these rules:

1. All stocks within the same sector were selected.
2. All stock prices were indexed to 100.
3. The spread of indexed values between two stocks at a time was computed.
4. An average of the spread through out the sample analysis period (04 January 2000 to 31 December 2015) was taken.
5. A graph of the average of the spread against the pairs' indexed values was drawn.
6. Pairs with a mean-reverting behavior were then selected for the strategy implementation.

For an example, below are two stocks, AngloGold Ashanti Ltd and Merafe Resources Ltd, from the same sectors that passed the abovementioned requirements:

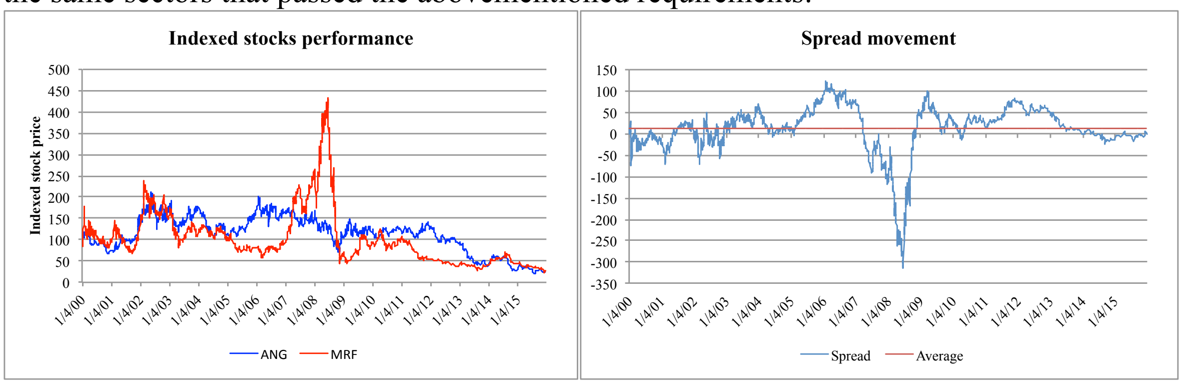


Figure 7.1: Performance of AngloGold Ashanti Ltd (ANG) vs. Merafe Resources Ltd (MRF) and their indexed prices spread movement from 04 January 2000 to December 2015.

The next pair failed to meet the above-mentioned requirements. The spread graph should revert to the red line, however, the FirstRand Ltd vs. RMB Holding Ltd spread failed to revert in the time period of testing. In this regard, the trader might have to wait for longer for convergence to occur which could lead to a lot of money being lost.

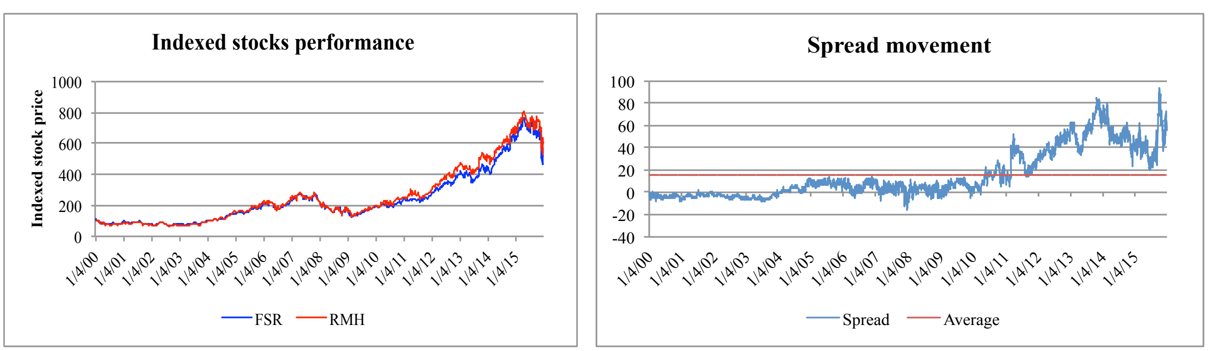


Figure 7.2: Performance of FirstRand Ltd (FSR) vs. RMB Holdings Ltd (RMH) and their indexed prices spread movement from 04 January 2000 to December 2015.

7.2.3 Indexation

Before the methodology implementation, it is important that the stocks time series are brought to a particular unit through indexation. Indexation is done by bringing all stocks within the same sector to a unit of 100 as the starting point, and re-calculate the stock price daily move from the indexed value of 100.

$$\ddot{x}_{it} = \ddot{x}_{it-1}\,(1 + r_{it}).$$

where

$\ddot{x}_{it}$ is the indexed price of stock i at time t.

r_{it} is the return of stock i between time t and t-1. $r_{it} = \ln \frac{x_{it}}{x_{it-1}}$.

7.2.4 Pairs trading implementation

Finally, the strategy was implemented based on the following set of rules:

1. Trade signal is triggered when the difference between the spread and the average value was greater than 1.
2. When the trade signal was triggered, a stock that was overvalued got sold and the stock that was undervalued got bought.
3. For simplicity and to minimize trading costs, each stock kept the same profile throughout the analysis. This means that a trade signal was triggered again when the stock that was undervalued before became undervalued again and vice versa.
4. It is assumed that trading was done at end of day closing prices.
5. The trade got exited when the difference between the spread and the average fell below the threshold of 1.
6. Trading costs were capped at 20 bps per trade, to avoid too much trading costs.

7.3 Results

The results are presented by using an average return per trade as a performance metric for the Pairs trading strategy. The results presented took 20 bps trading costs into account but ignored any balance costs.

If $r_c < |r_s| + |r_l|$, then r_c is unprofitable

where

r_c is the pairs trading (combination of long and short) average return per trade.

r_l is the long only average return per trade.

r_s is the short only average return per trade.

All pairs that formed part of the analysis produced profitable results, net of trading costs. There were instances where a long only strategy would have produced better results than the Pairs trading strategy. This suggests that on average, such pairs were more upward trending during the testing period. It was also observable that the pairs' spread oscillation around the mean varied with time and the general state of the market. The strategy also had periods of underperformance as the spread widens further and required longer holding period. This observation supports the theory of Adaptive Market Hypothesis (AMH), which states that markets are not always efficient but usually competitive and adaptive, varying in their degree of efficiency as the environment and investors change over time. Investment strategies undergo cycles of profitability and loss in response to changing business conditions, number of competitors, and arbitrage opportunities do exist from time to time but disappear as they are discovered and exploited (Lo, 2012). Results of the pairs' performance are displayed in Table (7.1) below.

Sector	Pair	Average returns	Number of trades
Basic Material	AGL vs. ANG	2.31%	140
	AGL vs. GFI	3.16%	172
	AMS vs. GFI	3.89%	92
	ANG vs. MRF	5.04%	240
	GFI vs. ANG	2.36%	164
	GFI vs. MRF	5.65%	260
	NHM vs. IMP	3.13%	120
Consumer Goods	AVI vs. OCE	3.67%	176
	CFR vs. SNH	7.68%	64
	ILV vs. TON	2.07%	124
	SAB vs. CFR	2.99%	220
	TBS vs. CFR	18.02%	48
Consumer Services	CBS vs. WHL	11.26%	104
	SUI vs. ADH	6.45%	196
	TSH vs. SUI	3.04%	260
	WHL vs. MRP	6.48%	172
Financials	BGA vs. SBK	1.82%	252
	NED vs. INL	2.79%	88
Health Care	NTC vs. MDC	2.47%	256
Industrials	AEL vs. AEG	6.52%	136
	AEN vs. AEG	9.78%	248
	BAW vs. IPL	2.66%	136
	BAW vs. NPK	3.04%	124
	PPC vs. AEG	11.02%	124
	PPC vs. AEL	3.61%	152
	PPC vs. AEN	4.85%	144
	PPC vs. GRF	4.37%	328
	TRE vs. BVT	3.94%	200

Table 7.1: Pairs trading strategy average cumulative returns net trading cost, from 04 January 2000 to 31 December 2015.

7.4 Summary

The results showed that the Pairs trading strategy was profitable on average. However, there were other pairs where the Pairs trading strategy underperformed long only holdings, these pairs were found to be upward trending on average, during the testing period. The results showed that the Pairs trading strategy can yield positive returns after considering trading costs of 20 bps, and the strategy does go through different cycles of outperformance and underperformance. On average the strategy made at least 2.00% per trade. Graphical results showed that the performance of the strategy goes through different cycles. There were periods where the Pairs trading strategy outperformed and underperformed long and/or short only. This observation supports the Adaptive Market Hypothesis that investment strategies go through cycles of outperformance and underperformance (Lo, 2004). The main draw back about this strategy is that it is backward looking, however, it is possible to modify it to be forward-looking.

Chapter 8: Conclusion

In this dissertation market efficiency of ninety JSE-listed stocks and four major equity indices was investigated, using data sample of daily stock and index prices collected between 4 January 2000 and 31 December 2015. In order to have an inclusive and comprehensive study, four major equity indices were selected, namely, FTSE/JSE Africa All Share Index, FTSE/JSE Africa Top 40 Index, FTSE/JSE Africa Mid Cap Index and FTSE/JSE Africa Small Cap Index, and all stocks that were listed on the Johannesburg Stock Exchange that met the following requirements:

- Listed and traded on the Johannesburg Stock Exchange (JSE) as of 04 January 2000.
- No trading disruptions, such as stock suspension except for market wide suspensions during the assessment period.
- Companies did not merge or were not acquired by another during the assessment period.

The aim of the dissertation was to present two schools of thought namely, the Efficient Market Hypothesis (EMH) and the Adaptive Market Hypothesis (AMH) in a unifying manner. The analysis was subdivided into two categories, namely, the Efficient Market Hypothesis (Weak Form) and the Adaptive Market hypothesis.

Under the Efficient Market Hypothesis, the focus was on the Weak Form test. Stocks and indices were tested for stationarity, independence and normality in order to conclude on whether the stocks and indices forming part of the analysis were Weak Form efficient. These three tests are pivotal in financial time series and have previously been used together to test for the Weak Form efficiency. The Phillips-Perron test was used to assess stationarity in daily stocks and indices price. To assess independency in daily returns the Autocorrelation and Ljung-Box tests were used, while normality was assessed through the implementation of the Shapiro-Wilk and Jarque-Bera tests.

For the Adaptive Market Hypothesis (AMH) structural breaks in stocks and indices were tested through the implementation of structural breaks test called the Bai-Perron and potential drivers of structural breaks were also studied. Furthermore, eligible stocks were tested for potential arbitrage opportunities through the implemention of the Pairs trading strategy. In addition, it was assessed whether those opportunities did yield profits, after taking transaction costs into consideration.

Stationarity – empirical results for the Weak Form test found that most stocks and indices failed to reject the null hypothesis of non-stationarity at 95% and 99% confidence levels. A stationary process implies existence of mean-reversion and a higher likelihood of predictability of price or returns. A predictable time series violates the Weak Form efficiency test, which states that if a market is Weak Form efficient it is impossible to use historical prices to predict future prices. However, given the low power of the Phillips-Perron test in the presence of structural breaks, the structural breaks test was incorporated to draw a more solid conclusion. After incorporating structural breaks test, the Bai-Perron, it was found that besides Mr Price Group Ltd, all stocks exhibited structural breaks, driven primarily by company-specific events.

Independence – the empirical results showed that more stocks rejected the null hypothesis of independence. In terms of size, more large caps daily returns were tilted towards independence, relative to smaller caps. This observation revealed a positive relationship between market capitalization and independence. As market capitalization increased so did the likelihood of independence. Lack of independence in smaller caps was likely driven by thin-trading. Overall, most stocks and indices rejected the null hypothesis of independence, exhibited negative serial-correlation and tended to revert to the mean.

Normality – the empirical results showed that the stocks and indices profile distribution of daily returns is non-normal. The Shapiro-Wilk test, which is the most powerful between the two tests of normality deployed, rejected the null hypothesis of normality for all stocks and indices. Most stocks and indices were positively skewed with heavy tails, which implied log-normal distribution. It was also found that as size decreased more leptokurtosis among stocks was observable. Inversely, as size increased more platykurtosis was observable.

Structural breaks – empirical results showed that besides Mr Price Group Ltd (MRP), the rest of the stocks and indices had at least one structural break. More structural breaks were observable at lower confidence levels. These structural breaks were driven primarily by company-specific events, such as, securities dealings by company directors, mergers and acquisitions, and more. Furthermore, no strong correlation was found between structural breaks and stocks within the same sub-sector classification.

Pairs trading – overall Pairs trading strategy produced at least 2.00% per trade on average of cumulative returns, net of 20bps trading costs. However, there were a few stocks that were more upward trending during the sampling period, which means a long only strategy would have performed better than the Pairs trading strategy. The strategy did also reveal different cycles of underperformance and outperformance as time varied.

Given all these empirical results, it was concluded that JSE-listed stocks and indices, that formed part of this dissertation, were not always Weak Form efficient, rather went through different cycles of efficiency and inefficiency. This conclusion supports the Adaptive Market Hypothesis (AMH). The use of arbitrage strategies can yield positive returns, net of trading costs, from time to time. The conclusion is congruent with that of Appiah-Kusi and Menyah (2003), who also concluded that the JSE is Weak Form inefficient. Smith and Dyakova (2014) tested eight African stock markets including South Africa, and found that stock markets go through cycles of efficiency and inefficiency. Future research would explore the impact of structural breaks on arbitrage strategies profitability, using a forward-looking arbitrage strategy method.

Bibliography

Appiah-Kusi, Joe and Kojo Menyah, “Return Predictability in African Stock Markets.” *Review of Financial Economics*,12 (2003), pp. 247–70.

Andrews, Donald W.K., “Tests for Parameter Instability and Structural Change with Unknown Change Point.” *Econometrica*, 61:4 (July 1993), pp. 821–856.

Andrews Donald W. K., Inpyo Lee and Wener Ploberger, “Optimal Changepoint Tests For Normal Linear Regression.” *Journal of Econometrics*, 70:1(January 1996), pp. 9–38.

Augustine, Cecelia, “Pairs Trading: A Copula Approach.” Dissertation presented for the degree of Master of Philosophy in Mathematical Finance, Faculty of Commerce, University of Cape Town (2014).

Bachelier, Louis, “Théorie de la Spéculation.” (*Paris: Gauthier-Villars*), reprinted in English in [8](1900), pp. 17 – 78.

Bai, Jushan, “ Estimating Multiple Breaks One at a Time.” *Journal of Econometric Theory*, 13:3 (June 1997), pp. 315 – 352.

Bai, Jushan, “Estimation of a Change Point in Multiple Regression Models.” *Review of Economic and Statistics*, 79:4 (November 1997), pp. 551–563.

Bai, Jushan and Pierre Perron, “Estimating and Testing Linear Models with Multiple Structural Changes.” *Econometrica*, 66:1 (January 1998), pp. 47 – 78.

Bai, Jushan and Pierre Perron, “Computation and Analysis of Multiple Structural Changes Models.” *Journal of Applied Econometrics*, 18: (2003), pp. 1 – 22.

Bekaert, Geert and Campbell R. Harvey, “Research in Emerging Market Finance: Looking to the Future.” *Emerging Markets Review*, 3:4(September 2002), pp. 429-448.

Björk, Thomas, “Arbitrage Theory in Continuous Time.” Third edition, Oxford University Press Inc., New York, 2009. ISBN 978-0-19-957474-2.

Borges, Maria R., “Random walk test for the Lisbon stock market.” *Applied Economics*, 43 (2011), pp. 631 – 639.

Campbell, John Y., Andrew W. Lo and A. Craig MacKinlay, “The Econometrics of Financial Markets.” Princeton University Press, 1997.

Chow, Gregory C.,"Tests of Equality between Sets of Coefficients in Two Linear Regressions." *Econometrica*, 28:3(July 1960), pp. 591– 605.

Cohen, Kalman J., Gabriel A. Hawawini, Steven F. Maier, Robert A. Schwartz and David K. Whitcomb, “Friction in the Trading Process and the Estimation of Systematic Risk.” *Journal of Financial Economics*, 12 (1983), pp. 263 – 278.

Cooper, Suzanne J., “Multiple Regimes in U.S. Output Fluctuations.” *Journal of Business & Economic Statistics*, 16:1 (1998), pp. 92 – 100.

Cromwell, Jeff B., Walter C. Labys and Michel Terraza, Univariate Tests for Time Series Models, 1994.

DeJong, David N., John C. Nankervis, N.E. Savin and Charles H. Whiteman, “The Power Problems of Unit Root Test in Series with Autoregressive Errors.” *Journal of Econometrics*, 53:1 – 3(July – September 1992), pp. 323 – 343.

Delbaen, Freddy and Walter Schacharmayer, “Fundamental Theorem of Asset Pricing.” *Mathemathische Annalen*, 300 (1994), pp. 463 – 520.
Do, Binh and Robert Faff, “Are Pairs Trading Profits Robust to Trading Costs? “*The Journal of Financial Research*, 35:2 (June, 2012), pp. 261–287.
Enders, Walter, “Applied Econometric Time Series.” Second Edition, 2004.
Gundersen, Ruben J., “Statistical Arbitrage: High Frequency Pairs Trading.” MSc. Economics and Business Administration Financial Economics, Norwegian School of Economics, Spring (2014).
Ghazani, M. Majid and Mansour K. Araghi , “Evaluation of the Adaptive Market Hypothesis as an Evolutionary Perspective on Market Efficiency: Evidence from the Tehran Stock Exchange.” *Research in International Business and Finance*, 32 (2014), pp. 50–59.
Govender, Kieran, “Statistical Arbitrage in South African Financial Markets.” University of Cape Town (2011).
Fama, Eugene F., “Efficient Capital Markets: A Review of Theory and Empirical Work.” *The Journal of Finance*, 25:2 (May, 1970), pp. 383 – 417.
Fama, Eugene F. and Kenneth R. French, “Common Factors in the Serial Correlation of Stock Returns.” Center for Research in Security Prices, Working Paper 200, October 1986b.
Garcia, René and Pierre Perron, “An Analysis of the Real Interest Rate under Regime Shifts.” *Review of Economics and Statistics*, 78 (1996), pp. 111 – 125.
Fries, Christian, “Mathematical Finance: Theory, Modeling, Implementation.” Wiley, Hoboken, (2007). ISBN 978-0-470-04722-4.
Hamilton, James D., “Time Series Analysis.” Princeton University Press, Princeton, (1994). ISBN 0-691-04289-6.
Higgs, Helen and Andrew Worthington, “Random Walks and Market Efficiency in European Equity Markets.” *Global Journal of Finance and Economics*, 1:1 (2004), pp. 59 – 78.
Hudson, Robert and Viktor Manahov, “A Note on the Relationship between Market Efficiency and Adaptability – New Evidence from Artificial Stock Markets.” *An International Journal*, 41 (2014), pp. 7436 – 7454.
Jarrow, Robert A., and Martin Larsson, “The Meaning of Market Efficiency.” *Mathematical Finance*, 22:1 (January 2012), pp. 1 – 30.
Jarque, Carlos M and Anil K., “A Test for Normality of Observations and Regression Residual.” *International Statistical Review*, 55:2(August 1987), pp. 163 –172.
Jefferis, Keith and Graham Smith, “The Changing Efficiency of African Stock Markets.” *South African Journal of Economics*, 73:1 (March 2005), pp. 54 – 67.
Jovanovic, Franck and Philippe Le Gall, “Does God practice a random walk? The 'financial physics' of a nineteenth-century forerunner, Jules Regnault.” *Euro. J. History of Economic Thought*, 8:3 (Autumn, 2001), pp. 332–362.
Leung, Charles Ka Yui and Nan-Kuang Chen, “Stock Volatility, Negative Autocorrelation and the Consumption Wealth Ratio: The Case of Constant Fundamentals.” *Pacific Economic Review*, 15:2 (2010), pp. 224– 245.

Ljung, Greta M. and George E.P. Box, "On a Measure of Lack of Fit in Time Series Models." *Biometrika*, 65:2(1978), pp. 297 – 303.
Lo, Andrew W., "The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective." *The Journal of Portfolio Management*, 30th Anniversary Issue 2004.
Lo, Andrew W., " Adaptive Markets hypothesis and the New World Order." *Financial Analyst Journal*, 68:2 (2012), pp. 18 – 29.
Lo, Andrew W. and A. Craig MacKinlay, "Stock Market Prices do not follow Random Walks: Evidence from a Simple Specification Test." *The Review of Financial studies*, 1:1 (Spring, 1988), pp. 41– 66.
Magnusson, Magnus A. and Bruce Wydick, "How Efficient Are Africa's Emerging Stock Markets?" *Journal of Development Studies*, 38(2002), pp. 141–56.
Mangani, Ronald, "Distributional Properties of JSE Prices and Returns." *Investment Analysts Journal,* 66 (2007), pp. 57– 72.
Mashele, H.P., S.E. Terblanche and J.H. Venter, "Pairs Trading on the Johannesburg Stock Exchange." *Investment Analyst Journal*, 78 (2013), pp. 13 – 26.
Masindi, Khuthadzo, "Statistical Arbitrage in South African Equity Market." Dissertation presented for the degree of Master of Philosophy in Mathematical Finance, Faculty of Commerce, University of Cape Town (2014).
Mlambo, Chipo and Nicholas Biekpe, "The Efficient Market Hypothesis: Evidence from Ten African Stock Markets." *Investment Analysts Journal*, 66 (2007), pp. 5 – 18.
Mlambo, Chipo, Nicholas Biekpe and Eon vd M. Smit, "Testing the Random-Walk Hypothesis on Thinly-Traded Markets: The Case of Four African Stock Markets." *The African Finance Journal*, 5:1 (2003), pp. 16 –35.
Noda, Akihiko, "A test of the adaptive market hypothesis using a time–varying AR model in Japan." *Finance Research Letters*, 17 (2016), pp. 66–71.
Page, Mike, "The arbitrage pricing theory: An assessment of the robustness of empirical techniques employed under conditions of thin trading and in the presence of non-normalities." *Unpublished PhD Thesis*, University of Cape Town (1993).
Perlin, Marcelo S., "Evaluation of Pairs-Trading Strategy at the Brazilian Financial Market." *Journal of Derivatives & Hedge Funds*, 15 (2009), pp. 122–136.
Perron, Pierre, "Dealing with Structural Breaks." This paper was prepared for the Palgrave Handbook of Econometrics, Econometric Theory. Vol. 1(April 2005), Boston University.
Phillips, Peter C.B. and Pierre Perron, "Testing for a unit root in time series regression."*Biometrika*, 75:2(1988), pp. 335 – 346.
Phiri, Andrew, "Efficient Market Hypothesis in South Africa: Evidence from Linear and Nonlinear Unit Root Tests." *Managing Global Transitions*, 13:4 (2015), pp.369–387.
Poterba, James M. and Lawrence H. Summers, "Mean Reversion in Stock Prices: Evidence and Implications." Massachusetts Institute of Technology Department of Economics, Working Paper 457, August 1987.

Razali, Nornadiah Mohd and Yap Bee Wah, "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests." *Journal of Statistical Modeling and Analytics*, 2:1 (2011), pp 21 – 33.

Regnault, Jules, "Calcul des Chances et Philosophie de la Bourse." Mallet Bachelier and Castel, 1863.

Rice, John A., "Mathematical Statistics and Data Analysis.", Third edition, Duxbury, 2007, ISBN 0-534-39942-8.

Rodriguez, E., M., Aguilar–Cornejo R. Femat and J. Alvarez–Ramirez, J, "US Stock Market Efficiency over Weekly, Monthly, Quarterly and Yearly Time Scales." *Physica A*, 413 (2014), pp. 554–564.

Samuelson, Paul A., "Proof that Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review,* 6:2 (Spring,1965), pp. 41–49.

Shapiro, Samuel S. and Martin B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika*, 52:3/4 (1965), pp. 591– 611.

Smith, Graham and Aneta Dyakova,"African Stock Markets: Efficiency and Relative Predictability." *South African Journal of Economics*, 82:2 (June 2014), pp. 258 – 275.

Thadewald, Thorsten and Herbert Büning, "Jarque-Bera Test and its Competitors for Testing Normality – A Power Comparison." *Journal of Applied Statistics*, 34:1 (2007), pp. 87 – 105.

Tsay, Ruey S., "Analysis of Financial Time Series." Second Edition. Wiley, Hoboken, 2002. ISBN-13 978-0-471-69074-0.

Urquhart, Andrew and Frank McGroarty, "Calendar effects, market conditions and Adaptive Markey Hypothesis: Evidence from long–run U.S. data." *International Review of Financial Analysis*, 35(2014), pp. 154 –166.

Van Rensburg, Paul, "Macroeconomic identification of candidate APT factors on the Johannesburg Stock Exchange." *Journal for Studies in Economics and Econometrics*, 23:2 (1999), pp. 27–53.

Appendix I: Test results tables

1 Stationarity: Phillips-Perron Test results

	Critical Value			-3.41	-3.96
	Confidence Level			95%	99%
Share code	tau-statistic	Lag	p-value	H0	H0
AGL	-1.016	10	0.936	Accept	Accept
AMS	-1.404	10	0.831	Accept	Accept
ANG	-2.037	10	0.563	Accept	Accept
APN	-1.290	10	0.879	Accept	Accept
BGA	-3.397	10	0.054	Accept	Accept
BIL	-2.718	10	0.274	Accept	Accept
BVT	-1.328	10	0.863	Accept	Accept
CFR	-0.979	10	0.942	Accept	Accept
DSY	-0.377	10	0.987	Accept	Accept
FSR	-1.440	10	0.816	Accept	Accept
GRT	-2.465	10	0.382	Accept	Accept
IMP	-0.787	10	0.963	Accept	Accept
INL	-1.352	10	0.853	Accept	Accept
IPL	-1.804	10	0.661	Accept	Accept
ITU	-1.241	10	0.900	Accept	Accept
MDC	0.311	10	0.990	Accept	Accept
MRP	-1.294	10	0.878	Accept	Accept
MTN	-2.175	10	0.504	Accept	Accept
NED	-2.020	10	0.570	Accept	Accept
NPN	1.362	10	0.990	Accept	Accept
NTC	-1.166	10	0.912	Accept	Accept
OML	-0.692	10	0.971	Accept	Accept
RMH	-1.404	10	0.831	Accept	Accept
SAB	2.428	10	0.990	Accept	Accept
SBK	-3.196	10	0.089	Accept	Accept
SHP	-2.148	10	0.516	Accept	Accept
SLM	-1.652	10	0.726	Accept	Accept
SNH	0.967	10	0.990	Accept	Accept
SOL	-3.057	10	0.131	Accept	Accept
TBS	-2.838	10	0.224	Accept	Accept
WHL	-0.463	10	0.984	Accept	Accept

Table 3.2: FTSE/JSE Africa Top 40 Index constituents Phillips-Perron (PP) Test results.

	Critical Value			-3.41	-3.96
	Confidence Level			95%	99%
Share code	tau-statistic	Lag	p-value	H0	H0
ACL	-1.079	10	0.926	Accept	Accept
AFE	-2.081	10	0.544	Accept	Accept
ARI	-0.806	10	0.961	Accept	Accept
ASR	-0.343	10	0.989	Accept	Accept
AVI	-0.617	10	0.977	Accept	Accept
BAT	4.001	10	0.990	Accept	Accept
BAW	-1.758	10	0.681	Accept	Accept
CLS	-1.275	10	0.885	Accept	Accept
DST	-1.417	10	0.825	Accept	Accept
DTC	-4.796	10	0.010	Reject	Reject
EOH	0.024	10	0.990	Accept	Accept
FBR	-0.263	10	0.990	Accept	Accept
GFI	-2.432	10	0.395	Accept	Accept
GND	-1.106	10	0.922	Accept	Accept
HAR	-2.739	10	0.265	Accept	Accept
HYP	-1.288	10	0.880	Accept	Accept
ILV	-1.273	10	0.886	Accept	Accept
ITE	-0.796	10	0.962	Accept	Accept
KAP	-3.427	10	0.049	Reject	Accept
LBH	-2.195	10	0.496	Accept	Accept
LON	-1.441	10	0.815	Accept	Accept
NHM	-1.982	10	0.586	Accept	Accept
NPK	-2.169	10	0.507	Accept	Accept
OCE	-0.223	10	0.990	Accept	Accept
OMN	-1.460	10	0.807	Accept	Accept
PIK	-3.500	10	0.042	Reject	Accept
PPC	-0.843	10	0.958	Accept	Accept
PSG	1.026	10	0.990	Accept	Accept
RCL	-2.952	10	0.175	Accept	Accept
RLO	-2.677	10	0.292	Accept	Accept
SAC	-1.714	10	0.699	Accept	Accept
SAP	-1.977	10	0.588	Accept	Accept
SNT	-2.334	10	0.437	Accept	Accept
SPG	-0.977	10	0.943	Accept	Accept
SUI	-1.765	10	0.678	Accept	Accept
TFG	-2.563	10	0.340	Accept	Accept
TON	-2.348	10	0.431	Accept	Accept
TRE	-1.662	10	0.721	Accept	Accept
TRU	-2.720	10	0.273	Accept	Accept
TSH	-2.244	10	0.475	Accept	Accept

Table 3.3: FTSE/JSE Africa Mid Cap Index constituents Phillips-Perron (PP) Test results.

	Critical Value			-3.41	-3.96
	Confidence Level			95%	99%
Share code	**tau-statistic**	**Lag**	**p-value**	**H0**	**H0**
ACT	-2.109	10	0.532	Accept	Accept
ADH	1.768	10	0.990	Accept	Accept
ADR	-2.332	10	0.438	Accept	Accept
AEG	-0.385	10	0.987	Accept	Accept
AEL	-0.395	10	0.987	Accept	Accept
AEN	0.006	10	0.990	Accept	Accept
AFX	-1.186	10	0.909	Accept	Accept
BRN	-1.916	10	0.614	Accept	Accept
CAT	-3.267	10	0.076	Accept	Accept
CLH	-1.850	10	0.642	Accept	Accept
CSB	0.605	10	0.990	Accept	Accept
GRF	-1.264	10	0.890	Accept	Accept
HDC	-3.185	10	0.090	Accept	Accept
MRF	-2.220	10	0.485	Accept	Accept
MTA	-1.626	10	0.737	Accept	Accept
MUR	-1.025	10	0.935	Accept	Accept
OCT	-2.592	10	0.328	Accept	Accept
PGR	-0.388	10	0.987	Accept	Accept
WBO	-2.275	10	0.462	Accept	Accept

Table 3.4: FTSE/JSE Africa Small Cap Index constituents Phillips-Perron (PP) Test results.

	Critical Value			-3.41	-3.96
	Confidence Level			95%	99%
Index code	**tau-statistic**	**Lag**	**p-value**	**H0**	**H0**
Top 40	-2.292	9	0.455	Accept	Accept
JALSH	-2.268	9	0.465	Accept	Accept
MIDCAP	-2.723	8	0.272	Accept	Accept
JSMLC	-3.116	9	0.106	Accept	Accept

Table 3.5: Indices Phillips-Perron (PP) Test results.

2 Independency: Autocorrelation Test results

2 x Std error =	0.0316307		p-value =		0.05
Share Code	1st Order	1st Order p-value	1st lag with AC	1st lag with AC p-value	H0
AGL	0.025	0.111	3	-0.070	Accept
AMS	0.091	0.000			Reject
ANG	0.009	0.568	50	0.040	Accept
APN	0.033	0.035			Reject
BGA	0.031	0.051	2	-0.064	Accept
BIL	0.002	0.922	3	-0.066	Accept
BVT	-0.071	0.000			Reject
CFR	-0.007	0.668	3	-0.048	Accept
DSY	0.034	0.031			Reject
FSR	0.011	0.496	2	-0.076	Accept
GRT	-0.003	0.868	13	-0.043	Accept
IMP	0.072	0.000			Reject
INL	0.003	0.852	2	-0.036	Accept
IPL	0.016	0.310	2	-0.042	Accept
ITU	-0.013	0.396	6	-0.034	Accept
MDC	-0.058	0.000			Reject
MRP	0.048	0.003			Reject
MTN	0.050	0.002			Reject
NED	0.023	0.138	2	-0.061	Accept
NPN	0.058	0.000			Reject
NTC	-0.045	0.004			Reject
OML	-0.032	0.040			Reject
RMH	-0.078	0.000			Reject
SAB	-0.006	0.693	2	-0.049	Accept
SBK	-0.005	0.730	2	-0.092	Accept
SHP	-0.014	0.368	2	-0.063	Accept
SLM	-0.016	0.325	2	-0.075	Accept
SNH	0.005	0.736	2	-0.074	Accept
SOL	0.042	0.007			Reject
TBS	-0.012	0.451	2	-0.033	Accept
WHL	-0.026	0.106	19	-0.041	Accept

Table 4.2: FTSE/JSE Africa Top 40 Index constituents Autocorrelation test results.

2 x Std error =		0.0316307		p-value =	0.05
Share Code	1st Order	1st Order p-value	1st lag with AC	1st lag with AC p-value	H0
ACL	0.08	0.000			Reject
AFE	-0.005	0.751	6	-0.042	Accept
ARI	0.058	0.000			Reject
ASR	-0.019	0.236	17	0.040	Accept
AVI	0.010	0.525	2	-0.040	Accept
BAT	0.055	0.001			Reject
BAW	0.052	0.001			Reject
CLS	0.045	0.005			Reject
DST	-0.108	0.000			Reject
DTC	0.099	0.000			Reject
EOH	-0.042	0.008			Reject
FBR	-0.109	0.000			Reject
GFI	0.023	0.150	2	-0.047	Accept
GND	0.024	0.123	3	-0.039	Accept
HAR	0.042	0.008			Reject
HYP	-0.018	0.243	7	0.036	Accept
ILV	0.041	0.009			Reject
ITE	-0.101	0.000			Reject
KAP	-0.117	0.000			Reject
LBH	-0.06	0.000			Reject
LON	0.166	0.000			Reject
NHM	0.109	0.000			Reject
NPK	-0.049	0.002			Reject
OCE	-0.08	0.000			Reject
OMN	0.034	0.032			Reject
PIK	-0.027	0.091	2	-0.059	Accept
PPC	-0.028	0.143	2	-0.050	Accept
PSG	-0.093	0.000			Reject
RCL	-0.082	0.000			Reject
RLO	0.002	0.910	4	-0.073	Accept
SAC	-0.105	0.000			Reject
SAP	0.066	0.000			Reject
SNT	-0.058	0.000			Reject
SPG	-0.010	0.546	8	0.040	Accept
SUI	0.083	0.000			Reject
TFG	0.057	0.000			Reject
TON	-0.009	0.566	10	0.040	Accept
TRE	-0.042	0.008			Reject
TRU	0.021	0.188	6	-0.056	Accept
TSH	-0.092	0.000			Reject

Table 4.3: FTSE/JSE Africa Mid Cap Index constituents Autocorrelation test results.

2 x Std error =	0.0316307			p-value =	0.05
Share Code	1st Order	1st Order p-value	1st lag with AC	1st lag with AC p-value	H0
ACT	-0.073	0.000			Reject
ADH	-0.119	0.000			Reject
ADR	-0.095	0.000			Reject
AEG	0.073	0.000			Reject
AEL	-0.131	0.000			Reject
AEN	-0.089	0.000			Reject
AFX	-0.043	0.006			Reject
BRN	-0.192	0.000			Reject
CAT	-0.208	0.000			Reject
CLH	-0.007	0.643	8	-0.041	Accept
CSB	-0.039	0.013			Reject
GRF	0.05	0.002			Reject
HDC	-0.023	0.148	9	-0.034	Accept
MRF	-0.010	0.524	10	-0.046	Accept
MTA	-0.193	0.000			Reject
MUR	0.054	0.001			Reject
OCT	-0.228	0.000			Reject
PGR	-0.029	0.064	50	-0.034	Accept
WBO	0.016	0.306	10	0.033	Accept

Table 4.4: FTSE/JSE Africa Small Cap Index constituents Autocorrelation test results.

p-value = 0.05						
Indiex Code	1st Order	1st Order p-value	1st lag with AC	1st lag with AC p-value	H0	2 x std error
JALSH	-0.003	0.867	2	-0.033	Accept	0.034
Top 40	0.007	0.674	3	-0.063	Accept	0.034
MIDCAP	0.095	0.000			Reject	0.043
JSMLC	0.023	0.179	3	0.040	Accept	0.034

Table 4.5: Indices Autocorrelation test results.

3 Independency: Ljung-Box Test results

df = 1			Chi^2 = 3,841	
	Confidence level		95%	99%
Share Code	X-Squared	p-value	H0	H0
AGL	2.543	0.111	Accept	Accept
AMS	32.914	0.000	Reject	Reject
ANG	0.327	0.568	Accept	Accept
APN	4.429	0.035	Reject	Accept
BGA	3.801	0.051	Accept	Accept
BIL	0.010	0.922	Accept	Accept
BVT	19.960	0.000	Reject	Reject
CFR	0.184	0.668	Accept	Accept
DSY	4.670	0.031	Reject	Accept
FSR	0.464	0.496	Accept	Accept
GRT	0.028	0.868	Accept	Accept
IMP	20.792	0.000	Reject	Reject
INL	0.035	0.852	Accept	Accept
IPL	1.032	0.310	Accept	Accept
ITU	0.721	0.396	Accept	Accept
MDC	13.232	0.000	Reject	Reject
MRP	9.110	0.003	Reject	Reject
MTN	10.025	0.002	Reject	Reject
NED	2.204	0.138	Accept	Accept
NPN	13.458	0.000	Reject	Reject
NTC	8.094	0.004	Reject	Reject
OML	4.223	0.040	Reject	Accept
RMH	24.371	0.000	Reject	Reject
SAB	0.156	0.693	Accept	Accept
SBK	0.119	0.730	Accept	Accept
SHP	0.811	0.368	Accept	Accept
SLM	0.969	0.325	Accept	Accept
SNH	0.113	0.736	Accept	Accept
SOL	7.184	0.007	Reject	Reject
TBS	0.569	0.451	Accept	Accept
WHL	2.610	0.106	Accept	Accept

Table 4.8: FTSE/JSE Africa Top 40 Index constituents Ljung-Box Test results.

df = 1		Chi^2 = 3,841		
	Confidence level		95%	99%
Share Code	X-Squared	p-value	H0	H0
ACL	25.494	0.000	Reject	Reject
AFE	0.101	0.751	Accept	Accept
ARI	13.306	0.000	Reject	Reject
ASR	1.403	0.236	Accept	Accept
AVI	0.404	0.525	Accept	Accept
BAT	12.093	0.001	Reject	Reject
BAW	10.748	0.001	Reject	Reject
CLS	7.942	0.005	Reject	Reject
DST	46.249	0.000	Reject	Reject
DTC	39.485	0.000	Reject	Reject
EOH	6.990	0.008	Reject	Reject
FBR	47.714	0.000	Reject	Reject
GFI	2.076	0.150	Accept	Accept
GND	2.379	0.123	Accept	Accept
HAR	7.093	0.008	Reject	Reject
HYP	1.364	0.243	Accept	Accept
ILV	6.840	0.009	Reject	Reject
ITE	40.994	0.000	Reject	Reject
KAP	54.400	0.000	Reject	Reject
LBH	14.489	0.000	Reject	Reject
LON	109.840	0.000	Reject	Reject
NHM	47.620	0.000	Reject	Reject
NPK	9.786	0.002	Reject	Reject
OCE	25.293	0.000	Reject	Reject
OMN	4.606	0.032	Reject	Accept
PIK	2.862	0.091	Accept	Accept
PPC	2.147	0.143	Accept	Accept
PSG	34.744	0.000	Reject	Reject
RCL	26.583	0.000	Reject	Reject
RLO	0.013	0.910	Accept	Accept
SAC	44.366	0.000	Reject	Reject
SAP	17.319	0.000	Reject	Reject
SNT	13.666	0.000	Reject	Reject
SPG	0.364	0.546	Accept	Accept
SUI	27.471	0.000	Reject	Reject
TFG	13.144	0.000	Reject	Reject
TON	0.330	0.566	Accept	Accept
TRE	7.067	0.008	Reject	Reject
TRU	1.731	0.188	Accept	Accept
TSH	33.744	0.000	Reject	Reject

Table 4.9: FTSE/JSE Africa Mid Cap Index constituents Ljung-Box Test results.

df = 1			Chi^2 = 3,841	
Confidence level			95%	99%
Share Code	**X-Squared**	**p-value**	**H0**	**H0**
ACT	21.037	0.000	Reject	Reject
ADH	56.207	0.000	Reject	Reject
ADR	35.989	0.000	Reject	Reject
AEG	21.227	0.000	Reject	Reject
AEL	68.980	0.000	Reject	Reject
AEN	31.553	0.000	Reject	Reject
AFX	7.540	0.006	Reject	Reject
BRN	147.540	0.000	Reject	Reject
CAT	173.810	0.000	Reject	Reject
CLH	0.215	0.643	Accept	Accept
CSB	6.210	0.013	Reject	Accept
GRF	9.987	0.002	Reject	Reject
HDC	2.091	0.148	Accept	Accept
MRF	0.406	0.524	Accept	Accept
MTA	148.790	0.000	Reject	Reject
MUR	11.768	0.001	Reject	Reject
OCT	207.230	0.000	Reject	Reject
PGR	3.418	0.064	Accept	Accept
WBO	1.048	0.306	Accept	Accept

Table 4.10: FTSE/JSE Africa Small Cap Index constituents Ljung-Box Test results.

df = 1			Chi^2 = 3,841	
Confidence level			95%	99%
Index Code	**X-Squared**	**p-value**	**H0**	**H0**
Top 40	0.177	0.674	Accept	Accept
JALSH	0.028	0.867	Accept	Accept
MIDCAP	19.381	0.000	Reject	Reject
JSMLC	1.809	0.179	Accept	Accept

Table 4.11: Indices Ljung-Box Test results.

4 Normality: Jarque-Bera Test results

Critical Value =				5.991	9.21
Confidence Level =				95%	99%
Share Code	**Skewness**	**Kurtosis**	**Jarque-Bera Statistic**	**H0**	**H0**
AGL	0.131	3.363	33	Reject	Reject
AMS	0.032	2.447	52	Reject	Reject
ANG	0.409	3.227	120	Reject	Reject
APN	0.736	8.207	4877	Reject	Reject
BGA	0.046	4.513	383	Reject	Reject
BIL	0.426	4.351	425	Reject	Reject
BVT	0.082	3.456	39	Reject	Reject
CFR	1.220	20.087	49630	Reject	Reject
DSY	0.251	5.464	1053	Reject	Reject
FSR	0.111	3.164	13	Reject	Reject
GRT	0.354	10.211	8745	Reject	Reject
IMP	0.072	2.186	114	Reject	Reject
INL	0.199	3.843	145	Reject	Reject
IPL	0.102	2.540	42	Reject	Reject
ITU	-0.203	4.529	417	Reject	Reject
MDC	0.263	3.038	46	Reject	Reject
MRP	-0.987	17.077	33662	Reject	Reject
MTN	0.427	5.082	844	Reject	Reject
NED	0.193	2.712	39	Reject	Reject
NPN	0.066	3.953	154	Reject	Reject
NTC	0.347	2.899	82	Reject	Reject
OML	0.144	6.017	1530	Reject	Reject
RMH	0.244	2.630	62	Reject	Reject
SAB	0.594	6.246	1990	Reject	Reject
SBK	0.203	2.858	31	Reject	Reject
SNH	0.301	5.069	773	Reject	Reject
SHP	0.134	2.297	94	Reject	Reject
SLM	0.196	2.721	39	Reject	Reject
SOL	0.196	2.538	61	Reject	Reject
TBS	0.161	2.281	103	Reject	Reject
WHL	0.208	2.319	106	Reject	Reject

Table 5.4: FTSE/JSE Africa Top 40 Index constituents Jarque-Bera Test results.

Critical Value =				5.991	9.21
Confidence Level =				95%	99%
Share Code	**Skewness**	**Kurtosis**	**Jarque-Bera Statistic**	**H0**	**H0**
ACL	0.850	20.522	51627	Reject	Reject
AFE	0.478	5.131	909	Reject	Reject
ARI	-0.028	4.068	191	Reject	Reject
ASR	1.151	14.185	21722	Reject	Reject
AVI	1.194	33.303	153922	Reject	Reject
BAT	0.285	7.573	3538	Reject	Reject
BAW	-0.018	3.057	1	Accept	Accept
CLS	0.027	2.731	13	Reject	Reject
DST	0.249	13.396	18046	Reject	Reject
DTC	-0.650	15.663	26992	Reject	Reject
EOH	1.016	11.171	11810	Reject	Reject
FBR	1.683	33.066	152471	Reject	Reject
GFI	0.471	4.009	317	Reject	Reject
GND	0.640	7.973	4392	Reject	Reject
HAR	0.379	3.486	135	Reject	Reject
HYP	-0.061	10.984	10622	Reject	Reject
ILV	0.165	3.179	23	Reject	Reject
ITE	1.526	34.827	170293	Reject	Reject
KAP	1.407	33.891	160287	Reject	Reject
LBH	1.194	21.538	58198	Reject	Reject
LON	4.372	113.564	2049102	Reject	Reject
NHM	-0.113	10.072	8340	Reject	Reject
NPK	-0.214	3.673	106	Reject	Reject
OCE	0.334	7.919	4106	Reject	Reject
OMN	1.921	37.393	199505	Reject	Reject
PIK	0.407	4.156	333	Reject	Reject
PPC	0.168	3.348	39	Reject	Reject
PSG	-0.320	51.119	385779	Reject	Reject
RCL	1.720	23.768	73821	Reject	Reject
RLO	0.257	3.017	44	Reject	Reject
SAC	0.021	5.727	1239	Reject	Reject
SAP	0.230	6.620	2218	Reject	Reject
SNT	0.198	6.343	1888	Reject	Reject
SPG	-0.228	14.202	20936	Reject	Reject
SUI	0.090	4.008	175	Reject	Reject
TFG	-0.228	4.220	283	Reject	Reject
TON	-0.070	4.607	434	Reject	Reject
TRE	0.143	14.478	21959	Reject	Reject
TRU	0.164	2.439	70	Reject	Reject
TSH	0.582	8.290	4888	Reject	Reject

Table 5.5: FTSE/JSE Africa Mid Cap Index constituents Jarque-Bera Test results.

Critical Value =				5.991	9.21
Confidence Level =				95%	99%
Share Code	**Skewness**	**Kurtosis**	**Jarque-Bera Statistic**	**H0**	**H0**
ACT	4.464	76.377	910197	Reject	Reject
ADH	1.224	14.451	22840	Reject	Reject
ADR	0.636	32.199	142296	Reject	Reject
AEG	-0.043	4.393	324	Reject	Reject
AEL	0.477	10.811	10315	Reject	Reject
AEN	0.231	6.629	2229	Reject	Reject
AFX	-0.063	3.983	164	Reject	Reject
BRN	7.255	210.110	7180607	Reject	Reject
CAT	0.929	19.151	44030	Reject	Reject
CLH	0.112	9.558	7172	Reject	Reject
CSB	1.055	31.089	132171	Reject	Reject
GRF	0.449	22.232	61749	Reject	Reject
HDC	1.063	13.327	18517	Reject	Reject
MRF	0.365	3.929	232	Reject	Reject
MTA	4.230	136.678	2988752	Reject	Reject
MUR	0.110	2.935	9	Reject	Accept
OCT	4.743	177.046	5061158	Reject	Reject
PGR	0.415	6.094	1710	Reject	Reject
WBO	0.288	3.667	129	Reject	Reject

Table 5.6: FTSE/JSE Africa Small Cap Index constituents Jarque-Bera Test results.

Critical Value =				5.991	9.21
Confidence Level =				95%	99%
Share Code	**Skewness**	**Kurtosis**	**Jarque-Bera Statistic**	**H0**	**H0**
Top 40	0.014	3.294	12	Reject	Reject
JALSH	0.692	17.295	29188	Reject	Reject
MIDCAP	-0.212	5.255	470	Reject	Reject
JSMLC	-45.278	2446.235	871982062	Reject	Reject

Table 5.7: Indices Jarque-Bera Test results.

5 Normality: Shapiro-Wilk Test results

Confidence Level =			95%	99%
Share Code	**W statistic**	**p-value**	**H0**	**H0**
AGL	0.967	2.2E-16	Reject	Reject
AMS	0.977	2.2E-16	Reject	Reject
ANG	0.966	2.2E-16	Reject	Reject
APN	0.945	2.2E-16	Reject	Reject
BGA	0.961	2.2E-16	Reject	Reject
BIL	0.961	2.2E-16	Reject	Reject
BVT	0.968	2.2E-16	Reject	Reject
CFR	0.936	2.2E-16	Reject	Reject
DSY	0.941	2.2E-16	Reject	Reject
FSR	0.974	2.2E-16	Reject	Reject
GRT	0.869	2.2E-16	Reject	Reject
IMP	0.982	2.2E-16	Reject	Reject
INL	0.959	2.2E-16	Reject	Reject
IPL	0.976	2.2E-16	Reject	Reject
ITU	0.952	2.2E-16	Reject	Reject
MDC	0.959	2.2E-16	Reject	Reject
MRP	0.917	2.2E-16	Reject	Reject
MTN	0.957	2.2E-16	Reject	Reject
NED	0.974	2.2E-16	Reject	Reject
NPN	0.959	2.2E-16	Reject	Reject
NTC	0.967	2.2E-16	Reject	Reject
OML	0.934	2.2E-16	Reject	Reject
RMH	0.974	2.2E-16	Reject	Reject
SAB	0.958	2.2E-16	Reject	Reject
SBK	0.973	2.2E-16	Reject	Reject
SHP	0.976	2.2E-16	Reject	Reject
SLM	0.974	2.2E-16	Reject	Reject
SNH	0.958	2.2E-16	Reject	Reject
SOL	0.974	2.2E-16	Reject	Reject
TBS	0.974	2.2E-16	Reject	Reject
WHL	0.977	2.2E-16	Reject	Reject

Table 5.8: FTSE/JSE Africa Top 40 Index constituents Shapiro-Wilk Test results.

Confidence Level =			95%	99%
Share Code	**W statistic**	**p-value**	**H0**	**H0**
ACL	0.884	2.2E-16	Reject	Reject
AFE	0.933	2.2E-16	Reject	Reject
ARI	0.949	2.2E-16	Reject	Reject
ASR	0.774	2.2E-16	Reject	Reject
AVI	0.896	2.2E-16	Reject	Reject
BAT	0.926	2.2E-16	Reject	Reject
BAW	0.968	2.2E-16	Reject	Reject
CLS	0.969	2.2E-16	Reject	Reject
DST	0.818	2.2E-16	Reject	Reject
DTC	0.860	2.2E-16	Reject	Reject
EOH	0.803	2.2E-16	Reject	Reject
FBR	0.743	2.2E-16	Reject	Reject
GFI	0.959	2.2E-16	Reject	Reject
GND	0.886	2.2E-16	Reject	Reject
HAR	0.964	2.2E-16	Reject	Reject
HYP	0.833	2.2E-16	Reject	Reject
ILV	0.961	2.2E-16	Reject	Reject
ITE	0.612	2.2E-16	Reject	Reject
KAP	0.595	2.2E-16	Reject	Reject
LBH	0.896	2.2E-16	Reject	Reject
LON	0.704	2.2E-16	Reject	Reject
NHM	0.930	2.2E-16	Reject	Reject
NPK	0.962	2.2E-16	Reject	Reject
OCE	0.859	2.2E-16	Reject	Reject
OMN	0.810	2.2E-16	Reject	Reject
PIK	0.959	2.2E-16	Reject	Reject
PPC	0.956	2.2E-16	Reject	Reject
PSG	0.829	2.2E-16	Reject	Reject
RCL	0.846	2.2E-16	Reject	Reject
RLO	0.961	2.2E-16	Reject	Reject
SAC	0.926	2.2E-16	Reject	Reject
SAP	0.940	2.2E-16	Reject	Reject
SNT	0.914	2.2E-16	Reject	Reject
SPG	0.870	2.2E-16	Reject	Reject
SUI	0.940	2.2E-16	Reject	Reject
TFG	0.971	2.2E-16	Reject	Reject
TON	0.948	2.2E-16	Reject	Reject
TRE	0.809	2.2E-16	Reject	Reject
TRU	0.975	2.2E-16	Reject	Reject
TSH	0.879	2.2E-16	Reject	Reject

Table 5.9: FTSE/JSE Africa Mid Cap Index constituents Shapiro-Wilk Test results.

Confidence Level =			95%	99%
Share Code	**W statistic**	**p-value**	**H0**	**H0**
ACT	0.531	2.2E-16	Reject	Reject
ADH	0.812	2.2E-16	Reject	Reject
ADR	0.805	2.2E-16	Reject	Reject
AEG	0.953	2.2E-16	Reject	Reject
AEL	0.866	2.2E-16	Reject	Reject
AEN	0.893	2.2E-16	Reject	Reject
AFX	0.947	2.2E-16	Reject	Reject
BRN	0.502	2.2E-16	Reject	Reject
CAT	0.787	2.2E-16	Reject	Reject
CLH	0.888	2.2E-16	Reject	Reject
CSB	0.793	2.2E-16	Reject	Reject
GRF	0.880	2.2E-16	Reject	Reject
HDC	0.834	2.2E-16	Reject	Reject
MRF	0.949	2.2E-16	Reject	Reject
MTA	0.559	2.2E-16	Reject	Reject
MUR	0.963	2.2E-16	Reject	Reject
OCT	0.624	2.2E-16	Reject	Reject
PGR	0.920	2.2E-16	Reject	Reject
WBO	0.943	2.2E-16	Reject	Reject

Table 5.10: FTSE/JSE Africa Small Cap Index constituents Shapiro-Wilk Test results.

Confidence Level =			95%	99%
Index Code	**W statistic**	**p-value**	**H0**	**H0**
Top 40	0.965	2.2E-16	Reject	Reject
JALSH	0.919	2.2E-16	Reject	Reject
MIDCAP	0.948	2.2E-16	Reject	Reject
JSMLC	0.190	2.2E-16	Reject	Reject

Table 5.11: Indices Shapiro-Wilk Test results.

6 Structural Breaks results

Sector	Share Code	First break	Second break	Third break	Fourth break	Fifth break
Basic Materials	AGL	14/07/2003	11/04/2007	03/09/2009	20/05/2013	
	AMS	28/05/2007	19/10/2009	25/03/2013		
	ANG	08/10/2003	16/07/2007	07/12/2009	29/05/2013	
	BIL	19/12/2002	20/03/2007	17/08/2009		
	IMP	14/07/2003	02/04/2007	28/08/2009	12/02/2013	
	SOL	12/06/2003	24/04/2007	16/09/2009		
Finacials	BGA	08/01/2003	22/05/2007	14/10/2009		
	DSY	11/05/2007	02/10/2009	20/02/2013		
	FSR	14/11/2002	19/12/2006	25/05/2009		
	GRT	07/06/2002	20/01/2006	20/05/2009	26/07/2012	
	INL	24/07/2003	20/06/2007	11/11/2009		
	ITU	06/08/2007	04/01/2010			
	NED	22/04/2004	22/05/2007	14/10/2009		
	OML	13/10/2003	08/06/2007	30/10/2009		
	RMH	14/11/2002	11/07/2005	18/01/2008	14/06/2010	
	SBK	20/12/2002	19/06/2007	10/11/2009		
	SLM	06/05/2003	31/01/2006	18/05/2009	13/05/2013	
Industrials	BVT	30/09/2002	07/02/2007	16/07/2009		
	IPL	29/10/2002	11/05/2006	02/12/2009		
Consumer Goods	CFR	17/07/2003	11/12/2007	10/06/2010	26/06/2013	
	SAB	09/06/2003	12/03/2007	17/08/2009		
	SNH	30/05/2007	21/10/2009			
	TBS	04/09/2002	05/10/2005	02/06/2009	18/01/2013	
Consumer Services	NPN	30/05/2002	11/08/2005	30/07/2009	06/05/2013	
	SHP	20/12/2002	30/04/2007	25/09/2009	14/08/2012	
	WHL	27/09/2002	03/05/2006	05/05/2009		
Health Care	APN	30/05/2002	18/11/2004	30/11/2007	15/06/2010	11/02/2013
	MDC	04/12/2007	13/07/2010	09/05/2013		
	NTC	06/09/2002	17/05/2007	14/05/2013	30/10/2009	
Telecommunications	MTN	29/08/2002	09/01/2007	04/06/2009		

Table 6.5: FTSE/JSE Africa Top 40 Index constituents structural breaks based on 04 January 2000 to 31 December 2015 absolute returns.

Sector	Share Code	First break	Second break	Third break	Fourth break	Fifth break
Industrials	BAW	30/05/2002	26/10/2004	17/05/2007	08/10/2009	
	GND	11/12/2007	20/05/2010	19/02/2013		
	KAP	31/05/2004				
	NPK	20/06/2002	02/01/2007	28/09/2009	27/05/2013	
	PPC	20/06/2002	10/01/2006	12/11/2009	24/04/2013	
	RLO	31/05/2002	27/06/2007	18/11/2009	25/07/2012	
	SPG	18/11/2002	03/03/2008	03/09/2010		
	TRE	18/11/2002	19/06/2007			
Basic Materials	ACL	23/08/2002	29/06/2005	18/01/2008	23/07/2010	01/08/2013
	AFE	03/06/2002	26/11/2007	08/07/2010	02/01/2013	
	ARI	29/11/2006	15/07/2009	28/03/2013		
	ASR	05/01/2005	18/02/2013			
	GFI	31/03/2003	13/09/2005	16/07/2009	28/03/2013	
	HAR	30/01/2003	15/05/2007	06/10/2009	02/04/2013	
	LON	03/06/2003	23/01/2006	24/06/2008		
	NHM	06/06/2002	03/01/2008	14/06/2010	02/04/2013	
	OMN	31/07/2008	14/03/2011			
	SAP	30/06/2003	07/06/2007	29/10/2009	28/05/2013	
Consumer Goods	AVI	02/09/2002	10/02/2005	30/07/2007	22/12/2009	28/06/2012
	DST	03/04/2009	10/11/2011			
	ILV	31/05/2002	29/03/2007	25/08/2009	10/07/2013	
	OCE	11/05/2007	03/06/2013			
	RCL	30/07/2002	24/03/2009			
	TON	30/05/2002	23/12/2009	26/03/2013		
Consumer Services	CLS	31/05/2002	20/02/2006	03/12/2009		
	FBR	08/12/2003	04/10/2011			
	ITE	25/10/2007	14/05/2010			
	PIK	30/05/2002	27/02/2006	01/07/2009	27/05/2013	
	SUI	02/09/2002	21/05/2007	14/10/2009	30/01/2013	
	TFG	12/08/2002	10/01/2005	24/08/2009	18/06/2012	
	TRU	07/08/2003	11/05/2006	09/10/2009	10/05/2013	
	TSH	02/02/2004	15/01/2008	26/10/2011		
Financials	BAT	09/01/2008	17/06/2010	16/05/2013		
	HYP	30/08/2004	01/06/2009	11/04/2013		
	LBH	16/07/2003	11/12/2007	18/06/2010	14/03/2013	
	PSG	18/10/2006	18/12/2009	17/05/2013		
	SAC	02/10/2002	26/07/2005	07/07/2009	25/10/2012	
	SNT	18/09/2002	31/07/2007	04/01/2010		
Technology	DTC	02/06/2003	11/01/2008	26/07/2010		

Table 6.6: FTSE/JSE Africa Mid Cap Index constituents structural breaks based on 04 January 2000 to 31 December 2015 absolute returns.

Sector	Share Code	First break	Second break	Third break	Fourth break	Fifth break
Health Care	ACT	27/02/2006	11/01/2010	15/08/2012		
Industrials	ADR	16/09/2002	15/01/2007	07/10/2011		
	AEG	23/07/2007	14/12/2009	03/06/2013		
	AEL	09/03/2006	03/09/2009	23/05/2013		
	AEN	01/08/2007	24/12/2009	14/03/2013		
	GRF	29/08/2002	02/06/2006	21/09/2009	13/08/2012	
	HDC	22/03/2005				
	MUR	11/06/2002	28/06/2007	07/04/2010	12/06/2013	
	WBO	30/05/2002	21/02/2007	13/09/2010		
Consumer Services	ADH	06/04/2004	07/04/2009	06/11/2012		
	CAT	28/07/2003	06/08/2007	19/03/2010		
	CLH	24/12/2007				
	CSB	26/02/2003	17/04/2009	08/05/2013		
Basic Materi als	AFX	05/08/2002	25/04/2005	10/01/2008	02/07/2010	23/04/2013
	MRF	31/05/2002	16/01/2008	10/06/2010	31/12/2012	
Financials	BRN	30/05/2002				
	OCT	19/12/2008	04/08/2011			
	PGR	03/06/2002	31/07/2009			
Consumer Goods	MTA	16/02/2007				

Table 6.7: FTSE/JSE Africa Small Cap Index constituents structural breaks based on 04 January 2000 to 31 December 2015 absolute returns.

Appendix II: Probability Theory

Definition 1. (Fries, 2007) (Probability Space, $\boldsymbol{\sigma}$ Algebra):

Let Ω denote a set and $\mathcal{F}$ denote a family of subsets of Ω. $\mathcal{F}$ is a $\boldsymbol{\sigma}$-algebra if

1. $\emptyset \in \mathcal{F}$.
2. $F \in \mathcal{F} \Longrightarrow \Omega \setminus F \in \mathcal{F}$.
3. $F_1, F_2, F_3, \ldots \in \mathcal{F} \Longrightarrow \bigcup_{i=1}^{\infty} F_i \in \mathcal{F}$.

The pair $(\Omega , \mathcal{F})$ is a measurable space. A function P: $\mathcal{F} \longrightarrow [0, \infty)$ is a probability measure if

1. $\mathrm{P}(\emptyset) = 0, \mathrm{P}(\Omega) = 1$.
2. For $F_1, F_2, F_3, \ldots \in \mathcal{F}$ mutually disjoint (i.e. $i \neq j \Longrightarrow F_i \cap F_j = \emptyset$) we have

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(F_i).$$

The triple $(\Omega, \mathcal{F}, \mathrm{P})$ is a called probability space, however, if

1. $\mathrm{P}(\emptyset) = 0$ only.
2. Same as pair.

P is called measure and $(\Omega, \mathcal{F}, \mathrm{P})$ is called measure space.
$\boldsymbol{\Omega}$-may be interpreted as the set of elementary events, only one event may occur.
$\boldsymbol{\sigma}$-algebra may be interpreted as a set of properties.

Definition 2. (Fries, 2007) (Independence, Conditional Probability)
Let $(\Omega, \mathcal{F}, \mathrm{P})$ denote a probability space and A, B $\in \mathcal{F}$.

1. We say that A and B are independent if
   $\mathrm{P}(\mathrm{A} \cap \mathrm{B}) = \mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B})$.
2. For $\mathrm{P}(\mathrm{B}) > 0$ we define the conditional probability of A under the hypothesis B as
   $\mathrm{P}\,(\mathrm{A}|\mathrm{B}) = \frac{\mathrm{P}(\mathrm{A} \cap \mathrm{B})}{\mathrm{P}(\mathrm{B})}$.

Definition 3. (Fries, 2007) (Filtration)
Let $(\Omega, \mathcal{F})$ denote a measurable space. A family of $\boldsymbol{\sigma}$-algebras $\{\mathcal{F}_t | t \geq 0\}$, where

$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \leq s \leq t$, is called a filtration on $(\Omega, \mathcal{F})$.

$\{\mathcal{F}_t | t \geq 0\}$; $\mathcal{F}_t$ is the information known at time t.

Definition 4. (Fries, 2007) (Martingales)
The stochastic process $\{x(t), \mathcal{F}_t; 0 \leq t < \infty\}$ is called a martingale w.r.t the filtration $\{\mathcal{F}\}$ and measure P if
$x_s = E(x(t)|\mathcal{F}_s)$ P-almost surely (a.s.), $\forall\, 0 \leq s \leq t < \infty$

if $x_s \leq E(x(t)|\mathcal{F}_s)$ the stochastic process is called a submartingale.
if $x_s \geq E(x(t)|\mathcal{F}_s)$ the stochastic process is called a supermartingale.
Theorem 1. (Rice, 2007) (Central Limit Theorem)
Let $X_1, X_2, \ldots X_n$ be a time series of independent random variables with mean μ and variance σ^2.

Let each X_i have the *cdf* $\mathrm{P}(X_i \leq x) = F(x)$ and the moment generating function $M(t) = E(e^{tX_i})$.
$S_n = \sum_{i=1}^{n} X_i$.

Then $\lim_{n\to\infty} \mathrm{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}\right) = \Phi(x)$, for $-\infty < x < \infty$.

Proof: If $\mu = 0$
Let $Y_1 + Y_2 + \ldots + Y_n = T_n$, then we have
$\lim_{n\to\infty} \mathrm{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \mathrm{P}\left(\frac{T_n}{\sigma\sqrt{n}} \leq x\right)$.

Let $Z_n = \frac{S_n}{\sigma\sqrt{n}}$
Since S_n is a sum of independent random variables
$M_{S_n}(t) = [M(t)]^n$.

Proposition: If X is a random variable with *mgf* M_x and $Y = a + bX$, then $M_y(t) = e^{at}M_x(bt)$.
Therefore, $M_{Z_n}(t) = \left[M\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$
$\lim_{n\to\infty} n log M\left(\frac{t}{\sigma\sqrt{n}}\right) = \frac{t^2}{2}$.

Appendix III: Statistical Moments and Tests Derivation

1. Statistical moments

Let ι^{th} be the ι moment of a time series $\{ x_t \}$

$$m_l = E(x^l) = \int_{-\infty}^{\infty} x^l f(x)dx.$$

$f(x)$ is the probability density function of x_t

$$m_1 = E(x - \bar{x})^1 = \int_{-\infty}^{\infty} (x - \bar{x})^1 f(x)dx,$$

$$m_2 = E(x - \bar{x})^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)dx,$$

$$m_n = E(x - \bar{x})^n = \int_{-\infty}^{\infty} (x - \bar{x})^n f(x)dx.$$

$\bar{x}$ is the mean of x_t

2. Log-normal returns

Log-normal returns of a risky asset with mean μ and variance σ^2 are given by:

$E(x_t) = e^{(\mu + \frac{\sigma^2}{2})} - 1.$

$Var(x_t) = e^{(2\mu + \sigma^2)e^{\sigma^2} - 1}.$

3. Difference-stationarity

Let's assume a Random Walk process:

$$x_t = \alpha + \beta_1 t + \beta_2 x_{t-1} + \varepsilon_t.$$

where

$\varepsilon_t \sim IID(0, \sigma^2)$ is white noise.

If $\alpha = 0, \beta_1 = 0$ and $\beta_1 = 1$ then $x_t = x_{t-1} + \varepsilon_t$ (Pure Random Walk).

$$\Delta x_t = \varepsilon_t.$$

$$\text{E}(\Delta x_t) = E(\varepsilon_t),$$
$$= 0.$$

$$\text{Var}(\Delta x_t) = Var\ (\varepsilon_t),$$
$$= \sigma^2.$$

Therefore Random Walk without a drift is difference-stationary, since its mean and variance are constant (time-invariant).

If $\alpha \neq 0, \beta_1 = 0$ and $\beta_2 = 1$ then $x_t = \alpha + x_{t-1} + \varepsilon_t$ (Random Walk with a drift).

$$\Delta x_t = \alpha + \varepsilon_t.$$

$$\text{E}(\Delta x_t) = E(\alpha + \varepsilon_t),$$
$$= E\ (\alpha) + E(\varepsilon_t),$$
$$= \alpha + 0,$$
$$= \alpha.$$

$$\text{Var}(\Delta x_t) = Var\ (\alpha + \varepsilon_t),$$
$$= Var(\alpha\ ) + Var(\varepsilon_t),$$
$$= \alpha + \sigma^2.$$

Therefore Random Walk with a drift is difference-stationary, since its mean and variance are constant (time-invariant).

If $\alpha \neq 0, \beta_1 \neq 0$ and $\beta_2 = 0$ then $x_t = \alpha + \beta_1 t + \varepsilon_t$ (Random Walk with a deterministic trend).

$$\Delta x_t = x_{t-1} + \alpha + \beta_1 t + \varepsilon_t.$$

$$\text{E}(\Delta x_t) = E\ (x_{t-1} + \alpha + \beta_1 t + \varepsilon_t),$$
$$= E\ (x_{t-1}) + E\ (\alpha) + E\ (\beta_1 t) + E(\varepsilon_t),$$

$$= E\ (x_{t-1}) + \alpha + \ \beta_1 t + 0,$$
$$= E\ (x_{t-1}) + \alpha + \ \beta_1 t.$$

$$\text{Var}(\Delta x_t) = Var\ (x_{t-1} + \alpha + \ \beta_1 t + \varepsilon_t),$$
$$= Var(\ x_{t-1}) + Var(\alpha\ ) + \ Var\ (\beta_1 t) + \ Var(\varepsilon_t),$$
$$= Var(\ x_{t-1}) + \ \alpha + \beta_1 t + \ \sigma^2.$$

Therefore Random Walk with a deterministic trend is difference-non-stationary, since its mean and variance are time varying.

If $\alpha \neq 0, \beta_1 \neq 0$ and $\beta_2 = 1$ then $x_t = \alpha + \ \beta_1 t + x_{t-1} + \ \varepsilon_t$ (Random Walk with a drift and deterministic trend).

$$\Delta x_t = \ \alpha + \ \beta_1 t + \ \varepsilon_t.$$

$$\text{E}(\Delta x_t) = E\ (\alpha + \ \beta_1 t + \ \varepsilon_t\ ),$$
$$= E\ (\alpha) + \ E(\ \beta_1 t) + \ E(\varepsilon_t),$$
$$= \alpha + \ \beta_1 t + 0,$$
$$= \alpha + \ \beta_1 t.$$

$$\text{Var}(\Delta x_t) = Var\ (\alpha + \ \beta_1 t + \ \varepsilon_t),$$
$$= Var(\alpha\ ) + \ Var(\ \beta_1 t) + \ Var(\varepsilon_t),$$
$$= \ \alpha + \beta_1 t + \ \sigma^2.$$

Therefore Random Walk with a drift and deterministic trend is difference-non-stationary, since its mean and variance are time varying.

4. Stationarity Test Derivation

Definition 4.1. (Fries, 2007) A process $\{x_t\}$ is called stationary if for every n, the process $\{x_t\}_{t=n}^{\infty}$ has the same distribution as $\{x_t\}_{t=1}^{\infty}$, that is

P $\{(x_1, x_2, \ldots) \in \ \mathcal{B}^\infty\}$ = P $\{(x_{n+1}, x_{n+2}, \ldots) \in \ \mathcal{B}^\infty\}$

where

P is the probability measure.

$\mathcal{B}$ is the Borel σ −algebra.

Definition 4.2. (Fries, 2007) A stationary process is ergodic if every invariant event has

P (A) = 0 or 1.

where

An event A is called invariant.

Theorem .4.1. (Fries, 2007)

If $x_1, x_2, \ldots$, is a stationary and ergodic process and $\mathrm{E}|x_1| < \infty$, then

$$\frac{1}{T}\sum_{t=1}^{T} x_t \longrightarrow \mathrm{E}\, x_1 \ a.s.$$

Phillips-Perron Test

Assume the regression model

$x_t = \alpha + \rho x_{t-1} + \varepsilon_t; \qquad \varepsilon_t \sim IID(0, \sigma^2).$

$H_0: I(0)$ against covariance-stationary process

$H_1: I(1)$.

Phillips-Perron test estimates

$$x_t = \alpha + \rho x_{t-1} + \varepsilon_t \ by\ OLS \qquad \text{(I)}$$

Under the assumption

$$\alpha = 0, \qquad \rho = 1\ and\ \varepsilon_t \sim IID \quad \mathbb{N}(0, \sigma^2)$$

OLS formulae are used to compute $\hat{\rho}$ and the standard error of regressions $\hat{\sigma}_{\hat{\rho}}$.

Let's consider two least-squares regression equations

$$x_t = \hat{\varepsilon} + \hat{\rho}\, x_{t-1} + \hat{\varepsilon}_t \qquad \text{(II)}$$

$$x_t = \tilde{\varepsilon} + \tilde{\beta}\left(t - \frac{1}{2}T\right) + \tilde{\rho} x_{t-1} + \hat{\varepsilon}_t \qquad \text{(III)}$$

where

$(\hat{\varepsilon}, \hat{\rho})$ and $(\tilde{\varepsilon}, \tilde{\rho}, \tilde{\beta})$ are conventional least−squares regression coefficients.

Regression to statistics is given by:

$$t_{\hat{\rho}} = (\hat{\rho} - \rho)\frac{\sqrt{\sum(x_{t-1} - \bar{x}_{-1})^2}}{\hat{s}},$$

$$t_{\hat{\varepsilon}} = (\hat{\varepsilon} - \varepsilon)\frac{\sqrt{\frac{\sum(x_{t-1} - \bar{x}_{-1})^2}{\sum x^2{}_{t-1}}}}{\hat{s}},$$

$$t_{\tilde{\varepsilon}} = \frac{\tilde{\varepsilon} - \varepsilon}{\sqrt{\tilde{S}^2 c_1}},$$

$$t_{\tilde{\beta}} = \frac{\tilde{\beta} - \beta}{\sqrt{\tilde{S}^2 c_2}},$$

$$t_{\tilde{\rho}} = \frac{\tilde{\rho} - \rho}{\sqrt{\tilde{S}^2 c_3}}.$$

where

$\hat{S}$ and $\tilde{S}$ are the standard errors of regressions (II) and (III).

c_i is the i[th] diagonal element of the matrix $(X'X)^{-1}$ where X is the T x 3 matrix.

$\bar{x}_{-1} = \frac{\sum x_{t-1}}{T}$.

So under equation (II)

$$x_t = S_t + x_0,$$

Thus as T → ∞

$$T^{\frac{-3}{2}}\sum x_t \rightarrow \sigma\int W,$$

$$T^{-2}\sum x^2{}_t \rightarrow \sigma^2\int W^2,$$

$$T^{\frac{-5}{2}}\sum t x_{t-1} \rightarrow \sigma\int rW \ ,$$

$$T^{-1}\sum x_{t-1}\,\varepsilon_t \rightarrow \sigma^2\int WdW + \lambda.$$

5. Independence Tests Derivation

Dickey-Fuller Test

Consider the following regression models for financial time series x_t

$$x_t = \theta_1 x_{t-1} + \varepsilon_t,$$

$$x_t = \theta_0 + \theta_1 x_{t-1} + \varepsilon_t.$$

Hypothesis test

$H_0: \theta_1 = 1.$

$H_1: \theta_1 < 1.$

The Dickey-Fuller test statistic is given by

$$DF = \frac{\theta_1 - 1}{SE(\theta_1)} = \frac{\sum_{t=1}^{N} x_{t-1}\varepsilon_t}{\sigma_\varepsilon \sqrt{\sum_{t=1}^{N} x^2{}_{t-1}}}.$$

Ljung-Box Test

Box and Ljung (1978)

Consider a stationary autoregressive-moving average model

$\phi(B) w_t = \theta(B)\varepsilon_t$

where

$\{w_t\}$ is a discrete time series generated by a stationary autoregressive-moving average model.

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_q B^p,$$

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q.$$

$$B^k w_t = w_{t-k}.$$

$\{\varepsilon_t\}$ is white noise,

$$\varepsilon_t \sim IID \quad \mathbb{N}(0, \sigma^2).$$

The autocorrelation of white noise (ε_t)

$$\rho_i = \frac{\sum_{t=i+1}^{n} \varepsilon_t \varepsilon_{t-i}}{\sum_{t=1}^{n} {\varepsilon_t}^2}. \quad (i = 1,2, \dots)$$

For residuals $\varepsilon_1, \dots, \varepsilon_n$

Therefore Ljung−Box(Q) test statistic is given by

$$\tilde{Q}(\rho) = n(n+2)\sum_{i=1}^{\iota} \frac{{\rho_i}^2}{n-i},$$

where

$$\rho_k = \frac{\sum_{t=i+1}^{n} \varepsilon_t \varepsilon_{t-i}}{\sum_{t=1}^{n} {\varepsilon_t}^2}.$$

$n \sim {\chi_\iota}^2.$

$$E\{\, Q(\rho)\} = n\sum_{i=1}^{\iota} E({\rho_i}^2) = \frac{\iota n}{n+2}\left(1 - \frac{\iota+1}{2n}\right).$$

$$Var\{\, Q(\rho)\} = n^2\sum_{i=1}^{\iota} Var({\rho_i}^2) + 2n^2 \sum_{i=1}^{\iota-1}\sum_{j=i+1}^{\iota} \frac{cov({\rho_i}^2{\rho_j}^2)}{(n-i)(n-j)}.$$

where

For fixed $n, cov({\rho_i}^2{\rho_j}^2)$ is nonzero.

6. Normality Test Derivation

Shapiro-Wilk (W) Test

Shapiro and Wilk (1965)

Let $m' = [m_1 \dots m_n]$ denote the vector of the expected values of standard normal order statistics

Let $V = (v_{ij})$ a $n\ x\ n$ Covariance matrix then

$E(x_i) = m_i \; (i = 1,2, \dots, n).$

$Cov(x_i, x_j) = v_{ij} \; (i, j = 1,2, \dots, n).$

For $x_1 \leq x_2 \leq \cdots x_n \sim \mathbb{N}\,(0,1)$

Let

$y' = [y_1 \dots y_n]$ denote a vector of ordered random observations, assume that $\{y_i\}$ are a normal sample

$y_i = \mu + \sigma x_i \ (i = 1,2, \dots n)$

$$\hat{\mu} = \frac{m'v^{-1}(m1' - 1m')v^{-1}y}{1'v^{-1}1m'v^{-1}m - (1'v^{-1}m)^2}.$$

$$\hat{\sigma} = \frac{1'v^{-1}(1m' - m1')v^{-1}y}{1'v^{-1}1m'v^{-1}m - (1'v^{-1}m)^2}.$$

$1'v^{-1}m = 0$ for symmetric distributions

Therefore $\hat{\mu} = \frac{1}{n}\sum_{i=0}^{n} y_i = \bar{y}$ and $\hat{\sigma} = \frac{m'v^{-1}y}{m'v^{-1}m}$.

W test statistic for normality is defined by

$W = \frac{R^4\hat{\sigma}^2}{c^2 s^2} = \frac{(a'y)^2}{s^2} = \frac{(\sum_{i=1}^{n} a_i y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

where

s^2 is the sample variance

$R^2 = m'v^{-1}m$

$c^2 = m'v^{-1}v^{-1}m$

$a' = [a_1 \dots a_n] = \frac{m'v^{-1}}{\sqrt{m'v^{-1}v^{-1}m}}$

7. Structural Breaks Test Derivation

Bai-Perron Test

Assume a simple AR(1) with mean and variance breaks

$$y_t = \beta y_t - 1 + \mu_t + \varepsilon_t,$$

$$\varepsilon_t \sim \mathbb{N}\ (0, \sigma_t^2).$$

$$E(y_t) = \frac{\mu_t}{1 - \beta}.$$

$$\text{Var}(y_t) = \frac{\sigma_t{}^2}{1-\beta^2}.$$

The full structural breaks model is given by:

$$y_t = \beta_1{}'\mathbb{X}_t + \varepsilon_t, \qquad t \leq T_1$$

$$y_t = \beta_2{}'\mathbb{X}_t + \varepsilon_t, \qquad t > T_1$$

$$\Rightarrow y_t = \beta_1{}'\mathbb{X}_t\mathbf{1}(t \leq T_1) + \beta_2{}'\mathbb{X}_t\mathbf{1}(t > T_1) + \varepsilon_t .$$

The partial structural model is given by:

$$y_t = \beta_0{}'\mathbb{Z}_t + \beta_1{}'\mathbb{X}_t\mathbf{1}(t \leq T_1) + \beta_2{}'\mathbb{X}_t\mathbf{1}(t > T_1) + \varepsilon_t$$

where T_1 is the break date. The partial structural break model has been used for the purpose of this paper.

The sum of squared error (SSE) is given by

$$S(\beta_1\ T_1) = \frac{1}{n}\sum_{t=1}^{n}(y - \beta_0{}'\mathbb{Z}_t - \beta_1{}'\mathbb{X}_t\mathbf{1}(t \leq T_1) - \beta_2{}'\mathbb{X}_t\mathbf{1}(t > T_1))^2$$

$$\left(\widehat{\beta_1}\ \widehat{T_1}\right) = argmin S(\beta_1\ T_1)$$

$$= argmin_{T_1}\, min_{\beta}\, S(\beta_1\ T_1)$$

$$= argmin_{T_1}\ S(T_1).$$

where:

$$S(T_1) = min_{\beta}\, S(\beta_1\ T_1) = \frac{1}{n}\sum_{t=1}^{n}\widehat{\varepsilon_t}\,(T_1)^2.$$

$\widehat{\varepsilon}_t\,(T_1)$ are the OLS residuals from

$$y_t = \hat{\beta}_0{}'\mathbb{Z}_t + \hat{\beta}_1{}'\mathbb{X}_t\mathbf{1}(t \leq T_1) + \hat{\beta}_2{}'\mathbb{X}_t\mathbf{1}(t > T_1) + \hat{\varepsilon}_t(T_1)$$

and T_1 is fixed

$$\widehat{T_1} = argmin_{T_1} S(T_1)$$

Distribution of break date estimator, Bai (1997)

$$\frac{(\delta' Q\delta)^2}{\delta'\Omega\delta}\left(\hat{T}_1 - T_1\right) \to d\zeta = \ argmax_s \left[W(S) - \frac{|S|}{2}\right].$$

If

$\delta \to 0$

where

W(S) is a double-sided Brownian motion.

$Q = E(\mathbb{X}_t \mathbb{X}'_t\ ).$

$\Omega = E(\mathbb{X}_t \mathbb{X}'_t \varepsilon_t{}^2\ ).$

$\zeta = \beta_2 - \beta_1.$

The distribution of ζ for x ≥ 0 is

$$G(x) = \ 1 + \ \sqrt{\frac{x}{2\pi}}\exp\left(-\frac{x}{8}\right) - \ \frac{x+5}{2}\Phi\left(-\frac{\sqrt{x}}{2}\right) + \frac{3e^x}{2}\Phi\left(-\frac{3\sqrt{x}}{2}\right)$$

and

$$G(x) = \ 1 - G(-x).$$

If $\quad \varepsilon_t \sim \ IID$ then

$$\Omega_1 \ = Q_1\,\sigma_1{}^2 \quad \text{and}$$

$$\frac{\delta' Q_1 \delta}{\sigma_1{}^2}\left(\,\hat{T}_1 - T_1\,\right) \to d\zeta.$$

Break date confidence interval Theorem

$\hat{T}_1 \sim T_1 + \frac{\delta'\Omega\delta}{(\delta' Q\delta)^2}\zeta$ then

$\hat{T}_1 \pm \frac{\hat{\delta}'\hat{\Omega}\hat{\delta}}{(\hat{\delta}'\hat{Q}\hat{\delta})^2}c$

where

c is the critical value of ζ.

$\hat{\delta} = \hat{\beta}_2 - \hat{\beta}_1$.

$\hat{Q} = \frac{1}{n} \sum_{t=1}^{n} \mathbb{X}_t \mathbb{X}'_t$.

$\widehat{\Omega} = \frac{1}{n-k} \sum_{t=1}^{n} \mathbb{X}_t \mathbb{X}'_t \hat{\varepsilon}_t{}^2 + \frac{1}{n-k} \sum_{j=0}^{h-1} 1 \sum_{t=1}^{T_1-j} \mathbb{X}_t \mathbb{X}'_{t+j} \hat{\varepsilon}_t \hat{\varepsilon}_{t+j} + \mathbb{X}_{t+j} \mathbb{X}'_t \hat{\varepsilon}_{t+j} \hat{\varepsilon}_t$.

$c \sim \frac{1}{G(x)}$.

Confidence interval	c
80%	4.7
90%	7.7
95%	11

Appendix IV: Structural Breaks Dates vs. Company News

Dates highlighted in red are not the exact dates of the structural break. They are a few days before or after the break date. There are also instances where I didn't highlight the date in red, even though it doesn't match the break date. This is where there is a two business days difference.

Section A: Allows heterogeneity and autocorrelation in the residuals

Top 40 Stocks

APN

31/05/2002 – Appointment of a non-executive.

18/11/2004 – Sale of shares by a director (dealing in securities).

04/02/2013 – Announcement of a potential acquisition.

IMP

02/04/2007 – Dealing in securities by a director.

28/08/2009 – release of financial statements (27/08/2009).

12/02/2013 –Trading statement on financial statements release.

SBK

10/11/2009 – Disclosure of Basel II capital adequacy.

NED

21/05/2007 – Dealing in securities by directors (options exercise).

14/10/2009 – Announcement of acquisition (15/10/2009).

FSR

13/11/2002 – Company's subsidiary expansion plans.

19/12/2006 – Dealing in securities by a director (purchase).

22/05/2009 – Reserve bank announces plans to probe banks' lending rates.

DSY

11/05/2007 – Dealing in securities by BEE partner.

20/02/2013 – Unaudited interim results announcement and cash dividend declarations for the six months ended 31 December 2012 (21/02/2013).

BGA

08/01/2003 – Take-over announcement.

23/05/2007 – Purchases 10% of 'B' Shares of an investment company.

25/05/2007 – A private equity consortium led by Absa Capital buys a stake in another company.

SLM

31/01/2006 – SLM sells stake in its fund to BEE fund.

26/01/2009 – Credit ratings agency changes company outlook to negative.

17/05/2013 – Concludes acquisition of another company.

INL

24/07/2003 – Offers to buy assets in another company.

20/06/2007 – Disclosure on dealing in securities.

11/11/2009 – Disclosure on dealing in securities.

GRT

10/01/2006 – Resignation of a director.

18/05/2009 – Buys control of international company.

ITU

06/08/2007 – Buys its own shares.

OML

14/10/2003 – Appointment of a CEO to a subsidiary (Nedcor).

06/08/2007 – Announcement regarding an investment project.

30/10/2009 – Timetable for proposed acquisition of another company (Mutual & Federal).

AGL

14/07/1003 – News about expected demand in chrome and a rise in prices.

11/04/2007 – Announcement about transaction in own shares (buyback).

03/09/2009 – Dealing in securities by persons with interest in securities representing 1% or more.

20/05/2013 – Announcement of an investment project.

AMS

28/05/2007 – Company declines to increase pay for workers.

15/10/2009 – Another company (Xstrata) dropped its bid for a nil premium merger. Formal bid was expected on 20/10/2009.

12/03/2013 – Wildcat strikes.

SOL

12/06/2003 – Refutes claims made in parliament that one of its refineries received a subsidy from government.

17/04/2007 – Announcement about the development of technology to boost output.

14/09/2009 – Release of financial results.

ANG

07/10/2003 – Potential bid by another company to ANG's Ghana subsidiary.

12/07/2009 – Cuts non-essential workforce in Guinea operations.

29/05/2013 – Potential job cuts announcement.

BIL

19/12/2002 – Dealing in securities.

20/03/2007 – Credit agency affirms company's outlook as stable.

IPL

22/10/2002 – Announcement of offer price for odd-lot.

12/05/2006 – Changes in executive management.

01/12/2009 – Update in trading environment.

BVT

05/09/2002 – Buys another company.

CFR

02/06/2010 – Acquires shares in another company.

26/06/2013 – International bank recommends stock as a long-term portfolio investment.

SNH

29/05/2007 – Merger discussion termination.

21/10/2009 – Dealings in securities by a director (sale).

TBS

03/09/2009 – Circular regarding odd-lot offer and amendment to the articles of association.

29/092005 – Sanctioning of the scheme (employee incentive plan) arrangement by the High Court.

28/05/2009 – Fulfillment of conditions precedent and disposal of another company's stake holdings.

NPN

04/08/2005 – No change statement and notice of annual general meeting.

30/07/2009 – No change statement and notice of annual general meeting.

SHP

04/12/2002 – Another company sells its stores to SHP.

16/09/2007 – Dealing in securities by a director (purchase).

12/08/2012 – Exposé of the chairman's tax liabilities.

WHL

29/04/2009 – Dealing in securities by a director (sale).

MDC

02/07/2010 – Rights offer finalization announcement.

07/05/2013 – Economic Development ministry announces that the Competition Commission will investigate private medical-care industry.

NTC

29/08/2002 – Dealing in securities by directors (purchase and sale).

14/04/2007 – First-half profits drop.

14/05/2013 – Disclosure of acquisition of a beneficial interest in preference shares an asset manager.

MTN

29/08/2002 – Share fell on fears that company network in Nigeria would suffer if the country defaults on its foreign debt.

04/06/2009 – Indian government seeks details about a proposed merger of MTN and Indian telecommunications company.

Midcap Stocks

GND

05/12/2007 and 13/12/2007 – Dealing in securities by directors (exercise of options and sale).
20/05/2010 – Results of annual general meeting and retirement of director.

BAW
27/05/2002 – Dealing in securities by director (Options).
17/05/2007 – Unbundling of another company shares and dealing in securities by director (options and purchase).
01/10/2009 – Change to board of directors.

NPK
20/06/2002 – Declaration of cumulative of preference shares dividends.
28/05/2013 – Interim results release.

PPC

06/01/2006 – Share falls on expenditure plan on new equipment.
12/11/2009 – Equity research company downgrade PPC Free Cash Flow Valuation.

RLO

19/06/2007 – Dealing in securities – zero cost collar option trigger.
18/11/2009 – Release of audited results.
25/07/2012 – Dealing in securities by director (option and sale).

SPG

12/11/2002 – Share price falls after reporting first-half profit decline.
30/03/2008 – Release of interim financial results.

TRE

20/11/2002 – Change in directorate.

ACL

23/08/2002 – Share jumps on a potential takeover bid by another company.
29/06/2005 – Announcement of steel prices cuts by company.
18/01/2008 – Announcement of steel prices increase by company.
23/07/2010 – Further cautionary announcement about a proposed BEE ownership transaction.
01/08/2013 – Company seeks iron-ore supply from its own property.

AFE

28/11/2007 – Share falls on profit decline forecast.
08/07/2010 – Trading statement, expects HEPS higher than prior year.

ARI

28/11/2006 – Company's joint venture receives mining rights.
10/07/2009 – Appointment of company secretary.
22/03/2013 – Disclosure of beneficial interests in securities.

ASR

13/02/2013 – HEPS decline by 47%.

GFI

16/07/2009 – Safety improvements at its one of its shafts.
28/03/2013 – CEO pay jumps as share declines news.

HAR

30/01/2003 – Sells stake in Canadian rival, more than three times what it originally paid.
14/05/2007 – A company buying its shafts posts a loss.

LON

03/06/2003 – Notification from an asset manager that it no longer hold notifiable interest in the company shares.
23/01/2006 – Notification from a bank that it no longer holds notifiable interest in the company's ordinary shares.

NHM

11/06/2010 – Company resumes convert operations at one of its mines.

OMN

31/07/2008 – Dealing in securities by director (purchase).
02/03/2011 – Three employees killed by explosion at one of its Plants.

SAP

01/06/2007 – Dealing in securities by directors of listed company (sale).
30/10/2009 – Announces intentions to close pulp mill and restructuring of forestry business in Swaziland.
24/05/2009 – Credit ratings agency changes rating from positive to stable.
29/05/2013 – Company may close more mills in Europe, said the CEO.

AVI

27/06/2012 – Acceptance of share options by company directors through the executive incentive scheme.

ILV

21/05/2002 – Profits fall on costs of Mozambique operations write-off.
02/04/2007 – Potential expansion in Zambia.
24/08/2009 – Posting of rights issue offer circular.

OCE

11/05/2007 – First-half profit climbs 34%.
04/06/2013 – Voluntary announcement about acquisition of another company and withdrawal of cautionary announcement.

RCL

01/06/2002 – Expects costs to increase if currency continues to weaken.

21/11/2009 – Changes to the board of directors.

TON
20/05/2002 – Zimbabwe police arrest strikers threatening to burn company's sugar cane.
10/12/2009 – Company secretary change.
19/03/2013 – Raises R117m in accelerated bookbuild offer.

CLS
07/12/2009 – Dealing in securities by executive directors (purchase on open market)

ITE
17/05/2010 – Changes to the board of directors.

PIK
28/05/2002 – Company plans to buyback R134m shares.
29/06/2009 – Share falls on antitrust probe.
24/05/2013 – Dealing in securities by directors (off market sale and purchases).

SUI
09/10/2009 – Dealing in securities by director (sale).
29/01/2013 – Changes to the board of directors.

TFG
19/06/2012 – Dealing in securities by directors (sale).

TRU
05/05/2006 – Dealing in securities by directors (sale).
03/05/2013 – Dealing in securities by directors (options + sale).

BAT
09/06/2010 – Dealing in securities by non-executive director (purchase).

HYP
27/05/2009 – Changes to the board of directors.

LBH
14/03/2013 – Dealing in securities by CEO and company secretary (sale).

03/12/2007 – Dealing in securities by company secretary (sale).

PSG
18/10/2006 - Dealing in securities by directors (sale + purchase).
16/05/2013 – Dealing in securities by a director of a major subsidiary company (sale on market).
07/12/2009 – Makes an offer of up to 49% to buy another company.

SAC
30/06/2005 – Dealing in securities by a director (option exercise).
24/10/2012 – Amended of exclusivity clauses for SAC transaction by another company, recommended by the Competition Commission.

DTC
15/01/2008 – Company sees strong second-half to meet full year earnings estimate.
26/07/2010 – International bank assigns "Overweight" rating to DTC.

Smallcap Stocks

ADR
11/09/2002 – Changes to the board of directors.
12/01/2007 – Sells research surveys to another company.
05/10/2011 – Release of voluntary trading statement.

AEG
02/07/2007 – Profits double.
03/06/2013 – Announcement of the settlement agreement with the Competition Commission regarding anti-competitive practices.

AEL
24/02/2006 – Cautionary announcement.
17/05/2013 – Makes an offer to buy another company.

AEN
31/07/2007 – Dealing in securities by CEO and CFO (Options exercised on open market).
21/12/2009 – Dealing in securities by a director (sale on open market).

GRF

23/08/2002 – Annual profits jump 64%.
29/08/2002 – Government expected to increase spending on infrastructure.
29/05/2006 – Renewal of cautionary trading.
01/09/2009 – Company faces anti-trust investigation.
13/08/2012 – Announces plans to expand into west and east Africa.

HDC
23/03/2005 – Acquires another company's businesses.

WBO
21/02/2007 – SA government rejects calls for more funding to build stadiums.
10/09/2010 – International bank recommends stock as a buy from hold.

ADH
30/03/2004 – Dealing in securities by a director (sale).

CAT
23/07/2007 – Asks South Africa's Competition Tribunal to ban a possible merger with a bigger company.
19/03/2010 – Dealing in securities by a director (sale).

CLH
21/12/2007 – Announcement about a proposed BEE transaction.

CSB
12/02/2003 – First-half earnings release.
09/05/2013 – Third quarter operations update announcement.

AFX
29/07/2002 – Results of capitalization award and right of election to receive cash dividends.
21/04/2005 – CEO accepts invitation for his contract to be extended for another year.
29/04/2005 – Company releases solid first-half results.
07/01/2008 – Changes to the board of directors.
11 – 12/04 / 2013 – Dealing in securities by directors (sale + purchases, off market).
25/04/2013 – Changes to the board of directors.

MRF
18/01/2008 – Release of trading statement.
03/06/2010 – Announcement about ten furnaces to undergo routine maintenance.

19/12/2012 – Announcement about price increase in ferrochrome.

BRN
20/05/2002 – Announcement regarding price manipulation probe in the company's share by the Financial Services Board.

OCT
04/08/2011 – Dealing in securities by an associate of a director (purchase on open market).

PGR
04/06/2002 – Cautionary announcement and increased shareholding in another company.
31/07/2009 – Dealing in securities by a director of a major subsidiary (sale).

MTA
16/02/2007 – Revision of salient dates for subdivision of share capital.

Section B: Disallows heterogeneity and autocorrelation in the residuals.

These are additional breaks when I disallowed heterogeneity and autocorrelation in the residuals.

Top 40 Stocks

SOL
20/05/2013 – Dealing in securities by an associate of a director of a major subsidiary of the company (sale on open market).

ITU
31/05/2012 – Announcement about voting rights.

IPL
08/05/2013 – Dealing in securities by a director.

MDC
02/02/2004 – Cautionary announcement about a potential acquisition.

Midcap Stocks

KAP
11/12/2009 – Dealing in securities by a director (sale).

RCL
25/01/2013 – Finalization announcement regarding an increase in the number of authorized shares and right.

ITE
25/07/2013 – Announcement of the new CEO.

SNT
01/06/2012 – Dealing in securities by an executive director (sale on open market).

Smallcap Stocks

AEG
12/03/2003 – Plans for expansion.

AEN
28/01/2005 – Cautionary announcement.

HDC
14/07/2009 – Dealing in securities by directors and the company secretary.
30/07/2013 – Dealing in securities by directors and the company secretary.

WBO
07/06/2013 – Business update.

BRN
23/02/2009 – Results review.

MTA
14/07/2009 – Trading statement release.

Source: Company websites and Bloomberg.

Appendix V: Source code

RStudio code

Phillips-Perron Test

```
PP.test(Top.40$AGL)
PP.test(Top.40$AMS)
PP.test(Top.40$ANG)
PP.test(Top.40$APN)
PP.test(Top.40$BGA)
PP.test(Top.40$BIL)
PP.test(Top.40$BVT)
PP.test(Top.40$CFR)
PP.test(Top.40$DSY)
PP.test(Top.40$FSR)
PP.test(Top.40$GRT)
PP.test(Top.40$IMP)
PP.test(Top.40$INL)
PP.test(Top.40$IPL)
PP.test(Top.40$ITU)
PP.test(Top.40$MDC)
PP.test(Top.40$MRP)
PP.test(Top.40$MTN)
PP.test(Top.40$NED)
PP.test(Top.40$NPN)
PP.test(Top.40$NTC)
PP.test(Top.40$OML)
PP.test(Top.40$RMH)
PP.test(Top.40$SAB)
PP.test(Top.40$SBK)
PP.test(Top.40$SHP)
PP.test(Top.40$SLM)
PP.test(Top.40$SNH)
PP.test(Top.40$SOL)
PP.test(Top.40$TBS)
PP.test(Top.40$WHL)
PP.test(Midcap$ACL)
PP.test(Midcap$AFE)
PP.test(Midcap$ARI)
PP.test(Midcap$ASR)
PP.test(Midcap$AVI)
PP.test(Midcap$BAT)
PP.test(Midcap$BAW)
PP.test(Midcap$CLS)
```

```
PP.test(Midcap$DST)
PP.test(Midcap$DTC)
PP.test(Midcap$EOH)
PP.test(Midcap$FBR)
PP.test(Midcap$GFI)
PP.test(Midcap$GND)
PP.test(Midcap$HAR)
PP.test(Midcap$HYP)
PP.test(Midcap$ILV)
PP.test(Midcap$ITE)
PP.test(Midcap$KAP)
PP.test(Midcap$LBH)
PP.test(Midcap$LON)
PP.test(Midcap$NHM)
PP.test(Midcap$NPK)
PP.test(Midcap$OCE)
PP.test(Midcap$OMN)
PP.test(Midcap$PIK)
PP.test(Midcap$PPC)
PP.test(Midcap$PSG)
PP.test(Midcap$RCL)
PP.test(Midcap$RLO)
PP.test(Midcap$SAC)
PP.test(Midcap$SAP)
PP.test(Midcap$SNT)
PP.test(Midcap$SPG)
PP.test(Midcap$SUI)
PP.test(Midcap$TFG)
PP.test(Midcap$TON)
PP.test(Midcap$TRE)
PP.test(Midcap$TRU)
PP.test(Midcap$TSH)
PP.test(Smallcap$ACT)
PP.test(Smallcap$ADH)
PP.test(Smallcap$ADR)
PP.test(Smallcap$AEG)
PP.test(Smallcap$AEL)
PP.test(Smallcap$AEN)
PP.test(Smallcap$AFX)
PP.test(Smallcap$BRN)
PP.test(Smallcap$CAT)
PP.test(Smallcap$CLH)
PP.test(Smallcap$CSB)
PP.test(Smallcap$GRF)
PP.test(Smallcap$HDC)
PP.test(Smallcap$MRF)
```

```
PP.test(Smallcap$MTA)
PP.test(Smallcap$MUR)
PP.test(Smallcap$OCT)
PP.test(Smallcap$PGR)
PP.test(Smallcap$WBO)
PP.test(Index$Top40Index)
PP.test(Index$MidcapIndex)
PP.test(Index$SmallcapIndex)
```

Ljung-Box Test

```
Box.test(Top.40$AGL, type = c("Ljung-Box"))
Box.test(Top.40$AMS, type = c("Ljung-Box"))
Box.test(Top.40$ANG, type = c("Ljung-Box"))
Box.test(Top.40$APN, type = c("Ljung-Box"))
Box.test(Top.40$BGA, type = c("Ljung-Box"))
Box.test(Top.40$BIL, type = c("Ljung-Box"))
Box.test(Top.40$BVT, type = c("Ljung-Box"))
Box.test(Top.40$CFR, type = c("Ljung-Box"))
Box.test(Top.40$DSY, type = c("Ljung-Box"))
Box.test(Top.40$FSR, type = c("Ljung-Box"))
Box.test(Top.40$GRT, type = c("Ljung-Box"))
Box.test(Top.40$IMP, type = c("Ljung-Box"))
Box.test(Top.40$INL, type = c ("Ljung-Box"))
Box.test(Top.40$IPL, type = c("Ljung-Box"))
Box.test(Top.40$ITU, type = c("Ljung-Box"))
Box.test(Top.40$MDC, type = c("Ljung-Box"))
Box.test(Top.40$MRP, type = c("Ljung-Box"))
Box.test(Top.40$MTN, type = c("Ljung-Box"))
Box.test(Top.40$NED, type = c("Ljung-Box"))
Box.test(Top.40$NPN, type = c("Ljung-Box"))
Box.test(Top.40$NTC, type = c("Ljung-Box"))
Box.test(Top.40$OML, type = c("Ljung-Box"))
Box.test(Top.40$RMH, type = c("Ljung-Box"))
Box.test(Top.40$SAB, type = c("Ljung-Box"))
Box.test(Top.40$SBK, type = c("Ljung-Box"))
Box.test(Top.40$SHP, type = c("Ljung-Box"))
Box.test(Top.40$SLM, type = c("Ljung-Box"))
Box.test(Top.40$SNH, type = c("Ljung-Box"))
Box.test(Top.40$SOL, type = c("Ljung-Box"))
Box.test(Top.40$TBS, type = c("Ljung-Box"))
Box.test(Top.40$WHL, type = c("Ljung-Box"))
Box.test(Midcap$ACL, type = c("Ljung-Box"))
Box.test(Midcap$AFE, type = c("Ljung-Box"))
```

```
Box.test(Midcap$ARI, type = c("Ljung-Box"))
Box.test(Midcap$ASR, type = c("Ljung-Box"))
Box.test(Midcap$AVI, type = c("Ljung-Box"))
Box.test(Midcap$BAT, type = c("Ljung-Box"))
Box.test(Midcap$BAW, type = c("Ljung-Box"))
Box.test(Midcap$CLS, type = c("Ljung-Box"))
Box.test(Midcap$DST, type = c("Ljung-Box"))
Box.test(Midcap$DTC, type = c("Ljung-Box"))
Box.test(Midcap$EOH, type = c("Ljung-Box"))
Box.test(Midcap$FBR, type = c("Ljung-Box"))
Box.test(Midcap$GFI, type = c("Ljung-Box"))
Box.test(Midcap$GND, type = c("Ljung-Box"))
Box.test(Midcap$HAR, type = c("Ljung-Box"))
Box.test(Midcap$HYP, type = c("Ljung-Box"))
Box.test(Midcap$ILV, type = c("Ljung-Box"))
Box.test(Midcap$ITE, type = c("Ljung-Box"))
Box.test(Midcap$KAP, type = c("Ljung-Box"))
Box.test(Midcap$LBH, type = c("Ljung-Box"))
Box.test(Midcap$LON, type = c("Ljung-Box"))
Box.test(Midcap$NHM, type = c("Ljung-Box"))
Box.test(Midcap$NPK, type = c("Ljung-Box"))
Box.test(Midcap$OCE, type = c("Ljung-Box"))
Box.test(Midcap$OMN, type = c("Ljung-Box"))
Box.test(Midcap$PIK, type = c("Ljung-Box"))
Box.test(Midcap$PPC, type = c("Ljung-Box"))
Box.test(Midcap$PSG, type = c("Ljung-Box"))
Box.test(Midcap$RCL, type = c("Ljung-Box"))
Box.test(Midcap$RLO, type = c("Ljung-Box"))
Box.test(Midcap$SAC, type = c("Ljung-Box"))
Box.test(Midcap$SAP, type = c("Ljung-Box"))
Box.test(Midcap$SNT, type = c("Ljung-Box"))
Box.test(Midcap$SPG, type = c("Ljung-Box"))
Box.test(Midcap$SUI, type = c("Ljung-Box"))
Box.test(Midcap$TFG, type = c("Ljung-Box"))
Box.test(Midcap$TON, type = c("Ljung-Box"))
Box.test(Midcap$TRE, type = c("Ljung-Box"))
Box.test(Midcap$TRU, type = c("Ljung-Box"))
Box.test(Midcap$TSH, type = c("Ljung-Box"))
Box.test(Smallcap$ACT, type = c("Ljung-Box"))
Box.test(Smallcap$ADH, type = c("Ljung-Box"))
Box.test(Smallcap$ADR, type = c("Ljung-Box"))
Box.test(Smallcap$AEG, type = c("Ljung-Box"))
Box.test(Smallcap$AEL, type = c("Ljung-Box"))
Box.test(Smallcap$AEN, type = c("Ljung-Box"))
Box.test(Smallcap$AFX, type = c("Ljung-Box"))
Box.test(Smallcap$BRN, type = c("Ljung-Box"))
```

```
Box.test(Smallcap$CAT, type = c("Ljung-Box"))
Box.test(Smallcap$CLH, type = c("Ljung-Box"))
Box.test(Smallcap$CSB, type = c("Ljung-Box"))
Box.test(Smallcap$GRF, type = c("Ljung-Box"))
Box.test(Smallcap$HDC, type = c("Ljung-Box"))
Box.test(Smallcap$MRF, type = c("Ljung-Box"))
Box.test(Smallcap$MTA, type = c("Ljung-Box"))
Box.test(Smallcap$MUR, type = c("Ljung-Box"))
Box.test(Smallcap$OCT, type = c("Ljung-Box"))
Box.test(Smallcap$PGR, type = c("Ljung-Box"))
Box.test(Smallcap$WBO, type = c("Ljung-Box"))
Box.test(Index$Top40Index, type = c("Ljung-Box"))
Box.test(Index$MidcapIndex, type = c("Ljung-Box"))
Box.test(Index$SmallcapIndex, type = c("Ljung-Box"))
```

Shapiro-Wilk Test

```
shapiro.test(Top.40$AGL)
shapiro.test(Top.40$AMS)
shapiro.test(Top.40$ANG)
shapiro.test(Top.40$APN)
shapiro.test(Top.40$BGA)
shapiro.test(Top.40$BIL)
shapiro.test(Top.40$BVT)
shapiro.test(Top.40$CFR)
shapiro.test(Top.40$DSY)
shapiro.test(Top.40$FSR)
shapiro.test(Top.40$GRT)
shapiro.test(Top.40$IMP)
shapiro.test(Top.40$INL)
shapiro.test(Top.40$IPL)
shapiro.test(Top.40$ITU)
shapiro.test(Top.40$MDC)
shapiro.test(Top.40$MRP)
shapiro.test(Top.40$MTN)
shapiro.test(Top.40$NED)
shapiro.test(Top.40$NPN)
shapiro.test(Top.40$NTC)
shapiro.test(Top.40$OML)
shapiro.test(Top.40$RMH)
shapiro.test(Top.40$SAB)
shapiro.test(Top.40$SBK)
shapiro.test(Top.40$SHP)
shapiro.test(Top.40$SLM)
```

```
shapiro.test(Top.40$SNH)
shapiro.test(Top.40$SOL)
shapiro.test(Top.40$TBS)
shapiro.test(Top.40$WHL)
shapiro.test(Midcap$ACL)
shapiro.test(Midcap$AFE)
shapiro.test(Midcap$ARI)
shapiro.test(Midcap$ASR)
shapiro.test(Midcap$AVI)
shapiro.test(Midcap$BAT)
shapiro.test(Midcap$BAW)
shapiro.test(Midcap$CLS)
shapiro.test(Midcap$DST)
shapiro.test(Midcap$DTC)
shapiro.test(Midcap$EOH)
shapiro.test(Midcap$FBR)
shapiro.test(Midcap$GFI)
shapiro.test(Midcap$GND)
shapiro.test(Midcap$HAR)
shapiro.test(Midcap$HYP)
shapiro.test(Midcap$ILV)
shapiro.test(Midcap$ITE)
shapiro.test(Midcap$KAP)
shapiro.test(Midcap$LBH)
shapiro.test(Midcap$LON)
shapiro.test(Midcap$NHM)
shapiro.test(Midcap$NPK)
shapiro.test(Midcap$OCE)
shapiro.test(Midcap$OMN)
shapiro.test(Midcap$PIK)
shapiro.test(Midcap$PP)
shapiro.test(Midcap$PSG)
shapiro.test(Midcap$RCL)
shapiro.test(Midcap$RLO)
shapiro.test(Midcap$SAC)
shapiro.test(Midcap$SAP)
shapiro.test(Midcap$SNT)
shapiro.test(Midcap$SPG)
shapiro.test(Midcap$SUI)
shapiro.test(Midcap$TFG)
shapiro.test(Midcap$TON)
shapiro.test(Midcap$TRE)
shapiro.test(Midcap$TRU)
shapiro.test(Midcap$TSH)
shapiro.test(Smallcap$ACT)
shapiro.test(Smallcap$ADH)
```

```
shapiro.test(Smallcap$ADR)
shapiro.test(Smallcap$AEG)
shapiro.test(Smallcap$AEL)
shapiro.test(Smallcap$AEN)
shapiro.test(Smallcap$AFX)
shapiro.test(Smallcap$BRN)
shapiro.test(Smallcap$CAT)
shapiro.test(Smallcap$CLH)
shapiro.test(Smallcap$CSB)
shapiro.test(Smallcap$GRF)
shapiro.test(Smallcap$HDC)
shapiro.test(Smallcap$MRF)
shapiro.test(Smallcap$MTA)
shapiro.test(Smallcap$MUR)
shapiro.test(Smallcap$OCT)
shapiro.test(Smallcap$PGR)
shapiro.test(Smallcap$WBO)
shapiro.test(Index$Top40Index)
shapiro.test(Index$MidcapIndex)
shapiro.test(Index$SmallcapIndex)
```

Bai-Perron (structural breaks) Test MATLAB code

```
% version 3, October 11, 2004
% These routines can be used by and distributed for non-profit academic
% purposes without any royalty except that the users must cite:
% Bai, Jushan and Pierre Perron (1998): "Estimating and Testing Linear
% Models with Multiple Structural Changes," Econometrica, vol 66,ÔøΩ@47-78
% and
%  Bai, Jushan and Pierre Perron (2003): "Computation and Analysis of
%  Multiple Structural Change Models," Journal of Applied Econometrics, 18, 1-22.
% For any other commercial use and/or comments, please contact Pierre
% Perron at perron at bu.edu.
% Even though we tried to make this program error-free we cannot be held
% responsible for any consequences that could result from remaining errors.
% Copyright, Pierre Perron (1999, 2004)

clear;
clc;

yyy=csvread('real.csv');  % read data
bigt=3996;                % set effective sample size
y=yyy(1:3996,1);          % set up the data, y is the dependent variable z is
                          % the matrix of regressors (bigt,q) whose coefficients
                          % are allowed to change, x is a (bigt,p) matrix of
                          % regressors with coefficients fixed across regimes.
```

```
                    % Note: initialize x to [] if p=0.
z=ones(bigt,1);

x=[];

q=1;                    % number of regressors z
p=0;                    % number of regressors x
m=5;                    % maximum number of structural changes allowed
eps1=.15;               % value of the trimming (in percentage) for the construction
                        % and critical values of the supF type tests (used in the
                        % supF test, the Dmax, the supF(l+1|l) and the sequential
                        % procedure). If these tests are used, h below should be set
                        % at int(eps1*bigt). But if the tests are not required,
                        % estimation can be done with an arbitrary h. There are five
                        % options: eps1= .05, .10, .15, .20, or .25. For each option,
                        % the maximal value of m above is: 10 for eps=.05, 8 for
                        % eps1=.10, 5 for eps1=.15, 3 for eps1=.20, and 2 for eps1=.25.

h=round(eps1*bigt);     % minimal length of a segment (h >= q). Note: if robust=1, h
                        % should be set at a larger value.

% The followings are options if p > 0 -------------------------------
fixb=0;                 % set to 1 if use fixed initial values for beta
betaini=0;              % if fixb=1, load the initial value of beta
maxi=20;                % maximum number of iterations for the nonlinear procedure to
                        % obtain global minimizers
printd=1;               % set to 1 if want the output from the iterations to be printed
eps=0.0001;             % criterion for the convergence
% --------------------------------------------------------------------
robust=1;               % set to 1 if want to allow for heterogeneity and autocorrelation
                        % in the residuals, 0 otherwise. The method used is Andrews(1991)
                        % automatic bandwidth with AR(1) approximation and the quadratic
                        % kernel. Note: Do not set to 1 if lagged dependent variables are
                        % included as regressors.
prewhit=1;              % set to 1 if want to apply AR(1) prewhitening prior to estimating
                        % the long run covariance matrix.
hetdat=1;               % option for the construction of the F tests. Set to 1 if want to
                        % allow different moment matrices of the regressors across segments.
                        % If hetdat=0, the same moment matrices are assumed for each segment
                        % and estimated from the ful sample. It is recommended to set
                        % hetdat=1 if p>0.
hetvar=1;               % option for the construction of the F tests.Set to 1 if want to allow
                        % for the variance of the residuals to be different across segments.
                        % If hetvar=0, the variance of the residuals is assumed constant
                        % across segments and constructed from the ful sample. This option
                        % is not available when robust=1.
```

```
hetomega=1;          % used in the construction of the confidence intervals for the break
                     % dates. If hetomega=0, the long run covariance matrix of zu is
                     % assumed identical across segments (the variance of the errors u
                     % if robust=0)
hetq=1;              % used in the construction of the confidence intervals for the break
                     % dates. If hetq=0, the moment matrix of the data is assumed identical
                     % across segments.
doglobal=1;          % set to 1 if want to cal the procedure to obtain global minimizers
dotest=1;            % set to 1 if want to construct the supF, UDmax and WDmax tests
                     % doglobal must be set to 1 to run this procedure.
dospflp1=1;          % set to 1 if want to construct the supF(l+1|l) tests where under
                     % the null the l breaks are obtained using global minimizers.
                     % doglobal must be set to 1 to run this procedure.
doorder=1;           % set to 1 if want to cal the procedure that selects the number of
                     % breaks using information criteria. doglobal must be set to 1 to
                     % run this procedure.
dosequa=1;           % set to 1 if want to estimate the breaks sequentialy and estimate
                     % the number of breaks using supF(l+1|l) test
dorepart=1;          % set to 1 if want to modify the break dates obtained from the
                     % sequential method using the repartition method of Bai (1995),
                     % Estimating breaks one at a time. This is needed for the confidence
                     % intervals obtained with estim below to be valid.
estimbic=1;          % set to 1 if want to estimate the model with the number of breaks
                     % selected by BIC.
estimlwz=0;          % set to 1 if want to estimate the model with the number of breaks
                     % selected by LWZ
estimseq=1;          % set to 1 if want to estimate the model with the number of breaks
                     % selected using the sequential procedure
estimrep=0;          % set to 1 if want to estimate the model with the breaks selected
                     % using the repartition method
estimfix=0;          % set to 1 if want to estimate the model with a prespecified number
                     % of breaks equal to fixn set below
fixn=0;

pbreak(bigt,y,z,q,m,h,eps1,robust,prewhit,hetomega,hetq,doglobal,dotest,dospflp1,doorder,doseq
ua,dorepart,estimbic,estimlwz,estimseq,estimrep,estimfix,fixb,x,p,eps,maxi,betaini,printd,hetdat,
hetvar,fixn)
```