Determining the impact of different forms of stationarity on financial time series analysis

Jan Adriaan van Greunen

Dissertation submitted in partial fulfilment of the requirements for the degree *Magister Commercii in Risk Management* at the Potchefstroom Campus of the North-West University

Supervisor: Dr André Heymans

Assistant supervisor: Dr Gary van Vuuren

November 2011

If you are, you breathe. If you breathe, you talk. If you talk, you ask. If you ask, you think. If you think, you search. If you search, you experience. If you experience, you learn. If you learn, you grow. If you grow, you wish. If you wish, you find and if you find, you doubt. If you doubt, you question. If you question, you understand and if you understand, you know. If you know, you want to know more and if you want to know more, you are alive...

Live Curious

- *National Geographic Channel* -

to my Mother, my Father and Alicia

Abstract

Since most time series data are non-stationary, the econometrician and financial analyst are required to make the data stationary before embarking on any econometric analysis in order to avoid spurious results. Although there are several different ways to render a non-stationary time series stationary, few econometricians and financial analysts look past the first differencing and log-differencing methods. Due to this "difference first, ask questions later" approach, this study aims to determine the impact of different forms of stationarity on financial time series analysis. Furthermore, this study aims to determine whether it is of any significance to consider one of the other methods of rendering a time series stationary rather than simply first differencing.

As a starting point, the literature on the different forms of stationarity as well as the tests for stationarity is reported. After an extensive review of the literature, it was found that there are at least five different forms of stationarity, each characterised by specific statistical properties of the particular time series. The literature also revealed that the most popular tests of stationarity are the DF-GLS, ADF and KPSS tests. Furthermore, the manner in which the fractional differencing parameter or fractional integration parameter of a time series is determined was reviewed. The methods used to determine the fractional differencing parameter, which were reported in this work, are that of the MRS and GPH methods. Incorporating all the tests and the GPH method, a novel process to determine the correct form of stationarity for a specific time series was introduced. The process was then applied to different types of time series data, which included stock prices, a stock index, consumer price index and an exchange rate. After finding that the time series do differ statistically and have different forms of stationarity, ARFIMA and OLS were employed. ARFIMA and OLS allowed each time series (in its own form of stationarity suggested by the relevant process) to be compared to the alternative form. For example, if a time series was found to be fractional difference stationary, its forecasting performance would be tested against its first differenced form. Results indicated that the form of stationarity found in a time series, after employing the relevant process, outperformed its alternative in every instance tested.

The results confirmed that it is indeed reckless to "difference first, and ask questions later". First differencing is not the only method that should be used to render a time series stationary, and it is imperative that econometricians and financial analysts begin exploring properties of the data and cease blindly following processes suggested for different datasets in the literature. The data should lead the analyst to the method that should be used to truly render a particular non-stationary time series stationary.

Opsomming

Aangesien die meeste tydreeksdata nie-stasionêr is, is ekonometrici vereis om die data stasionêr te maak voordat enige ekonometriese analise ondergaan word ten einde onwaar resultate te voorkom. Hoewel daar verskillende maniere is om ’n nie-stasionêre tydreeks stasionêr te maak, volg min ekonometrici ander metodes as eerste verskille en logaritmiese verskille. Weens hierdie "neem eerste verskille, vra vrae later"-benadering, poog hierdie studie om die impak van verskillende vorme van stasionariteit op finansiële tydreeksanalise te bepaal. Verder het hierdie studie ’n doel om te bepaal of dit van enige belang is om een van die ander metodes wat ’n tydreeks stasionêr maak, te volg, eerder as om eenvoudig eerste verskille te oorweeg.

As ’n beginpunt, is die literatuur oor die verskillende vorme van stasionariteit asook die toetse vir stasionariteit nageslaan. Na ’n uitgebreide oorsig van die literatuur is bevind dat daar ten minste vyf verskillende vorme van stasionariteit voorkom en elkeen verwys na die spesifieke statistiese eienskappe van die bepaalde tydreeks. Die literatuur het ook aan die lig gebring dat die gewildste toetse van stasionariteit die DF-GLS-, ADF- en KPSS-toetse is. Verder is die wyse waarop die fraksionele verskille parameter of fraksionele integrasie parameter van ’n tydreeks bepaal word, ondersoek. Die metodes wat gebruik word om die fraksionele verskille parameter of fraksionele integrasie parameter te bepaal, en wat in hierdie studie aangemeld is, is die MRS- en GPH-metodes. Deur van die verskillende toetse en die GPH-metode gebruik te maak, is ’n nuwe proses om die korrekte vorm van stasionariteit vir ’n spesifieke tydreeks vas te stel, ontwikkel. Die proses is dan toegepas op verskillende soorte tyd reeks data, wat aandele pryse, 'n aandele-indeks, verbruikersprysindeks en 'n wisselkoers insluit, Nadat bevind is dat die tydreeks statisties verskil en verskillende vorme van stasionariteit teenwoordig is, is ARFIMA en OLS ook in die studie gebruik. ARFIMA en OLS het dit moontlik gemaak om elke tydreeks in die vorm van stasionariteit, voorgestel deur die relevante proses, met die alternatiewe vorm te vergelyk. Byvoorbeeld, as gevind is dat ’n tydreeks stasionêr in fraksionele verskille is, sou sy voorspellingresultate getoets word teen die eerste verskille vorm van dié tydreeks. Resultate het daarop gedui dat die vorm van stasionariteit wat gevind is deur van die relevante proses gebruik te maak, in elke geval beter as die alternatiewe vorm presteer het.

Die resultate bevestig dat dit inderdaad roekeloos is om die "neem nou verskille en vra vrae later"-benadering te volg. Eerste verskille is nie die enigste metode wat gebruik moet word om ’n tydreeks stasionêr te maak nie en dit is noodsaaklik dat ekonometrici die eienskappe van die data begin ondersoek en staak om blindelings ’n proses te volg wat vir ’n ander datastel in die literatuur

voorgestel is. Die data moet analiste lei na die metode wat gebruik moet word om ’n spesifieke nie-stasionêre tydreeks werklik stasionêr te maak.

Acknowledgements

The journey leading up to the completion of my dissertation has been packed with people who have influenced, motivated and inspired me. As I took a seat to start writing these acknowledgements, I could not think of the words to explain my gratitude and respect to all of those who have had an impact on my academic career. That said, I thought the best would be to just start writing. So here goes...

- Dr André Heymans – Thank you for the guidance and thoughts regarding this dissertation. You were always willing to listen to me speaking in circles about stationary time series and in the end summarised my thoughts in a way that even I could understand them.
- Dr Chris van Heerden (Cousin) – Thank you for the support, your inspirational work ethic and morals that never seem to falter.
- Dr Gary van Vuuren – Thank you for your valuable insights and for helping me master the quantitative methods used in this dissertation.
- Prof Paul Styger – Just as all of those students before me, you have been an inspiration and a mentor to me. The fact that I have come this far in my academic career, and plan to go even further, has mainly been driven by your belief in my ability. I sincerely thank you for everything, but most of all, I thank you for instilling in me the desire to never stop learning.
- Alicia – Thank you for your never-ending support and patience through all the long days and nights. You have been with me right from the start of my academic career and have never ceased to motivate me in achieving new academic heights. Our journey called life has only just begun and I can’t wait for the rest of it.
- My family – Thank you for your support and always believing in me.
- Thank you to all my lecturers who have passed their knowledge on to me. Special thanks to Prof Marianne Matthee, Prof Andrea Saayman and Prof Waldo Krugell.
- To all my colleagues for the past year, thank you. I have thoroughly enjoyed working with you during this brief but exceptional chapter in my life.
- I would also like to thank the School of Economics for providing me with this opportunity to further my academic career.
- I would like thank Herman and Dianne Meere for their support and encouragement.
- I thank the Lord for giving me the ability to complete this dissertation, and
- Mom, thank you for all the sacrifices you have made to give me this opportunity. You have made this possible. Thank you.

- The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.
- Finally, thank you to Cecile van Zyl for the editing of this dissertation.

Table of contents

CHAPTER 1 – INTRODUCTION

1.1 BACKGROUND

Financial time series are said to follow a random walk (Gujurati, 2006:499). The explanation of a random walk process is often presented in the literature by a drunkard leaving a bar and every step taken in a direction contributing to a random walk (Gujurati, 2006:500). An alternative, less famous, explanation is that used by Murray (1994:37). Picture a new-born puppy just beginning to find its footing. Where it lies at a certain point in time, an interesting scent passes in front of the puppy's nose. Curiously, the puppy gets up and stumbles, taking random steps, in a direction, until a new scent leads it in a different direction. The result is a puppy taking a random walk while drifting off in different directions due to random hints of interesting scents (Murray, 1994:37). Such is the randomness of financial markets, the stocks[1] within in these markets and the decisions made by economic units involved in these markets.

Stock prices follow a random walk process, meaning that each realised price of a stock is represented by its previous realised value plus a random shock (Enders, 2010:4). This inherent property of stock prices has led to the econometric term known as non-stationary time series. A non-stationary time series refers to a process whose statistical properties change over time (Maddala & Kim, 2000:22). Alternatively, a non-stationary time series is described as a process containing a unit-root and being integrated to an order of n, written as an $I(n)$ process. It is, therefore, clear that stock prices or other financial time series are non-stationary. The problem arises when using such non-stationary time series in econometric modelling.

Granger and Newbold (1974:112) refer to the results of regressions with non-stationary data as "nonsense" or "spurious" regression results. A spurious regression is a name given to regressions that were conducted using time series variables that provided high R^2 (goodness of fit) estimates and low Durbin-Watson statistics. The results of the regression, therefore, provide the illusion of accuracy and robustness, but are in fact not significant (Gujurati, 2006:493). To overcome the problem of spurious regressions, the variables that are included in the regression must be stationary.

[1] Although in South Africa the term 'shares' is used this study will make use of the more internationally used term 'stocks'.

The literature refers to two types of stationary time series, namely strictly stationary and covariance stationary.[2] A strictly stationary time series refers to a time series whose properties remain unchanged regardless of the time period selected within that specific time series (Fielitz, 1971:1025). The second type of stationary time series is a covariance stationary time series. A covariance stationary time series has a less restricted definition than a strictly stationary time series. By definition, a covariance stationary time series should have a mean, variance and covariance that remain constant over time (Enders, 2010:54). In econometric modelling, covariance stationary is the most common form of stationary time series, and in general, when stating that a time series is stationary, it is in fact covariance stationary. However, as discussed earlier, most financial time series are non-stationary and need to be transformed into a stationary time series.

There are several different methods that can be used to render a non-stationary time series stationary. Some of the more popular methods among financial time series analysts are the first difference and log-difference methods.[3] These methods are so popular that they have become common practice to simply difference a non-stationary time series, obtaining a stationary time series, using a unit root test to confirm stationarity and to continue with modelling. Financial time series have been found to be inherently non-stationary and integrated to the order of $I(n)$, with n taking on the value of an integer. The value of n indicates the number of times it is required to difference a non-stationary time series in order to reduce it to stationarity (Burke & Hunter, 2005:23). Therefore, when a time series is $I(1)$, it is first differenced and becomes difference stationary and, when a time series is $I(2)$, second differences are taken and the series becomes differenced stationary. However, if a time series is differenced more than once, while only being integrated to the order of one, the time series becomes over-differenced (Ashley & Verbrugge, 2005:4).

A time series that is over-differenced becomes stripped of its statistical properties that are of use during econometric modelling (Plosser & Schwert, 1977; De Jong & Whiteman, 1993). Besides being integrated to the order of an integer, a time series can also be fractionally integrated and denoted as being $I(d)$. The fractional integration parameter or fractional dif-

[2] See Maddala and Kim (2000:10); Montgomery, Jennings and Kulahci (2008:25); Enders, (2010:54) and Asteriou and Hall (2007:231).

[3] See McCabe and Tremayne (1995:1015); Leybourne, McCabe and Tremayne (1996:45); Diebold and Yilmaz (2008:4); Lien and Yang (2009:142) and Joshi (2011:2).

ference parameter d, takes on the value of a non-integer. Similar to the over-differenced time series discussed above, a fractionally integrated time series (where $0 < d < 1$) that is first differenced will lead to that series being over-differenced (Gil-Alana, 2006:34).. If the time series in question is, however, fractionally differenced, it would produce a fractionally differenced stationary time series that has not been over-differenced (Burke *et al.*, 2005:31).

Since first differencing is a popular method of achieving stationarity, it is possible that some of the financial time series that are being first differenced could be fractionally integrated. Also, fractional differencing is only one alternative method of obtaining a stationary time series. Other forms of stationary time series include trend stationary, seasonal stationary and cyclical stationary time series (Burke *et al*., 2005:30; Montgomery, Jennings & Kulahci, 2008:39; Enders, 2010:191). This suggests that some of the data used in financial econometric analysis are over-differenced, providing results and inferences that are less accurate than expected.

1.2 PROBLEM STATEMENT AND RESEARCH QUESTION

The problem has become apparent based on the discussion thus far. Since a considerable amount of time series data are in fact fractionally integrated, there exists a real danger that many research studies did in fact arrive at inferior results because of over-differencing the data. In order to test the validity and extent of this problem, the following question should be answered. Does time series analysis with data made stationary with the wrong method lead to inferior results in terms of statistical significance and regression accuracy?

To assess the validity of this question, it is necessary to consider the *status quo* in time series analysis. Most econometricians and financial analysts simply take first differences in order to render a time series stationary. In doing so, important information and statistical properties in the data are lost, leading to inferior results.

1.3 RESEARCH AIM AND OBJECTIVES

The aim of this study is to determine the impact of different forms of stationarity on financial time series analysis. Furthermore, this study aims to determine whether it is of any significance to consider one of the other methods of rendering a time series stationary rather than simply first differencing. Objectives set to ultimately reach this aim are to, firstly, become familiar with the different methods that are available and can be used to render a non-stationary financial time series stationary. Secondly, to investigate the tests that are used to determine whether a time series is stationary or not. Thirdly, to apply this research to financial and economic time series to determine whether such time series do have different forms of stationarity. Lastly, to compare these different forms of stationarity to their conventionally used first differenced forms in order to determine whether there are significant differences in modelling performance between these forms.

1.4 DISSERTATION OUTLINE

The introductory chapter – Chapter 1 – is followed by another four chapters with Chapters 2, 3 and 4 presented as articles. All three of these articles have been submitted to Economic Research Southern Africa (ERSA) for publication in their working paper series and will thereafter be submitted for publication in national and international journals. The fifth chapter concludes and provides recommendations for future work.

Chapter 2 (Paper 1) focuses on the stationarity of financial time series. The concept of stationarity has always been central to econometric time series analysis, since most financial time series analysis necessitates that data be made stationary before any regressions can be performed. It has become common practice to transform non-stationary financial time series by either differencing data to the order *I(1)* or obtaining the log-normal returns. However, this process often leaves the data bereft of their descriptive value and, therefore, ineffective for financial time series analysis. This chapter challenges, by means of an extensive literature study, the common practice of differencing data indiscriminately by presenting an overview of what has been established through other methods of achieving stationarity, including fractional differencing and de-trending. The assumption that time series data are always first difference stationary is challenged.

Chapter 3 (Paper 2) analyses and tests the stationarity of financial time series. According to the literature, it should be assumed that time series data are first difference stationary.[4] In order to ensure that their time series data are stationary, many econometricians and finan-

[4] "The cornerstone of practical time series modelling is their acceptability of the difference stationary assumption" (McCabe & Tremayne, 1995:1015), "Much of modern applied econometric analysis is predicated on the assumption that data series concerned are non-stationary and that … they can be differenced to achieve stationarity" (Leybourne, McCabe & Tremayne, 1996:45) and "... taking the first difference of the non-stationary process has reduced it to stationarity" (Burke & Hunter, 2005:22) to quote only a few.

cial analysts are, therefore, led into merely taking first differences or logarithmic differences in order to make their data stationary. In order to test the validity of this "difference first, ask questions later" approach, this chapter will employ a rigorous process in order to determine the form of stationarity in financial time series data. By doing so, this chapter will enable econometricians and financial analysts to discern between those time series that are trend-stationary, those that are fractional difference stationary and those that are in fact first difference stationary.

Chapter 4 (Paper 3) focuses on the application of different forms of stationarity in financial time series analysis. Since most time series data are non-stationary, the econometrician and financial analyst are required to make the data stationary before embarking on any econometric analysis in order to avoid spurious results. Although there are several different ways to render a non-stationary time series stationary, few econometricians and financial analysts look past the first differencing and log-differencing methods. In order to determine whether this approach is indeed the correct one, a novel process is employed to test whether using the correct form of stationary data enhances the forecasting ability of models such as ARFIMA. The results corroborate this hypothesis in that all the time series that were found to be fractional difference stationary, outperformed their first difference form and *vice versa*.

Chapter 5 explores three different aspects regarding this dissertation. Firstly, to provide conclusions on the research conducted and the results that were obtained with the aim of determining the impact of different forms of stationarity on financial time series analysis. Secondly, to provide insight into whether it is of any significance to consider a method different from first differencing when transforming a non-stationary time series into a stationary process, and lastly, to provide recommendations and suggestions for further research in the field of financial econometrics and risk management.

1.5 CONCLUSION

First differencing a financial time series has become a commonly used method to render a non-stationary process stationary. This approach of first differencing will be scrutinised through comparison with other methods of achieving stationarity. The aim of this study is, therefore, to determine the impact of different forms of stationarity on financial time series analysis. The following chapter will focus on the stationarity of financial time series and will present an overview of what has been established through other methods of achieving stationarity. Such methods include fractional differencing and de-trending.

CHAPTER 2 |PAPER 1| THE STATIONARITY OF FINANCIAL TIME SERIES

The stationarity of financial time series

Jan van Greunen,[5] André Heymans[6] and Gary van Vuuren[7]

Abstract

The concept of stationarity has always been central to econometric time series analysis, since most financial time series analysis necessitates that data be made stationary before any regressions can be performed. It has become common practice to transform non-stationary financial time series by either differencing data to the order *I(1)* (first difference) or using log-normal returns. However, first differencing often leaves the data bereft of its descriptive value and, therefore, ineffective for financial time series analysis. This work challenges the common practice of differencing data indiscriminately by presenting an overview of what has been established through other methods of achieving stationarity, including fractional differencing and de-trending. The assumption that time series data are first difference stationary (and that the correct form of differencing should be performed before attempting any regression analysis or forecasting) is challenged.

JEL Classification: C22, G00.

Keywords: Fractionally differenced stationarity, unit roots, financial time series analysis.

1. Introduction

In applied econometric analysis concerning time series data, the prerequisite of stationarity is a well-known concept. A stationary time series is represented by data over time whose statistical properties remain constant regardless of a change in the time origin (Fielitz, 1971:1025). Research conducted on financial markets suggests that financial time series follow a random walk (Fama, 1970). A random walk process is inherently non-stationary, because of the presence of a unit root (Burke & Hunter, 2005:22). When a time series contains a unit root, it is necessary to difference the time series to render it stationary (Box & Jenkins, 1976). The first difference approach has also become popular mainly because many macroeconomic time series are difference stationary and not trend stationary (Nelson & Plosser, 1982). It has, therefore, become common practice to transform non-stationary fi-

[5] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: jan.vangreunen@nwu.ac.za.

[6] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: andre.heymans@nwu.ac.za.

[7] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: gary.vanvuuren@nwu.ac.za.

nancial time series by either differencing the data to the order *I(1)* or using the log-normal returns. Simple differencing could over-difference a time series (Burke *et al*., 2005:23), leaving the data void of their descriptive value and, therefore, ineffective in financial time series analysis (Plosser & Schwert, 1977; De Jong & Whiteman, 1993). This paper aims to challenge the common practice of differencing data indiscriminately by presenting an overview of what has been established through other methods of achieving stationarity. These methods include, but are not limited, to achieving stationarity through fractional differencing and detrending.

In Section 2, this paper briefly provides an overview of the different forms of stationarity found in financial time series data. Section 3 provides an in-depth review of the literature surrounding these different forms of stationarity and Section 4 draws attention to the different tests used to determine the stationarity of financial time series (including unit root tests and tests with stationarity as the null-hypothesis). A conclusion and critical overview of the findings are provided in Section 5.

2. A brief overview of the forms of stationarity in time series data

The terms non-stationary and stationary form a fundamental part of time series econometric analysis. A stationary time series refers to data whose statistical properties remain unchanged over time, regardless of the change in time origin (Fielitz, 1971:1025). Investigating time series data for the purpose of obtaining significant properties of these time series will be meaningless if the data are non-stationary or cannot be transformed to be stationary (Fielitz, 1971:1025). Using non-stationary time series in regression analysis can lead to spurious regression (Asteriou & Hall, 2007:293). Capturing and examining the properties of financial time series, therefore, require that univariate financial time series are stationary before examining such data.

There are two types of stationarity. Firstly, a time series is considered strictly stationary if the data properties remain unaffected by a shift in the time origin (Maddala & Kim, 2000:10; Montgomery, Jennings & Kulahci, 2008:25). Secondly, a time series is covariant stationary when it exhibits the following characteristics:

i) the time series is mean revertingor has long memory;[8] and

ii) a finite variance can be observed in the time series as the lag length increases when observing a correlogram of the time series. This theoretical correlogram diminishes faster than the theoretical correlogram of non-stationary time series (Asteriou *et al.*, 2007:231).

There are also different forms of stationarity that refer to the method used to render a non-stationary time series stationary. The different forms of stationarity can be described as follows: a) data may be trend-stationary, indicating that they comprise of stationary variances around a linear trend and they are made stationary by removing this linear trend (Burke *et al.*, 2005:30; Enders, 2010:191); b) data are difference stationary if they contain a unit root and can be rendered stationary by differencing them according to their level of integration (McCabe & Tremayne, 1995:1015; Laybourne, McCabe & Tremayne, 1996:435; Enders, 2010:192); c) a time series may be fractionally integrated and will require fractional differencing to reduce the time series to stationarity (Burke *et al.*, 2005:31); and d) cyclical and seasonal components might be present in data and they might still be stationary after both first differencing and seasonally differencing them, respectively (Enders, 2010:192; Montgomery *et al.*, 2008:39). A time series may, therefore, also be either cyclical-stationary or seasonal-stationary.

Subsequent sections of this work will provide a critical overview of the different forms of stationarity that may be observed in financial time series analysis by means of an extensive literature review.

3. Stationarity

3.1 Strictly stationary time series

The first type of stationarity to be reviewed is that of a strictly stationary process. According to Maddala *et al*. (2000:10) and Montgomery *et al.*, (2008:25), a time series is strictly stationary if its properties remain unaffected by a shift in the time origin. A time series is, therefore, strictly stationary if the distribution of the series x_t, x_{t+1},..., x_{t+n} is equal to the joint distribution of the series x_{t+k}, x_{t+k+1},..., x_{t+k+n}. A strictly stationary time series is further characterised as having a constant mean and constant variance (Maddala *et al.*, 2000:27).

[8] The series will fluctuate around a constant, long-run mean (Asteriou *et al.*, 2007:231).

3.2 Covariance stationary time series

A time series is defined as being covariance stationary (Enders, 2010:54) if it exhibits long memory and finite variance. Also, as the lag length increases when observing a correlogram of the time series, the theoretical correlogram diminishes faster than a theoretical correlogram of a non-stationary time series (Asteriou *et al*., 2007:231). Mathematically, the characteristics of a covariance stationary time series can be expressed as follows (Enders, 2010:54):

$$E(x_t) = E(x_{t-s}) = \mu, \tag{1}$$

$$E[(x_t - \mu)^2] = E[(x_{t-s} - \mu)^2] = \sigma_x^2, \tag{2}$$

$$E[(x_t - \mu)(x_{t-s} - \mu)] = E[(x_{t-j} - \mu)(x_{t-j-s} - \mu)] = \gamma_s, \tag{3}$$

where Equation 1 refers to the time series exhibiting a constant mean, Equation 2 describes constant variance of the time series, and Equation 3 establishes that the time series has a constant covariance over time. The subscript t is the time period, s is the shift in the time origin and j represents the number of lags. Covariance stationary time series are also known to be weakly-stationary, second-order stationary or wide-order stationary (Enders, 2010:54). A time series that does not exhibit these characteristics is, therefore, non-stationary, by definition.

3.3 Differencing and stationarity

A time series is differenced stationary if the series contains a unit root and can be rendered stationary by differencing the time series according to the level of integration (McCabe *et al*., 1995:1015, Laybourne *et al.,* 1996:435; Enders, 2010:192). The level of integration can be an integer value or a non-integer value.

3.3.1 Differenced stationary financial time series

The first difference approach has become popular mainly because of the work of Nelson *et al*. (1982), who argued that many microeconomic time series are difference stationary and not trend stationary. The widespread popularity of differencing a time series to achieve stationarity may be attributed to McCabe *et al*. (1995:1015),[9] Leybourne *et al*. (1996:45),[10] and Burke *et al*. (2005:22).[11]

The first difference of a process refers to the change within the process from one time period to the next (Burke *et al*., 2005:22). The first difference of a time series x_t is, therefore, $x_t - x_{t-1}$ and is denoted by Δx_t. According to Burke *et al*. (2005:22), $\Delta x_t = \varepsilon_t$ with ε_t being a white noise process and Δx_t, therefore, being a stationary process. It is also possible to take the second difference of a time series, i.e. $\Delta^2 x_t = \Delta\varepsilon_t$. However, if a time series that is *I(1)* is second differenced, the time series will be over-differenced (Ashley & Verbrugge, 2005:4). Care should, therefore, be taken that a time series is only differenced by the minimal number of times needed to render the series stationary (Burke *et al*., 2005:23).

Besides being integrated to the order of one or two and requiring first or second differencing to achieve stationarity, a time series may also be fractionally integrated. A fractionally integrated time series needs to be fractionally differenced in order to render the time series stationary.

3.3.2 Fractionally differenced stationary financial time series

First differencing is used by most econometricians and financial analysts as an alternative to fractional differencing, due to the difficulties associated with the latter method (Erfani & Samimi, 2009:1721). However, by replacing fractional differencing with first differencing, the data are often over-differenced and inherent properties of the data are lost (Plosser & Schwert, 1977; De Jong & Whiteman, 1993). The importance of fractional differencing, therefore, lies in the fact that data may be reduced to stationarity without over-differencing and, therefore, retaining properties essential to forecast modelling accuracy.

The order of differencing used during the fractional differencing of the financial time series is determined by calculating the specific series' fractional differencing parameter, *d*. Several different methods of determining *d* have emerged in the empirical literature. Authors of these methods include Hurst (1951), Mandelbrot (1972:259), Davies and Hart (1987:95), Hosking (1981:175), Lo (1991) and Peng, Havlin, Stanley and Goldberger (1994:82). The *d*-

[9] "The cornerstone of practical time series modelling is their acceptability of the difference stationary assumption".
[10] "Much of modern applied econometric analysis is predicated on the assumption that data series concerned are non-stationary and that … they can be differenced to achieve stationarity".
[11] "... taking the first difference of the non-stationary process has reduced it to stationarity".

parameter is derived from determining the Hurst coefficient (*H*) of a time series that indicates long memory properties of the time series (Garcia, Percival, Cannon, Raymond & Bassingwaighte, 1997:10). *H* may be estimated by using rescaled range analysis (R/S) and de-trended fluctuation analysis (DFA) (Caccia *et al.*, 1997:610; Erfani *et al.*, 2009:172). This paper provides a discussion of the R/S method followed by the subsequent steps needed to determine *d*: the DFA method is discussed in Erfani *et al.* (2009:1723) and falls outside the focus of this paper. The Hurst coefficient was developed in 1951 by applying rescaled analysis (R/S statistic) to the presence of long-range correlations within time series (Lo, 1991:1280). The R/S statistic is determined by dividing time series into windows, calculating the mean of each window, determining the range of the cumulative sum of the observations within each window, and dividing by the corresponding window's standard deviation (Lo, 1991:1286; Caccia *et al.*, 1997:614; Erfani *et al.*, 2009:173). The R/S statistic is denoted by R/S_n and defined as:

$$R/S_n \equiv \frac{1}{s_n}\left[\underset{0\le k\le n}{Max}\sum_{j=1}^{k}(X_t-\bar{X}_n)-\underset{0\le k\le n}{Min}\sum_{j=1}^{k}(X_t-\bar{X}_n)\right], \quad (4)$$

where s_n is the standard deviation estimator:

$$s_n \equiv \sqrt{\sum_j \frac{(X_j-\bar{X}_n)^2}{n}}. \quad (5)$$

Equations 4 and 5 are applied to each window and the *R/S* is obtained for all time periods (Erfani *et al.*, 2009:1723). After obtaining the R/S, the H is estimated by determining the slope of a linear regression, with log(R/S_n) as the dependent variable and the log of the window length *n*. *H* is, therefore, estimated by running an Ordinary Least Squares (OLS) regression:

$$\log(R/S_n) = \alpha + H\log(n). \quad (6)$$

If 0<*H*<1, the time series exhibits long memory and *d* is calculated by subtracting 0.5 from *H* (Hosking, 1981:167). The above explained method forms the most basic of the rescaled analysis methods. The R/S analysis has been shown to be superior to methods such as:

i) the analysis of auto-correlations;

ii) variance ratios; and

iii) spectral decompositions (Mandlebrot & Wallis, 1969a; Mandelbrot, 1972; Mandelbrot & Taqqu, 1979).

The findings reported above may be contested, but there is no doubt about the ability of R/S analysis to determine the long memory in time series (Lo, 1991:1288). A significant shortcoming of the R/S analysis is its large measure of sensitivity to short-range dependence. The latter implies that the R/S statistic might be a product of the short-term memory of a series rather than that of the long-term memory properties of the particular series.

Lo (1991:1289) suggested a new R/S analysis that provides the econometrician and financial analyst with the manner of distinguishing between short- and long-term dependence. As a result, the R/S statistic is modified to change due to long memory processes and to be independent of short memory processes. This method of rescaled range analysis has since been referred to as the modified R/S statistic (MRS) (Lo, 1991:1289; Erfani *et al.*, 2009:173). MRS is denoted by R'/S_n and is defined as (Lo, 1991:1289; Erfani *et al.*, 2009:173):

$$R'/S_n \equiv \frac{1}{\hat{\sigma}_n(q)}\left[\underset{0\le k\le n}{Max}\sum_{j=1}^{k}(X_t-\bar{X}_n)-\underset{0\le k\le n}{Min}\sum_{j=1}^{k}(X_t-\bar{X}_n)\right], \tag{7}$$

where:

$$\hat{\sigma}_n^2(q) \equiv \frac{1}{n}\sum_{j=1}^{n}(X_j-\bar{X})^2+\frac{2}{n}\sum_{j=1}^{q}\omega_j(q)\left[\sum_{i=j+1}^{n}(X_i-\bar{X}_n)(X_{i-j}-\bar{X}_n)\right], \tag{8}$$

$$\hat{\sigma}_n^2(q) \equiv \hat{\sigma}_x^2+2\sum_{j=1}^{q}\omega_j(q)\hat{\gamma}_j,$$

$$\omega_j(q) \equiv 1-\frac{j}{q+1}, \tag{9}$$

$$q<n.$$

The sample auto-covariance and variance of X in the above equations are $\hat{\gamma}_j$ and $\hat{\sigma}_n^2$, respectively. The difference between R/S in Equation 4 and the MRS in Equation 7 lies in the denominator. On the one hand, the R/S statistic is calculated by using the corresponding window's standard deviation, while on the other hand, the MRS is calculated by using the square root of the window's estimated variance and the weighted auto-covariance up to lag *q* (Lo, 1991:1290; Erfani *et al.*, 2009:1723). Following the calculation of the MRS for different windows of length *n*, the analysis follows that of the R/S statistic. Therefore, using the method of OLS, Equation 10 is regressed:

$$\log(R'/S_n) = \alpha + H \log(n). \tag{10}$$

The slope of the regression (Equation 10) represents the *H*-coefficient, which indicates the presence of long memory if 0<*H*<1. Furthermore, *d* is calculated by subtracting 0.5 from *H* (Hosking, 1981:167). Hosking (1981:175) devised, as an alternative to the methods discussed above, a maximum likelihood method that may be used to estimate *d* (Hosking, 1981:175).

After calculating *d,* it is used to reduce a non-stationary financial time series to a stationary financial time series by fractionally differencing the series with *d*. This method of using MRS in calculating *H* and *d*, and then fractional differencing to achieve stationarity was used by Erfani *et al*. (2009:1724) on a time series of the daily closing prices of the Tehran Stock Exchange index.

After establishing *d*, the fractionally differenced financial time series, w_t, is obtained as follows:

$$w_t = (1-L)^d x_t, \tag{11}$$

where, w_t is the fractionally differenced financial time series, x_t is the financial time series in levels, *d* is the fractional differencing parameter, L is the lag operator, and $(1-L)^d$ is the fractional difference operator defined as (Erfani *et al*., 2009:1724):

$$(1-L)^d \equiv \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} L^k. \tag{12}$$

Applying Equation 11 to non-stationary financial time series may reduce the particular time series to a fractionally differenced stationary financial time series (Erfani *et al*., 2009:1724).

In addition to differencing, a non-stationary time series may consist of equal variations, such as cyclical or seasonal variations, surrounding a deterministic trend. The presence of equal variations around the deterministic trend may indicate the presence of a stationary component being present in the time series. Therefore, by de-trending or filtering the time series, it will produce a time series that is stationary and has become stationary by not applying any method of differencing. The following section of this literature review reports the topics of trends, filtering and achieving stationarity without differencing.

3.4 Trend, seasonal and cyclical stationarity

Enders (2010:189) explains the difference between a time series containing a trend and a stationary time series by referring to the influences of shocks on a series. A stationary time series will only be affected temporarily by shocks and as time passes the series will revert back to its long-run mean. However, a shock to a series containing a trend will cause a series to deviate from its mean and will not return to its long-run level.

A financial time series may be non-stationary due to either a deterministic trend or a stochastic trend (Burke *et al*., 2005:30). A time series containing a deterministic trend may be de-trended to remove the trend and a series that contains a stochastic trend may be differenced to remove the trend (Enders, 2010:257). As a result, differencing is not an appropriate method for removing a deterministic trend and de-trending is not an appropriate manner of attempting to remove a stochastic trend (Enders, 2010:257). The importance of determining the type of trend present in financial time series, before simply differencing the series to obtain stationarity is highlighted by the latter. A non-stationary time series containing a stochastic trend that is integrated of order one, *I(1)*, is reduced to stationarity by differencing the series and is known as a difference-stationary time series (Clements & Hendry, 2001:S1). On the other hand, a trend stationary time series refers to a non-stationary time series that is rendered stationary only by de-trending the time series (Clements *et al*., 2001:S1).

A review of the literature and empirical study surrounding time series analysis reveals an array of different de-trending methods available to the time series analyst (e.g. Kalman, 1960; Hodrick & Prescott, 1997; Baxter & King, 1999). The methods established by these authors ranged from the simple methods (regression analysis) presented by Enders (2010:191) to the more complex methods (Hodrick-Prescott filter) presented by Hodrick *et al*. (1997). The most popular methods used by econometricians and financial analysts to de-trend macro-economic time series are the Hodrick-Prescott (HP) filter and the Baxter-King approximate band-pass (BK) filter (Aadland, 2002:2).

The process of de-trending financial time series containing a deterministic trend is explained by Enders (2010:191) as follows: Consider a time series consisting of a deterministic trend and a pure noise component:

$$y_t = y_0 + a_1 t + \varepsilon_t, \tag{13}$$

where y_0 refers to the initial condition for period zero, $a_1 t$ is the deterministic trend component and ε_t is the pure noise component. The time series, y_t, is de-trended by regressing Equation 13and obtaining the values of the series ε_t by subtracting the estimated values of y_t from the observed values in the time series. Alternatively, a financial time series may consist of a deterministic polynomial trend:

$$y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + ... + a_n t^n + e_t, \tag{14}$$

where e_t represents a stationary process (Enders, 2010:191). Under these circumstances, de-trending is achieved by means of regression with y_t as dependent variable and a deterministic polynomial time trend as independent variable. The *t*-tests, *F*-tests, and/or the Akaike Information Criterion (AIC) or Schwartz Bayesian Criterion (SBC) statistics are used to determine the correct degree of the polynomial.[12] The stationary and de-trended series e_t is the product of subtracting the estimated values of y_t from the observed values of y_t. The stationary de-trended time series yielded from the above methods may now be used in models used for time series analysis (Enders, 2010:191). Another method of de-trending employs certain unit root tests with stationarity as alternative or by using filters. Unit root tests form the central point of the discussion in the following section of this paper, while filters are discussed next.

Aadland (2005:290) explains that the HP filter is widely used in the de-trending of macroeconomic time series. In the HP filter, there is a trade-off between the squared deviations from a trend and a smoothness constraint. The HP filter is given by Hodrick *et al.* (1997) as:

$$h(L) = \frac{\lambda(1-L)^2(1-L^{-1})^2}{1+\lambda(1-L)^2(1-L^{-1})^2}, \tag{15}$$

[12] For a comprehensive explanation on how the *t*-tests, *F*-tests, and/or the Akaike Information Criterion (AIC) or Schwartz Bayesian Criterion (SBC) statistics are used to determine the correct degree of the polynomial, refer to Enders (2010:191).

where λ is an adjustable smoothness parameter. Apart from the HP filter, another widely used filter is the BK filter which is based on the theory of spectral band-pass filters. These filters may be used to remove a trend from a cyclical stationary component in a time series that is non-stationary due to the trend component.

Time series may also consist of a seasonal component, rather than a cyclical component, fluctuating around a trend. Montgomery *et al*. (2008:39) suggested the following for seasonal differencing:

$$\nabla_d y_t = (1 - B^d) = y_t - y_{t-d}, \tag{16}$$

where the lag-*d* is the seasonal difference operator. If a trend remains after seasonally differencing the data, Montgomery (2008:39) suggests continuing by first differencing the data. However, we suggest that Enders' (2010) approach is followed and that the method of de-trending the data should be followed before differencing the series. As discussed in Section 2 removing a trend from an otherwise stationary time series would render the time series trend stationary (Enders, 2010:191).

Up to this point, the different methods of rendering a financial time series stationary have been discussed. Subsequently, it is necessary to confirm stationarity by conducting a test for stationarity. The different tests used to confirm that a financial time series is stationary are discussed in the following section.

4. The testing of stationarity in financial time series

Testing stationarity in financial time series involves testing for the order of integration in the time series and, therefore, whether the time series possess a unit root. There are two principal tests popular among econometricians and financial analysts to test the null-hypothesis of a unit root to establish stationarity. These are the Augmented Dickey-Fuller (ADF) test for unit roots (a modification of the original Dickey-Fuller (DF) test) and the Phillips-Perron test (Asteriou *et al*., 2007:297). There are also several other tests to test stationarity with stationarity as the null-hypothesis. A popular such test is the KPSS test (after Kwaitkowski, Phillips, Schimdt & Shin, 1992). Other tests with stationarity as null-hypothesis include: Tanaka (1990), Park (1990), Saikkonen and Luukkonen (1993), Choi (1994), the Leybourne and McCabe test (1994) and Arellano and Pentula (1995).

4.1 Unit root tests

Dickey and Fuller (1979, 1981) developed a procedure to test non-stationarity based on the presence of a unit root. This procedure has become known as the Dickey-Fuller test (DF) for unit roots (Asteriou *et al.*, 2007:295). The DF test provides the econometrician and financial analyst with three different regressions that may be used to test for the presence of a unit root (Enders, 2010:206):

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t, \tag{17}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t, \tag{18}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \varepsilon_t. \tag{19}$$

The difference between Equations 17 to 19 is the presence of intercepts and a linear time trend, whereas Equation 18 includes an intercept and Equation 19 includes an intercept and a deterministic time trend. According to Asteriou *et al*. (2007:296) and Enders (2010:206), γ is the central focus point of all three equations and if $\gamma = 0$, the time series contains a unit root.[13] An alternative to the DF test is the Augmented Dickey-Fuller (ADF) test for stationarity.

In the unlikely event of a white noise error term, Dickey and Fuller extended the DF test to include extra lagged terms of the explanatory variable (Asteriou *et al.*, 2007:297). This extended version is known as the ADF test and the inclusion of extra lagged dependent variables eliminates the presence of auto-correlation. The ADF follows on the three different forms used in the DF test and can be conducted using the following three regressions (Asteriou *et al.*, 2007:297):[14]

$$\Delta y_t = \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t, \tag{20}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t, \tag{21}$$

[13] See Enders (2010:206), Asteriou *et al*. (2007:296) and Maddala *et al*. (2000:61) for an in-depth explanation of the DF test.

[14] See Enders (2010:215), Asteriou *et al*. (2007:297) and Maddala *et al*. (2000:75) for an in-depth explanation of the ADF test.

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t. \quad (22)$$

Asteriou (2007:295) explains a three step method for using the DF and ADF tests for unit roots with the aim of concluding that a series is stationary. The first is to test the time series for a unit root and, if none exists, the time series is stationary and *I(0)*, otherwise it contains a unit root and is *I(n)*. The second step is to take first differences and to test the first differenced series for a unit root and, if there is none, it is stationary and *I(1)*, otherwise it contains a unit root and is *I(n)*. The third step entails differencing the time series up to the point where the tests indicate no presence of a unit root. In the latter step, the time series will be integrated to the order of times needed to make the time series stationary.

Phillips and Perron (1988) developed the Phillips-Perron (PP) test, which is similar to the ADF test. The difference between the PP and ADF is that ADF corrects for auto-correlation by adding lagged values of the dependent variable. The PP test accounts for auto-correlation in e_t by making a correction to the *t*-stat of γ from the AR(1) regression. The PP test regression is given by Phillips and Perron (1988) as:

$$\Delta y_{t-1} = a_0 + \gamma y_{t-1} + e_t. \quad (23)$$

The DF, ADF and PP tests have been criticised for their lack of power in determining the presence of unit roots. More particularly, most of the criticism originates from these tests being useless when working with data frequencies greater than quarterly (Maddala *et al.*, 2000:45&92). This poses a problem for the econometrician analysing financial time series data that tend to be either daily or intra-day data. The ability of these tests to determine fractional unit roots is also unknown and poses a further problem for researchers aiming to determine whether a time series is fractionally differenced stationary after fractionally differencing the time series.

4.2 Test with stationarity as null-hypothesis

While tests for unit roots with stationarity as an alternative are popular among econometricians, there are also tests available that have stationarity as the null-hypothesis and a unit root as the alternative. The most popular among these tests is the KPSS test (Maddala *et al.*, (2000:120). The KPSS test was developed by Kwaitkowski, Phillips, Schimdt and Shin (1992) and incorporates the following model:

$$y_t = \delta_t + \zeta_t + \varepsilon_t, \qquad (24)$$

where ε_t represents a stationary process and ζ_t is a random walk specified as:

$$\zeta_t = \zeta_{t-1} + u_t, \quad u_t \sim iid(0, \sigma_u^2), \qquad (25)$$

and u_t is an independently and identically distributed (iid) error term.

The null-hypothesis is that of stationarity:

$H_0 : \sigma_u^2 = 0$ or ζ_t is a constant.

The Nabeya-Tanaka test statistic, also known as the LM test, for the hypothesis of the KPSS test is specified as (Maddala *et al.*, 2000:121):

$$LM = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma}_e^2}, \qquad (26)$$

where y_t is regressed on a constant and a time trend, where ε_t is the residuals. The residual variance of the regression is given by $\hat{\sigma}_e^2$ and s_t is the partial sum of ε_t defined by:

$$S_t = \sum_{i=1}^{t} e_t, \; t = 1, 2, ..., T. \qquad (27)$$

To determine whether the time series is stationary in levels, rather than testing trend stationarity, the test is conducted with a regression of y_t on an intercept. Nabeya and Tanaka derived the asymptotic distribution of the LM test statistic and the critical values are presented in Kwaitkowski *et al*. (1992:166).

4.3 The Dickey-Fuller test with GLS de-trending (DF-GLS)

As a method of overcoming the low power of the DF, ADF and PP unit root tests, Maddala *et al*. (2000:114) suggest using the DF-GLS test. The DF-GLS is a modification of the ADF and was proposed by Elliott, Rothenberg and Stock (ERS) (1996). The DF-GLS de-trends a time series y_t and produces a series y_t^d that replaces y_t in the original ADF test equation and is given as:

$$\Delta y_t^d = \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i}^d + \varepsilon_t. \qquad (28)$$

The critical values of the test statistic are provided in ERS (1996:825).

The discussion above includes a wide variety of tests available to the econometrician performing analysis on financial time series. The choice of which test to use is, therefore, made difficult. Kwaitkowski *et al.* (1992:176) and Choi (1994:721) suggest that it would be wise to use a combination of the tests discussed above. One such combination can comprise the use of the ADF and KPSS tests and is referred to as confirmatory analysis (Maddala *et al.*, 2000:126). Stationarity can be deduced if the null-hypothesis of the one test is rejected and the null-hypothesis of the other is not. The reason for this is that the null for the ADF is that of a unit root, while the null of the KPSS is that of stationarity.

5. Conclusion

In this work, it was found that the stationarity of time series is an essential characteristic required to be achieved before analysing time series data. Stationarity has been shown to be a prerequisite for performing applied econometric analysis. Due to this prerequisite, it has become common practice to difference financial time series data in order to achieve stationarity. The common practice of differencing data indiscriminately was challenged by presenting an overview of what has been established through other methods of achieving stationarity.

Different forms of stationarity have been reviewed. These include: 1) difference stationarity; 2) fractionally differenced stationarity; 3) trend stationarity; 4) cyclical; and 5) seasonal stationarity and are presented in Section 3.3 and 3.4. An important factor in rendering a time series stationary is to determine the presence of a stochastic trend, a deterministic trend, a unit root, a cyclical component or seasonal component. In the case of determining the presence of a stochastic trend or unit root, the econometrician must follow the differencing method for achieving stationarity. However, should a time series display a deterministic trend, the data must be de-trended and then tested for stationarity before any differencing can take place. A time series containing a cyclical component can be de-trended using the HP filter or BK filter to determine whether the series is cyclically stationary. A seasonal stationary time series can be achieved by removing a seasonal component from that time series by means of seasonal differencing. After a financial time series has been reduced to stationarity, it is necessary to confirm stationarity through tests. The popular tests available, that have a unit root as null-hypothesis with stationarity as an alternative, are the DF, ADF, PP and DF-GLS tests. The econometrician can also use the KPSS test where the null-hypothesis is that of stationarity. It has been suggested that the econometrician should use the KPSS test in conjunction with one of the unit root tests as a confirmatory analysis for the presence of stationarity.

The discussion above shows the econometrician and financial analyst that reducing a financial time series to stationarity is a much more integrate process than simply taking first difference, testing for stationarity with the ADF test, and proceeding with, for example, forecasting models. Therefore, when conducting applied econometric analysis to time series, it is necessary to investigate the properties of the time series before launching into differencing the data. Only thereafter should it be decided which method must be applied to reduce a financial time series to the correct form of stationarity.

6. References

AADLAND, D. 2005. De-trending time-aggregated data. *Economic Letters*, 89:287-93.

ARELLANO, C. & PENTULA, S.G. 1995. Testing for trend stationarity versus difference stationarity. *Journal of Time Series Analysis*, 16:147-164.

ASHLEY, R. & VERBRUGGE, R.J. Comments on "A critical investigation on detrending procedures for non-linear processes". *Virginia Tech Economics Department, Working paper.* http://ashleymac.econ.vt.edu/working_papers/ashley_verbrugge_comment.pdf. Date of Access: 4 May 2012. ASTERIOU, D. & HALL, S.G. 2007. Applied econometrics: a modern approach. New York: Palgrave Macmillan. 397 p.

BAXTER, M. & KING, R.G., 1999. Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81:575-593.

BOX, G.E.P. & JENKINS, G.E.M. 1976. Time series analysis: Forecasting and control. San Francisco: Holden Day. 575 p.

BURKE, S.P. & HUNTER, J. 2005. Modelling non-stationary economic time series: A multivariate approach. New York: Palgrave MacMillan. 253 p.

CACCIA, D.C., PERCIVAL, D., CANNON, M.J., RAYMOND, G. & BASSINGWAIGHTE, J.B. 1997. Analyzing exact fractal time series: Evaluating dispersional analysis and rescaled range methods. *Physica A*, 246(3-4):609-632.

CHOI, I. 1994. Residual-based tests for the null of stationarity with application to US macroeconomic time series. *Econometric Theory*, 10:720-746.

DAVIES, R.B. & HARTE, D.S. 1987. Tests for the Hurst effect. *Biometricka Trust*, 74(1):95-101. Mar.

DE JONG, D.N. & WHITEMAN, C.H. 1993. Estimating moving average parameters: Classical pileups and Bayesian posteriors. *Journal of Business and Economic Statistics*, 11(3):311-317.

DICKEY, D.A. & FULLER, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427-431. Jun.

DICKEY, D.A. & FULLER, W.A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057-172. Jul.

ENDERS, W. 2010. Applied econometric time series. Hoboken: Wiley. 517 p.

ERFANI, A. & SAMIMI, A.J. 2009. Long memory forecasting of stock price index using a fractionally differenced ARMA model. *Journal of Applied Sciences Research*, 5(10):1721-1731.
FAMA. E.F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383-417. May.
FIELITZ, B.D. 1971. Stationarity of random data: Some implications for the distribution of stock price changes. *Journal of Financial and Quantitative Analysis*, 6(3):1025-1034.
GUJARATI, D.N. 2006 Essentials of econometrics. New York: McGraw Hill. 553 p.
HODRICK, J.R. & PRESCOTT, E.C. 1997. Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1-16. Feb.
HOSKING, J.R.M. 1981. Fractional differencing. *Biometrika Trust*, 68(1):1025-1034.
HURST, H.E. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770-799.
KALMAN, R.E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1): 35-45.
KING, R.G. & REBELO, S.T. 1993. Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17:207-231.
KWAITKOWSKI, D. PHILLIPS, P.C.B., SCHMIDT, P. & SHIN, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159-178.
LEYBOURNE, S.J. & MCCABE, B.P.M. 1994. A consistent test for a unit root. *Journal of Business and Economic Statistics*, 12:157-166.
LEYBOURNE, S.J., MCCABE, B.P.M. & TREMAYNE, A.R. 1996. Can economic time series be differenced to stationarity? *Journal of Business & Economic Statistics*, 14(4):435-446. Oct.
LO, A.W. 1991. Long-term memory in stock market prices. *Econometrica*, 59(5):1279-1313. Sep.
MADDALA, G.S. & KIM, I.M. 2000. Unit roots, cointegration, and structural change. Cambridge: Cambridge. 505 p.
MANDELBROT, B. & TAQQU, M. 1979. Robust R/S analysis of long-run serial correlation. *Bulletin of the International Statistical Institute*, 42nd Conference Proceedings, Manila. 2:59-104.
MANDELBROT, B. & WALLIS, J. 1969a. Computer experiments with fractional Gaussian noises. *Water Resources Research*, 5:228-267. Feb.
MANDELBROT, B. 1972. Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Statistics*, 23(3):1015-1028.
MCCABE, B.P.M. & TREMAYNE, A.R. 1995. Testing time series for difference stationarity. *The Annals of Statistics*, 23(3):1015-1028.
MONTGOMERY, D.C., JENNINGS, C.L. & KULAHCI, M. 2008. Introduction to time series analysis and forecasting. Hoboken: Wiley. 445 p.
NELSON, C. & PLOSSER, C. 1982. Trends and random walks in macroeconomic time series models: some evidence and implications. *Journal of Monetary Economics*, 10(2):139-162.
PARK, J.Y. 1990. Testing for unit roots and cointegration by variable addition. *Advances in Econometrics*, 8:107-133.
PENG, C.K., HAVLIN, S., STANLEY, H.E. & GOLDBERGER, A.L. 1994. Quantification of scaling exponents and crossover phenomena in non-stationary heartbeat time series. *Chaos*, 5(1):82-87.
PLOSSER, C.I. & SCHWERT, G.W. 1977. Estimation of a noninvertible moving average process: The case of over-differencing. *Journal of Econometrics*, 6:199-224.
SAIKKONEN, P. & LUUKKONEN, R. 1993. Testing for a moving average unit root in ARIMA models. *Journal of the American Statistical Association*, 88:596-601.

CHAPTER 3 |PAPER 2| ANALYSING AND TESTING THE STATIONARITY OF FINANCIAL TIME SERIES

Analysing and testing the stationarity of financial time series

Jan van Greunen,[15] André Heymans[16] and Gary van Vuuren[17]

Abstract

According to the literature, it should be assumed that time series data are first difference stationary.[18] In order to ensure that their time series data are stationary, many econometricians are, therefore, led into merely taking first differences or logarithmic differences in order to make their data stationary. In order to test the validity of this 'difference first, ask questions later' approach, we employ a rigorous process in order to determine the form of stationarity in our time series data. By doing this, we are able to discern between those time series that are trend-stationary, those that are fractional difference stationary and those that are in fact first difference stationary. The results show that five out of the twelve time series tested were over-differenced after first differences were taken. This finding might hold serious implications for all studies where the stationarity of the data is a prerequisite before performing further analysis.

JEL Classification: C22, G00.

Keywords: Tests for stationarity, financial time series analysis.

1. Introduction

The first difference approach has become popular mainly because of the work of Nelson *et al*. (1982), who argued that many macroeconomic time series are difference stationary and not trend stationary. This finding caused econometricians to simply make use of the first difference approach in order to render data stationary. The works of McCabe and Tremayne (1995:1015),[19] Leybourne, McCabe and Tremayne (1996:45),[20] and Burke and Hunter

[15] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: jan.vangreunen@nwu.ac.za.

[16] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: andre.heymans@nwu.ac.za.

[17] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: gary.vanvuuren@nwu.ac.za.

[18] "The cornerstone of practical time series modelling is their acceptability of the difference stationary assumption" (McCabe & Tremayne, 1995:1015), "Much of modern applied econometric analysis is predicated on the assumption that data series concerned are non-stationary and that … they can be differenced to achieve stationarity" (Leybourne, McCabe & Tremayne, 1996:45) and "... taking the first difference of the non-stationary process has reduced it to stationarity" (Burke and Hunter, 2005:22) to quote only a few.

[19] "The cornerstone of practical time series modelling is the acceptability of the difference stationary assumption".

[20] "Much of modern applied econometric analysis is predicated on the assumption that data series concerned are non-stationary and that … they can be differenced to achieve stationarity".

(2005:22)[21] are examples of the popularity of the first difference approach. Since these tests were performed before the exponential growth in stock prices in 2006, the probability that the authors were correct in concluding that most financial time series are first difference stationary, could have been substantial.

During the last six years,[22] however, the properties of financial time series have changed considerably. An example has been the increase in volatility of financial time series during the 2007/08 financial crisis. This increased volatility could have led to a fundamental shift in the dynamics of financial time series analysis. It is, therefore, necessary to revisit the assumption that most time series are first difference stationary and to investigate the possibility that time series data might take on other forms of stationarity.

First differencing is used by most econometricians and financial analysts as an alternative to fractional differencing due to the difficulties associated with the latter method (Erfani & Samimi, 2009:1721). However, by replacing fractional differencing with first differencing the data are often over-differenced and the inherent properties of the data are lost (Plosser & Schwert, 1977; De Jong & Whiteman, 1993). The importance of fractional differencing, therefore, lies in the fact that data may be reduced to stationarity without over-differencing and, therefore, retaining properties essential to time series modelling.

In order to investigate the forms of stationarity that characterise financial time series data, this paper will commence with an overview of the methods applied to test for stationarity in financial time series (Section 2). The data are discussed in Section 3. In order to ensure that the data are rendered stationary, a novel process is introduced in Section 4 and applied in Section 5. The empirical results are also interpreted and explained in Section 5. Concluding remarks and recommendations are provided in Section 6.

2. Stationary time series

The majority of financial time series data are non-stationary in their raw form due to fluctuating means and covariances (Erfani *et al.*, 2009:1722). As a result, non-stationary financial time series need to be transformed in some manner to obtain a stationary series. The most popular method for transforming non-stationary data into stationary data is to simply take first differences of the time series. This, however, is where an important problem in applied

[21] "... taking the first difference of the non-stationary process has reduced it to stationarity".
[22] 2006 to 2011.

time series analysis arises (Erfani *et al*., 2009:1722). Time series analysts have become reluctant to accept differencing as an answer for achieving stationarity, contemplating whether properties of importance might be lost in the process (Erfani *et al*., 2009:1722). The latter thinking of econometricians and financial analysts can be justified by the variety of different methods that can be used to achieve stationarity. De-trending, differencing and fractional differencing are examples of methods used to render a non-stationary time series stationary and will be discussed and applied in Sections 2.2, 4 and 5. Attention will also be given to the tests used to determine whether a time series is stationary or non-stationary in Sections 2.1, 4 and 5.

2.1 Testing stationary financial time series

In order to test for stationarity in financial time series, it is necessary to test for the order of integration in the time series and, therefore, whether the time series possess a unit root. There are two principal tests popular among econometricians and financial analysts to test the null-hypothesis of a unit root to establish stationarity. These are the Augmented Dickey-Fuller (ADF) test for unit roots (a modification of the original Dickey-Fuller (DF) test) and the Dickey-Fuller Generalised Least Squares test (DF-GLS) (Maddala & Kim, 2000). There are also several tests to test stationarity with stationarity as the null-hypothesis. A popular test with stationarity as null-hypothesis is the KPSS test (after Kwaitkowski, Phillips, Schimdt and Shin, 1992). Other tests with stationarity as null-hypothesis include tests by Tanaka (1990), Park (1990), Saikkonen and Luukkonen (1993), Choi (1994), the Leybourne and McCabe test (1994) and Arellano and Pentula (1995). The DF and ADF tests are discussed in the next section, followed by a discussion of the KPSS and DF-GLS tests.

2.1.1 Unit root tests

Dickey and Fuller (1979, 1981) developed a procedure to test non-stationarity based on the presence of a unit root. This procedure has become known as the Dickey-Fuller test (DF) for unit roots (Asteriou & Hall, 2007:295). The DF test provides econometricians and financial analysts with three different regressions to test for the presence of a unit root (Enders, 2010:206):

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t, \qquad (1)$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t, \qquad (2)$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \varepsilon_t, \quad (3)$$

These regression equations are similar except for the presence of an intercept term and an intercept and trend in Equations 2 and 3, respectively. The central focus point of all three equations is γ and if $\gamma = 0$, the time series contains a unit root (Asteriou *et al*., 2007:296; Enders, 2010:206).[23] In the unlikely event of a white noise error term, Dickey and Fuller extended the DF test to include extra lagged terms of the explanatory variable (Asteriou *et al*., 2007:297). This extended version is known as the ADF test and the inclusion of extra lagged dependent variables eliminates the presence of auto-correlation. The ADF follows on the three different forms used in the DF test and can be conducted using the following three regressions (Asteriou *et al*., 2007:297):[24]

$$\Delta y_t = \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t, \quad (4)$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t, \quad (5)$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t. \quad (6)$$

Asteriou *et al*. (2007:295) explain a three-step method for using the DF and ADF tests for unit roots with the aim of concluding that a series is stationary. The first is to test the time series for a unit root and, if none exists, the time series is stationary and *I(0)*, otherwise it contains a unit root and is *I(n)*. The second step is to take first differences (in cases where a unit root is present), and testing the first differenced series for a unit root and, if there is none, it is stationary and *I(1)*, otherwise it contains a unit root and is *I(n)*. The third step entails differencing the time series up to the point where the tests indicate no presence of a unit root. In the latter step, the time series will be integrated to the order of times needed to difference the time series before it is stationary.

Although widely used, the DF and ADF tests have been criticised for their lack of power in determining the presence of unit roots. More particularly, most of the criticism regarding

[23] See Enders (2010:206), Asteriou *et al*. (2007:296) and Maddala *et al*. (2000:61) for an in-depth explanation of the DF test.

[24] See Enders (2010:215), Asteriou *et al*. (2007:297) and Maddala *et al*. (2000:75) for an in-depth explanation of the ADF test.

these tests stems from their inability to work with data frequencies greater than quarterly observations (Maddala *et al*., 2000:45). This poses a problem for the econometricians analysing financial time series data, which tends to be either daily or intra-day data. The ability of these tests to determine fractional unit roots is also unknown and poses a further problem for researchers aiming to determine whether a time series is fractional difference stationary after fractionally differencing the time series.

2.1.2 Test with stationarity as null-hypothesis

While tests for unit roots with stationarity as an alternative are popular among econometricians and financial analysts, there are also tests available that have stationarity as the null-hypothesis and a unit root as the alternative. The most popular among these tests is the KPSS test (Maddala *et al*., 2000:120). The KPSS test was developed by Kwaitkowski, Phillips, Schmidt and Shin (1992) and incorporates the following model:

$$y_t = \delta_t + \zeta_t + \varepsilon_t, \tag{7}$$

where ε_t represents a stationary process and ζ_t is a random walk specified as:

$$\zeta_t = \zeta_{t-1} + u_t,\ \ u_t \sim iid(0, \sigma_u^2). \tag{8}$$

and u_t is an independently and identically distributed (iid) error term.

The null-hypothesis is that of stationarity:

$H_0 : \sigma_u^2 = 0$ or ζ_t is a constant.

To determine whether the time series is stationary in levels, rather than testing trend stationarity, the test is conducted with a regression of y_t on an intercept. The asymptotic distribution of the LM test statistic was derived by Nabeya and Tanaka (1988) and the critical values are presented in Kwaitkowski *et al*. (1992:166).

2.1.3 The Dickey-Fuller test with GLS de-trending (DF-GLS)

As a method of overcoming the low power of the DF, ADF and PP unit root tests, Maddala *et al*. (2000:114) suggest using the DF-GLS test. The DF-GLS is a modification of the ADF and was proposed by Elliott, Rothenberg and Stock (1996). The DF-GLS de-trends a time series

y_t and produces a series y_t^d that replaces y_t in the original ADF test equation and is given by Elliott, *et al* (1996) as:

$$\Delta y_t^d = \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i}^d + \varepsilon_t \tag{9}$$

with the critical values of the test statistic provided in Elliot *et al*. (1996:825).

Since a wide variety of tests are available to the econometrician performing analysis on financial time series, the choice of which test to use is made difficult. Kwaitkowski *et al.* (1992:176) and Choi (1994:721) suggest that it would be wise to use a combination of the tests discussed above. One such combination can comprise the use the ADF and KPSS tests and is referred to as confirmatory analysis (Maddala *et al*., 2000:126). In this instance, stationarity can be deduced if the null-hypothesis of the ADF test is rejected and the null-hypothesis of the KPSS test is not.

2.2 Methods to obtain stationary financial time series

Three methods are explained that have the ability to render a non-stationary time series stationary. The method used will depend on the results obtained from the tests explained in Section 2.1. De-trending (Section 2.2.1), differencing with integers (Section 2.2.2) and fractional differencing (Section 2.2.3) are the three methods that will be incorporated and explained next.

2.2.1 De-trending

The process of de-trending financial time series containing a deterministic trend is explained by Enders (2010:191) as follows: Consider a time series consisting of a deterministic trend and a pure noise component:

$$y_t = y_0 + a_1 t + \varepsilon_t, \tag{10}$$

where y_0 refers to the initial condition for period zero, $a_1 t$ is the deterministic trend component and ε_t is the pure noise component. The time series, y_t, is de-trended by regressing Equation 10 and obtaining the values of the series ε_t by subtracting the estimated values of y_t from the observed values in the time series Enders (2010:191).

2.2.2 *Differencing with integers*

The first difference of a process refers to the change within the process from one time period to the next (Burke *et al.*, 2005:22). The first difference of a time series x_t is, therefore, $x_t - x_{t-1}$ and is denoted by Δx_t. It then follows that $\Delta x_t = \varepsilon_t$ with ε_t as a white noise process and Δx_t, therefore, being a stationary process (Burke *et al.*, 2005:22). It is also possible to take the second difference of a time series, i.e. $\Delta^2 x_t = \Delta\varepsilon_t$. However, if a time series that is *I(1)* is second differenced, the time series might be over-differenced so care should be taken that a time series is only differenced by the minimal number of times needed to render the series stationary (Burke *et al.*, 2005:23). Besides being integrated to the order of one or two, and requiring first or second differencing to achieve stationarity, a time series may also be fractionally integrated. A fractionally integrated time series needs to be fractionally differenced in order to render the time series stationary.

2.2.3 *Fractional differencing*

The order of differencing used during the fractional differencing of a financial time series is determined by calculating the specific series' fractional differencing parameter, d. Several different methods of determining d have emerged in the empirical literature. Authors of these methods include Hurst (1951), Mandelbrot (1972), Davies and Hart (1987), Hosking (1981), Geweke and Porter-Hudak (1983), Sowell (1990), Lo (1991), Peng, Havlin, Stanley and Goldberger (1994) and Robinson (1995).

The method selected for this paper is that of Geweke and Porter-Hudak (1983), known as the GPH procedure (Maddala *et al.*, 2000:300). The selection of the GPH procedure was based on the following: i) GPH is recognised as one of the more frequently used methods to determine d (Elder & Serletis, 2006:778); ii) the fractional integration parameter is calculated directly;[25] and iii) it is robust and non-sensitive to the short range dependence and variance non-stationarity that are generally found in financial time series data (Fang, Kon & Lai, 1994:170). The fractional integration parameter has been used for fractional differencing, determining stationarity and also detecting the presence of long memory.

[25] Unlike the MRS method, which calculates the Hurst exponent that can be used to indirectly calculate *d* (Hosking, 1981:167).

The GPH method estimates d by running the following ordinary least square regression (Maddala *et al.*, 2000:300):

$$\ln[I(w_j)] = c - d\ln[4\sin^2(w_j/2)] + \eta_j, \tag{11}$$

where $j = 1,...,n$, $w_j = 2\pi j/T$, $n = g(T) < T$ and $I(w_j)$ is the periodogram of X at frequency w_j that can be defined as:

$$I(w) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} e^{itw}(X_t - \bar{X})\right|^2. \tag{12}$$

According to Maddala *et al.* (2000:301), the value of n should be set at $n = \sqrt{T}$ and the estimated variance of d should be computed using the known variance of $\eta_j, \pi^2/6$. The GPH method is also used to determine the presence of a unit root with the null-hypothesis, $H_0 : d = 0$ and the alternative, $H_1 : d \neq 0$. If the null-hypothesis is rejected, the time series contains a unit root and is non-stationary, and if the null-hypothesis cannot be rejected, the time series does not contain a unit root and is stationary (Elder *et al.*, 2006:780).

It is, however, important to correctly determine and interpret the fractional integration parameter of a time series (Gil-Alana, 2006:32). As such, a variety of researchers have provided explanations on how to interpret *d.*[26] The interpretation of d can be summarised as follows: a) if $d = 0$, the time series is covariance stationary; b) if d is between 0 and 0.5, the time series remains covariance stationary; c) if d falls between 0.5 and 1, the time series is no longer considered covariance stationary, but still indicates that the series is mean reverting or has long memory; d) if $d \geq 1$ the series is non-stationary and shows no sign of mean reversion; and e) should $d < 0$ the interpretation of the negative values to the values corresponding to the values above is the same except that the series has been over-differenced (Gil-Alana, 2006:31; Doornik, Hendry, Arellano, Bond, Boswijk & Ooms, 2007:102; Styger, Viljoen & Van Vuuren, 2008:340; Gallegati, 2008:3071).

The discussion above surrounding the interpretation of d and the null-hypothesis of the GPH method has led to the conclusion that a d value closer to zero indicates a larger degree of stationarity in a time series (Gil-Alano, 2006:30). The latter is of importance when

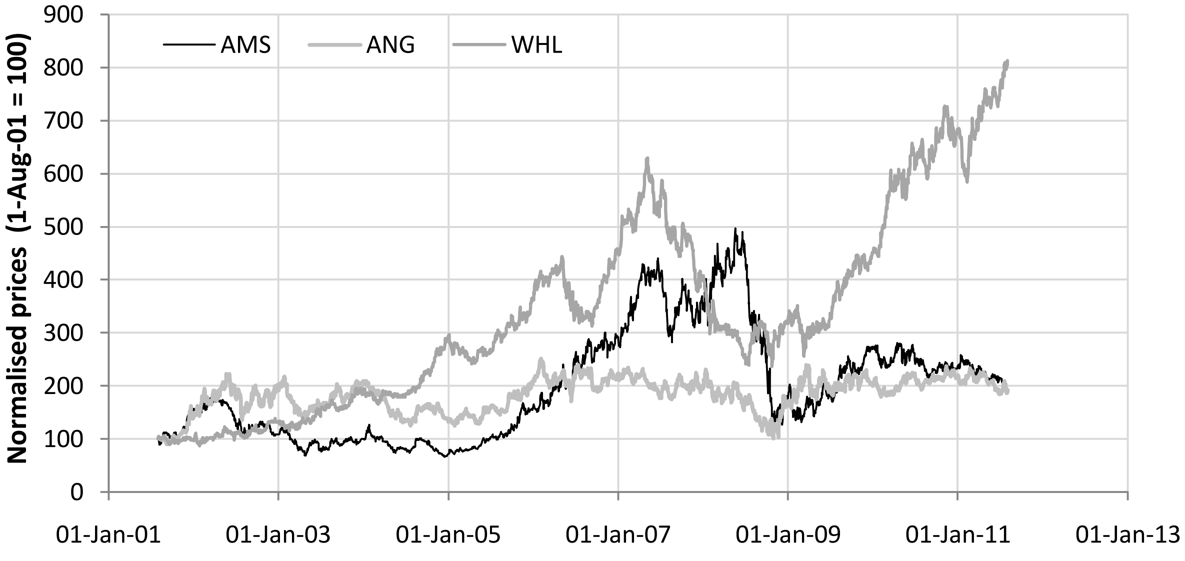
[26] See Doornik *et al.* (2007:102), Styger *et al.* (2008:340), and Gallegati (2008:3071).

the method used is explained and the results of the empirical analysis are reported. Before discussing the method used in this paper the data are discussed next.

3. Data

The data used in this study consist of the daily closing prices of the following stocks listed on the JSE: 1) Anglo Platinum Ltd. (AMS); 2) AngloGold Ashanti Ltd. (ANG); and 3) Woolworths Holdings Ltd. (WHL). The above-mentioned stocks were chosen to ensure that a variety of market capitalisation (large, medium and small) and sectors (AMS: Platinum and precious metals; ANG: Gold mining; WHL: Retail) are included in the study. Another factor playing a role in selecting these time series was the visual difference in the graphs of the three time series. The complete samples of the three stock price time series are presented in Figure 1, normalised to 100 at the start of the series. The differences observed in terms of possible trends and stationary fluctuations make these time series ideal for comparing different forms of stationarity.

Figure 1: Complete samples of AMS, AMG and WHL's normalised closing prices.



Source: McGregor BFA.

Each of the stock prices will be represented by four different sample periods. In order to distinguish between these sample periods, each will be labelled by its listing code followed by the numbers 1 to 4. The complete sample (period 1) is the same for all stocks and stretches from 2 August 2001 to 2 August 2011 to include a total of 2500 observations per stock. The three remaining sample periods differ between the stocks and are summarised, with period 1, in Table 1.

The process of determining the length for each sample period for each stock was conducted as follows: i) the length of sample period 2 was chosen from the start date of the complete sample up to the highest closing price before the 2007/08 financial crisis; ii) the next trading day after the high and up to the low before closing prices started to increase determined the third sample period, and iii) what remained up to the end of the complete sample was included as period 4.

Table 1: Summary of the sample periods.

Period	AMS	ANG	WHL
1	2001/08/02 to 2011/08/02 (2500)	2001/08/02 to 2011/08/02 (2500)	2001/08/02 to 2011/08/02 (2500)
2	2001/08/02 to 2008/05/19 (1967)	2001/08/02 to 2008/01/11 (1611)	2001/08/02 to 2007/05/07 (1438)
3	2008/05/20 to 2008/10/27 (113)	2008/01/14 to 2008/10/28 (200)	2007/05/08 to 2008/10/24 (371)
4	2008/10/28 to 2011/08/02 (690)	2008/10/29 to 2011/08/02 (689)	2008/10/08 to 2011/08/02 (691)

Note: *Number of observations in parenthesis.*

The total number of time series samples tested for different forms of stationarity was 12. Each of these time series differing in number of observations, time periods and statistical properties that make for a comprehensive empirical study of stationarity in financial time series.[27]

4. Method

The process of transforming a non-stationary financial time series into a stationary financial time series is more extensive than simply taking first differences, testing stationarity with a unit root test and continuing with research. There are a variety of tests with differing null-hypotheses that can be used to test the stationarity of a time series. The tests used in this paper include the ADF (Section 2.1.1), KPSS (Section 2.1.2) and DF-GLS (Section 2.1.3) tests.

In order to ensure that the data are rendered stationary, a novel process is used in this paper and is illustrated in Figure A.1 of the appendix. The process commences with the data in their raw form, testing whether the financial time series are stationary using the ADF and

[27] The inclusion of time series with different numbers of observations, time periods and stock prices was done with no significant role being played by the financial crisis. The choices were made in order to select twelve time series that differ significantly from each other.

KPSS tests. The time series are tested for stationarity using the three different ADF test equations and the two different KPSS test equations.[28] When estimatingthe ADF test, while not including an intercept or trend, one of the following may occur. The null-hypothesis is rejected and the series deemed stationary in levels around a zero mean. In this case, the data are strictly stationary and modelling may continue without any transformations.[29] Alternatively, the null-hypothesis cannot be rejected, indicating the presence of a unit root. In this case it is necessary to continue using the second ADF regression equation and the first KPSS regression equation.

When conducting the ADF and KPSS tests, while including an intercept, the following results may occur. The null-hypothesis is rejected in the ADF test and not rejected in the KPSS test. If so, the data are deemed stationary around a constant mean and modelling may continue. Alternatively, the null-hypothesis cannot be rejected in the ADF test and is rejected in the KPSS test, indicating the presence of a unit root. In the latter case, it is necessary to continue using the third ADF regression equation and the second KPSS regression equation. When conducting the ADF and KPSS tests with an intercept and trend in the regression equations, one of the following results may occur. The null-hypothesis is rejected in the ADF test and not rejected in KPSS test. If so, the data are stationary around a constant mean and a deterministic trend. This last result should be confirmed with the DF-GLS test while including an intercept term in the regression equation. The DF-GLS test offers two regression equations; one including an intercept term and the other including an intercept term and a trend.[30] The inclusion of an intercept term in the test equation, allows testing whether the time series is stationary around a constant mean (μ) with $\mu \neq 0$. Test results are, therefore, obtained for a de-trended time series. Rejecting the null-hypothesis while including an intercept term indicates that the particular time series is trend-stationary and confirmation with ADF and KPSS occurs. Before continuing with modelling, it is necessary to de-trend the time series and to repeat the ADF and KPSS tests while including an intercept term. If the null-hypothesis of the ADF test is rejected and the null-hypothesis of the KPSS test cannot be rejected, the time series is trend-stationary. Alternatively, the null-hypothesis cannot be

[28] ADF allows the inclusion of an intercept term, intercept and trend terms or none. KPSS does not allow the third situation.
[29] This is highly unlikely when using financial time series data.
[30] Note that the DF-GLS test regression uses a de-trended dependent variable.

rejected in the ADF test and is rejected in the KPSS test, indicating that non-stationarity persists.

It is also possible to include an intercept and trend in the DF-GLS regression. Rejecting the null-hypothesis in this instance indicates a trend stationary time series that will require the data to be de-trended first. In the case of not rejecting the null-hypothesis, econometricians and financial analysts may move to differencing methods. Therefore, only after following the whole process up to this point, may the econometrician and financial analyst start with the correct form of differencing.

Data can be differenced in two ways with the more popular of the two approaches being that of taking first differences of the data. The other approach would be to fractionally difference the data. Regardless of the data being tested, it is necessary to continue with the process of testing for stationarity by conducting both differencing methods and comparing the fractional integration parameters of both. Comparing the fractional differencing parameters will allow the econometrician and financial analyst to conclude whether the time series should be used in a first differenced stationary form or fractional difference stationary form.[31]

The methods discussed above are also used forfractional difference stationarity. Before starting with the process (conducted for each time series individually), first differences and fractional differences must be taken of the same time series. The method used to determine the fractional differencing parameter is the GPH method. In this case, d is used to obtain a fractionally differenced time series before proceeding with the process.

When conducting the ADF test (without an intercept or trend), the following results may occur. The null-hypothesis is rejected and the series deemed first difference/fractional difference stationary around a zero mean and modelling may continue without any further transformations. Alternatively, the null-hypothesis cannot be rejected, indicating the presence of a unit root. In this case, it is necessary to continue using the second ADF regression equation and the first KPSS regression equation. When conducting the ADF and KPSS tests, while including an intercept in the regression equations, one of the following results may occur. The null-hypothesis is rejected in the ADF test and not rejected in KPSS test, and the

[31] Refer to Section 2.2.3 for guidelines on interpreting the fractional integration parameter.

series is deemed first difference/fractional difference stationary around a constant mean. Before continuing with the analysis, the calculation of the fractional differencing parameter is required. By comparing the fractional differencing parameters of the two stationary time series it can be determined which time series should be used in further modelling.[32] Alternatively, the null-hypothesis cannot be rejected in the ADF test and is rejected in the KPSS test, indicating the presence of a unit root. In this case, it is necessary to continue using the third ADF regression equation and the second KPSS regression equation. When conducting the ADF and KPSS tests, while including an intercept and trend in the regression equations, one of the following results may occur. The null-hypothesis is rejected while including an intercept and trend after the time series has been first differenced/fractionally differenced. This indicates that the time series contains a trend. In the unlikely event of this occurring, the time series has to be de-trended, re-tested and the results confirmed with the ADF and KPSS tests.

The result of the process discussed above is that econometricians and financial analysts would now be in a position to report whether the time series under consideration are in fact trend stationary, first difference stationary or fractional difference stationary. Furthermore, analyses to continue after this process could yield more accurate and statistically significant results. The next section includes the results after applying the above method to three different stocks listed on the JSE.

5. Empirical results

The empirical results of the process of testing for stationarity are presented as follows: Table 2 presents the results for AMS1, AMS2, AMS3 and AMS4; Table 3 reports the results for ANG1, ANG2, ANG3 and ANG4; and Table 4 shows the results for WHL1, WHL2, WHL3 and WHL4.

The results in Table 2 indicate the following: AMS1 is non-stationary in levels, but stationary in first- and fractional difference form. The results are confirmed by the ADF and KPSS test results. It is, therefore, necessary to compare the fractional integration parameter of the first and fractional difference time series. The d of the fractional difference AMS1 series is closer to zero and it can be concluded that AMS1 is fractional difference stationary (Gil-Alana, 2006:31; Doornik, Hendry, Arellano, Bond, Boswijk & Ooms, 2007:102; Styger, Viljoen

[32] Refer to Section 2.2.3 for the interpretation of the *d* parameter.

& van Vuuren, 2008:340; Gallegati, 2008:3071). AMS2 is non-stationary in levels, but stationary in first and fractional difference form. The results are confirmed by the ADF and KPSS test results. The d parameter of the first difference time series is negative and, therefore, the time series is stationary, but over-differenced. The conclusion is, therefore, that AMS2 is fractional difference stationary.

Some unexpected results were obtained for AMS3. Firstly, the ADF test without a trend or intercept indicates stationarity in levels form at a 5% level of significance.[33] Secondly, KPSS test results indicate that AMS3 is trend-stationary, but this could not be confirmed with the parallel ADF test. The confirmed conclusion is, however, that the fractional difference time series is over-differenced, trend stationarity could not be confirmed by ADF and KPSS resulting in AMS3 being first difference stationary.

AMS4 gave conflicting results by illustrating that fractionally differenced stationarity could not be confirmed after including an intercept term in the ADF and KPSS tests. First differenced stationarity, however, was confirmed after including an intercept term resulting in AMS4 being over-differenced (due to a negative *d* parameter), but first difference stationary due to the absence of better results (no other confirmation). The results of ANG are discussed next.

The results for the four ANG time series, as reported in Table 3 can be concluded as follow: The complete sample period ANG1 was found to be trend-stationary and, therefore, needed to be de-trended to render the series stationary. This trend stationarity was confirmed by the results of the ADF and KPSS test equations that included an intercept and trend term.

The ANG2 time series results indicated that the first and fractional difference time series were stationary. After comparing the fractional integration parameter of the two time series it was found that the first differenced series was over-differenced (due to a negative *d* parameter). In conclusion, ANG2 was, therefore, found to be fractional difference stationary. The results obtained from tests conducted on ANG3 are similar to that of ANG2, and ANG4 was found to be stationary in levels form in two of the ADF test equations, but these results could not be confirmed by either the DF-GLS or KPSS tests. It was, therefore, necessary to

[33] Conclusions are based only on results with a 1% level of significance.

move on to differencing and conclude that ANG4 is fractional difference stationary and over-differenced (due to a negative *d* parameter) in first differenced form.

Table 2: Empirical results for all four AMS time series.

	DF-GLS		ADF			KPSS		GPH - *I(d)*
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 108**
AMS1	-1.011	-1.678	-0.497	-1.653	-1.605	2.807	0.580	1.097
(level)	(-2.566)	(-3.48)	(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	[0]
			[0.501]	[0.4553]	[0.7911]			
AMS1			-36.534	-36.529	-36.529	0.127	0.090	0.103
(first diff.)			(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	[0.157]
			[0]	[0]	[0]			
AMS1			-26.825	-26.819	-26.817	0.112	0.047	0.005
(frac. diff.)			(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	[0.938]
{1.097}			[0]	[0]	[0]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 85**
AMS2	1.654	-0.805	1.908	0.834	-0.836	3.559	1.187	0.956
(level)	(-2.566)	-3.480	(-2.566)	(-3.434)	(-3.963)	(0.739)	(0.216)	[0]
			[0.986]	[0.9946]	[0.9609]			
AMS2			-38.259	-38.311	-38.371	0.452	0.051	-0.024
(first diff.)			(-2.566)	(-3.434)	(-3.963)	(0.739)	(0.216)	[0.722]
			[0]	[0]	[0]			
AMS2			-40.290	-40.384	-40.454	0.311	0.134	0.010
(frac. diff.)			(-2.566)	(-3.434)	(-3.963)	(0.739)	(0.216)	[0.895]
{0.956}			[0]	[0]	[0]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 16**
AMS3	1.792	-2.412	-2.423	0.111	-2.344	1.184	0.089	0.996
(level)	(-2.586)	(-3.566)	(-2.586)	(-3.490)	(-4.042)	(0.739)	(0.216)	[0]
			[0.016]	[0.9652]	[0.4067]			
AMS3			-9.686	-10.181	-10.187	0.125	0.072	0.001
(first diff.)			(-2.586)	(-3.490)	(-4.043)	(0.739)	(0.216)	[0.995]
			[0]	[0]	[0]			
AMS3			-36.587	-37.379	-36.875	0.365	0.126	-0.003
(frac. diff.)			(-2.586)	(-3.490)	(-4.042)	(0.739)	(0.216)	[0.973]
{0.996}			[0]	[0.0001]	[0.0001]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 49**
AMS4	-0.708	-1.388	-0.033	-2.533	-2.126	1.573	0.686	0.916
(level)	(-2.566)	(-3.48)	(-2.566)	(-3.440)	(-3.971428)	(0.739)	(0.216)	[0]
			[0.671]	[0.1081]	[0.5296]			
AMS4			-25.846	-25.831	-25.864	0.296	0.024	-0.129
(first diff.)			(-2.566)	(-3.440)	(-3.971448)	(0.739)	(0.216)	[0.151]
			[0]	[0]	[0]			
AMS4			-30.588	-30.611	-30.733	0.806	0.083	0.049
(frac. diff.)			(-2.566)	(-3.440)	(-3.971)	(0.739)	(0.216)	[0.612]
{0.916}			[0]	[0]	[0]			

Note: critical values at 1% in (), p-values in [], test statistics normal and the fractional difference parameter in {}.

The last of the three stocks to be investigated was WHL and the results are reported in Table 4 and indicate the following: WHL1 and WHL2 were stationary in first- and fractional difference form. Comparing the two fractional integration parameters showed that both

WH1L and WHL2 are first difference stationary (*d* parameter closest to zero). The ADF test equation, including no intercept or trend, indicated that WHL3 is stationary in levels form. The latter result is highly unlikely and cannot be confirmed by any other unit root test. Therefore, it became necessary to move on to first and fractional difference methods, which resulted in WHL3 being fractionally differenced stationary (*d* parameter closest to zero), and the final sample investigated was that of WHL4. It was found that in first difference form the time series is over-differenced and that fractional differencing should be used to render the series stationary.

Table 3: Empirical results for all four ANG time series.

	DF-GLS		**ADF**			**KPSS**		
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	
ANG1	-1.096	-2.852	-0.291	-4.038	-4.353	1.801	0.156	
(level)	(-2.566)	(-3.48)	(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	
			[0.581]	[0.0013]	[0.0026]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 82**
ANG2	-0.482	-2.461	0.215	-2.910	-3.343	1.759	0.285	0.861
(level)	(-2.566)	(-3.48)	(-2.566)	(-3.434)	(-3.964)	(0.739)	(0.216)	[0]
			[0.749]	[0.0443]	[0.060]			
ANG2			-40.617	-40.618	-40.607	0.043	0.040	-0.073
(first diff.)			(-2.566)	(-3.434)	(-3.964)	(0.739)	(0.216)	[0.230]
			[0]	[0]	[0]			
ANG2			-39.996	-40.137	-40.168	0.475	0.211	0.153
(frac. diff.)			(-2.566)	(-3.434)	(-3.964)	(0.739)	(0.216)	[0.037]
{0.861}			[0]	[0]	[0]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 23**
ANG3	0.767	-2.375	-1.832	-1.195	-2.692	1.503	0.293	0.812
(level)	(-2.577)	(-3.4612)	(-2.577)	(-3.463)	(-4.005)	(0.739)	(0.216)	[0]
			[0.067]	[0.6766]	[0.241]			
ANG3			-13.309	-13.439	-13.406	0.057	0.053	-0.077
(first diff.)			(-2.577)	(-3.463)	(-4.005)	(0.739)	(0.216)	[0.483]
			[0]	[0]	[0]			
ANG3			-35.354	-35.228	-36.220	0.718	0.155	0.094
(frac. diff.)			(-2.577)	(-3.463)	(-4.005)	(0.739)	(0.216)	[0.294]
{0.812}			[0]	[0.0001]	[0.0001]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 49**
ANG4	-0.457	-1.336	0.181	-4.553	-4.417	1.001	0.225	0.913
(level)	(-2.568)	(-3.48)	(-2.568)	(-3.440)	(-3.972)	(0.739)	(0.216)	[0]
			[0.739]	[0.0002]	[0.002]			
ANG4			-20.556	-20.555	-20.612	0.220	0.047	-0.142
(first diff.)			(-2.568)	(-3.440)	(-3.971)	(0.739)	(0.216)	[0.173]
			[0]	[0]	[0]			
ANG4			-18.271	-18.403	-18.518	0.622	0.111	0.001
(frac. diff.)			(-2.577)	(-3.440)	(-4.005)	(0.739)	(0.216)	[0.991]
{0.913}			[0]	[0]	[0]			

Note: critical values at 1% in (), p-values in [], test statistics normal and the fractional difference parameter in {}.

Table 4: Empirical results for all four WHL time series.

	DF-GLS		ADF			KPSS		GPH - *I(d)*
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 108**
WHL1	1.959	-1.265	2.098	0.458	-1.220	4.613	0.352	0.979
(level)	(-2.566)	(-3.48)	(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	[0]
			[0.992]	[0.985]	[0.9054]			
WHL1			-51.428	-51.515	-51.536	0.233	0.116	0.085
(first diff.)			(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	[0.1765]
			[0.0001]	[0.0001]	[0]			
WHL1			-52.460	-52.577	-52.600	0.189	0.145	0.098
(frac. diff.)			(-2.566)	(-3.433)	(-3.962)	(0.739)	(0.216)	[0.1510]
{0.979}			[0.0001]	[0.0001]	[0]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 77**
WHL2	3.809	0.047	3.753	2.398	0.032	4.355	0.625	0.928
(level)	(-2.567)	(-3.48)	(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0]
			[1]	[1]	[0.9966]			
WHL2			-39.778	-40.022	-40.197	0.735	0.115	0.058
(first diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.4849]
			[0]	[0]	[0]			
WHL2			-41.376	-41.926	-42.151	0.490	0.193	0.117
(frac. diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.1118]
{0.928}			[0]	[0]	[0]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 33**
WHL3	1.344	-1.632	-2.766	-1.620	-2.311	2.253	0.395	0.955
(level)	(-2.572)	(-3.477)	(-2.572)	(-3.448)	(-3.983)	(0.739)	(0.216)	[0]
			[0.006]	[0.471]	[0.4263]			
WHL3			-19.914	-20.217	-20.232	0.131	0.041	0.145
(first diff.)			(-2.572)	(-3.448)	(-3.983)	(0.739)	(0.216)	[0.161]
			[0]	[0]	[0]			
WHL3			-74.052	-74.217	-73.824	0.342	0.155	0.017
(frac. diff.)			(-2.572)	(-3.448)	(-3.983)	(0.739)	(0.216)	[0.670]
{0.955}			[0.0001]	[0.0001]	[0.0001]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**n = 49**
WHL4	1.738	-2.911	2.177	-0.571	-2.909	3.036	0.243	0.965
(level)	(-2.568)	-3.480	(-2.568)	(-3.440)	(-3.971)	(0.739)	(0.216)	[0]
			0.993	0.874	0.160			
WHL4			-26.450	-26.666	-26.647	0.036	0.035	-0.066
(first diff.)			(-2.568)	(-3.440)	(-3.971)	(0.739)	(0.216)	[0.393]
			[0]	[0]	[0]			
WHL4			-37.583	-37.951	-37.880	0.204	0.091	0.067
(frac. diff.)			(-2.568)	(-3.440)	(-3.971)	(0.739)	(0.216)	[0.451]
{0.965}			[0]	[0]	[0]			

Note: critical values at 1% in (), p-values in [], test statistics normal and the fractional difference parameter in {}.

A summary of all the results from the process used on the twelve time series is reported in Table 5. The results show that trend-stationarity was found only once, first difference stationarity was found in four of the time periods, and fractional difference stationarity was found in seven of the time series.

Table 5: Final results: The stationarity of the twelve time series.

	Trend	First difference	Fractional difference
AMS1			X
AMS2		O	X
AMS3		X	O
AMS4		X	
ANG1	X		
ANG2		O	X
ANG3		O	X
ANG4		O	X
WHL1		X	
WHL2		X	
WHL3			X
WHL4		O	X

Note: Stationary time series are indicated by X and over-differenced time series are indicated by O.

Also worth taking note of is the regular occurrence of over-differenced time series after first differences were taken. A total of five time series were over-differenced after first differencing. The results, therefore, confirm that econometricians and financial analysts will have to adapt their approach when dealing with non-stationary data based on the statistical properties of the data and the sample period used. The transformation process to determine the correct method to achieve stationarity is highlighted by the results and also highlights the importance of not simply first differencing data. The risk of over-differencing a time series can be avoided if researchers practise due diligence.

6. Conclusions and recommendations

This paper challenges the assumption that most financial time series are first difference stationary. The common 'difference first, ask questions later' approach ought to be revisited by taking a more systematic approach when analysing the statistical properties of financial time series data. To this end, a systematic process (Appendix: Figure A.1)was developed to determine which method best suited the specific time series in order to render the data stationary. The process therefore focuses on, and is led by, the underlying statistical properties of a particular time series.

After applying our process[34]and methods[35] of reducing a time series to stationarity, the following results were reported: one time series is trend stationary, four of the time series are first difference stationary, a total of five time series were over-differenced after first differencing, and more than half of the time series are fractional difference stationary.

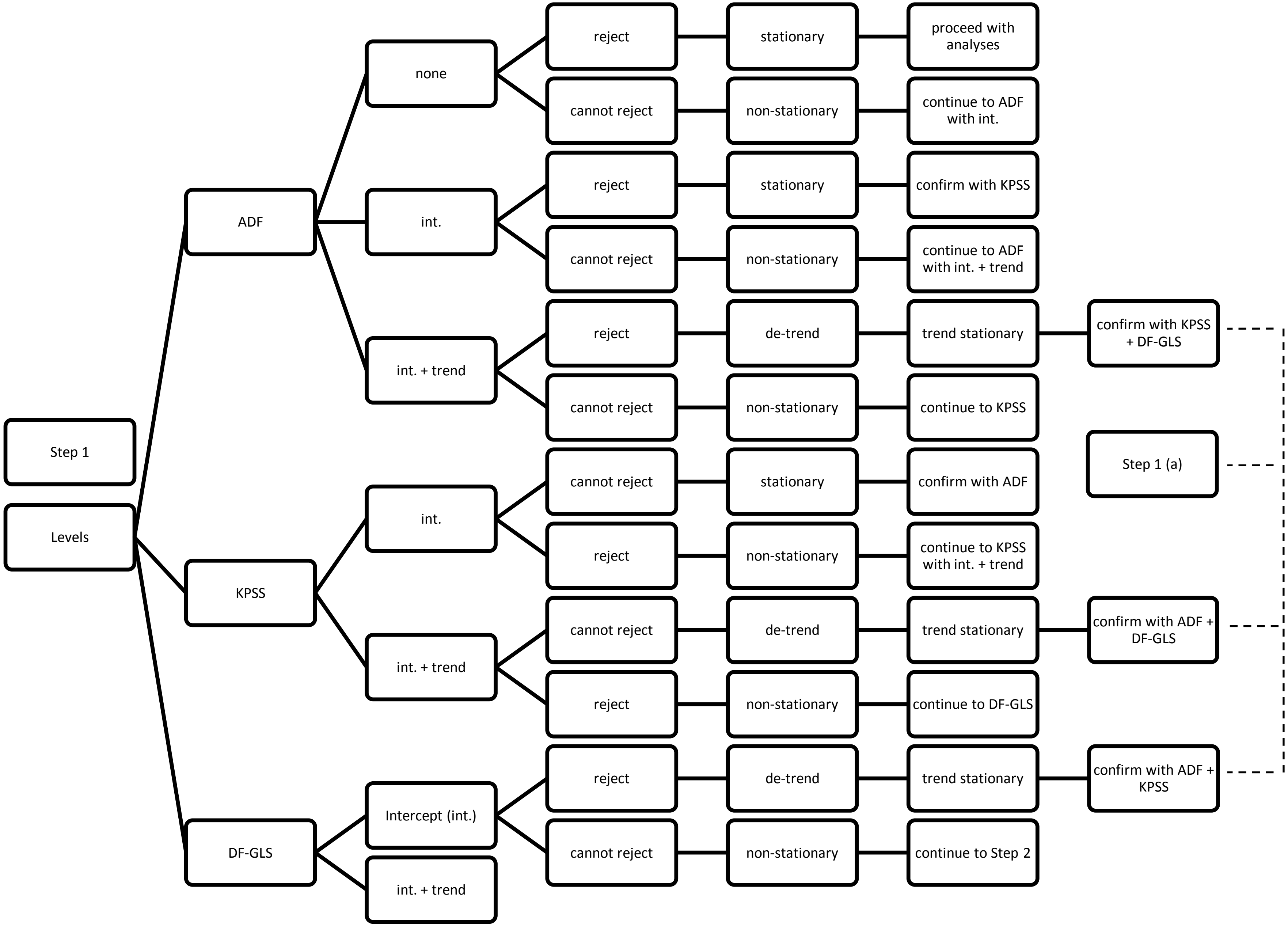
[34] Testing for stationarity by employing DF-GLS, ADF and KPSS.

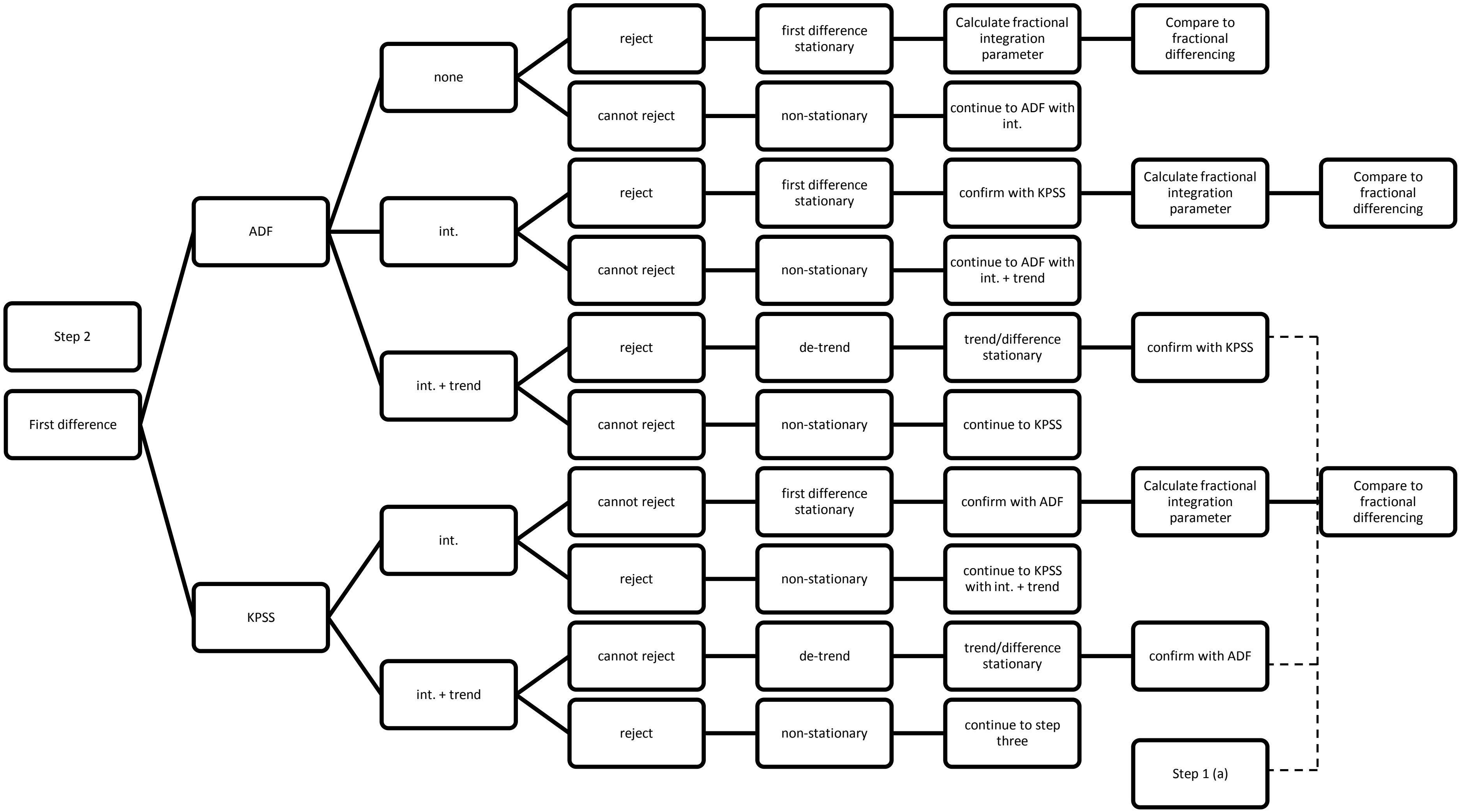
[35] De-trending, first differencing and fractional differencing.

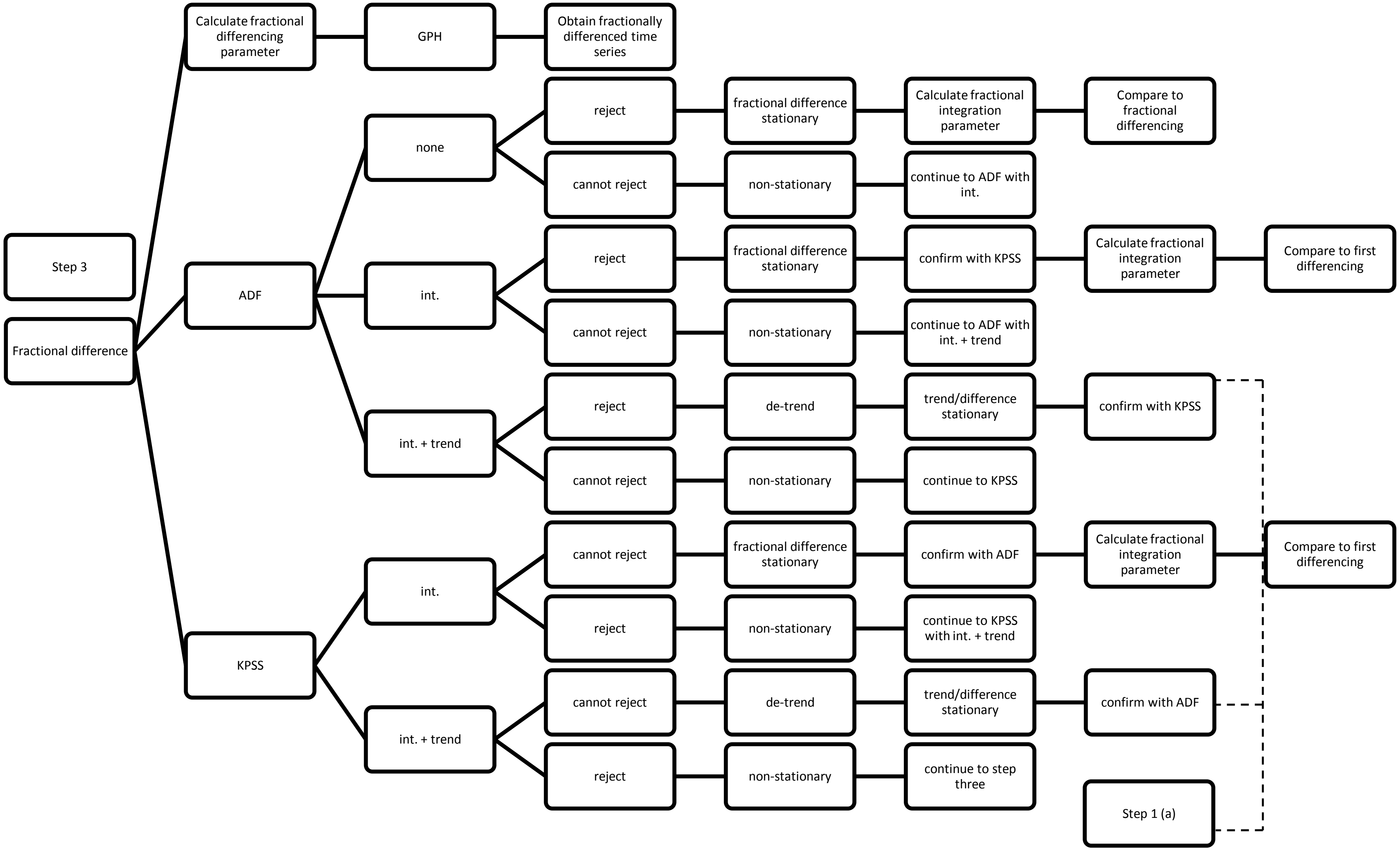
The results confirm that it is imperative that the statistical properties of the specific data series be determined before deciding on the actions to be taken. The assumption of most time series being first difference stationary cannot hold any longer due to the risk of over-differencing a time series. If researchers practise due diligence and allow the statistical properties of a time series to lead them in rendering the series stationary, the occurrence of over-differencing can be avoided and improved empirical results may be achieved.

8. Appendix

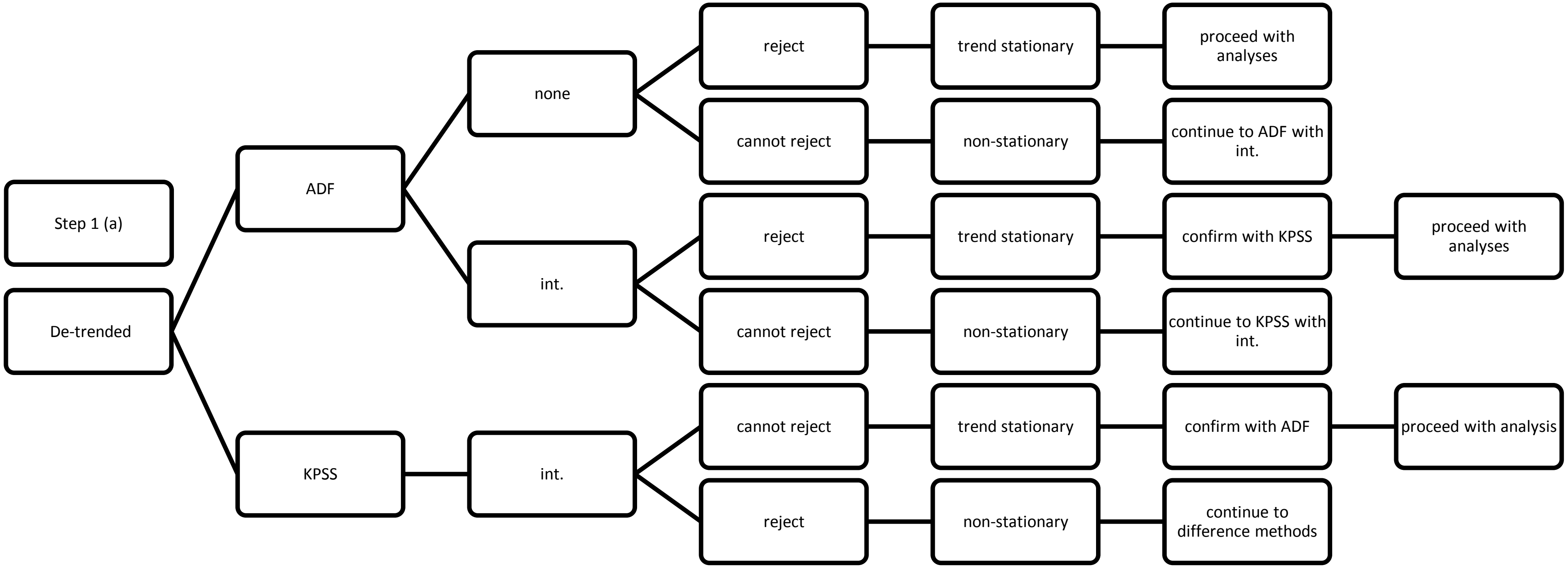
Figure A.1: The process of rendering a non-stationary time series stationary (Step1, Step 1.a, Step 2 and Step 3)




Step 2
First difference
ADF
KPSS
none
int.
int. + trend
int.
int. + trend
reject
cannot reject
reject
cannot reject
reject
cannot reject
cannot reject
reject
cannot reject
reject
first difference stationary
non-stationary
first difference stationary
non-stationary
de-trend
non-stationary
first difference stationary
non-stationary
de-trend
non-stationary
Calculate fractional integration parameter
continue to ADF with int.
confirm with KPSS
continue to ADF with int. + trend
trend/difference stationary
continue to KPSS
confirm with ADF
continue to KPSS with int. + trend
trend/difference stationary
continue to step three
Compare to fractional differencing
Calculate fractional integration parameter
confirm with KPSS
Calculate fractional integration parameter
confirm with ADF
Step 1 (a)
Compare to fractional differencing
Compare to fractional differencing



Step 3
Fractional difference
Calculate fractional differencing parameter
GPH
Obtain fractionally differenced time series
ADF
none
reject
fractional difference stationary
Calculate fractional integration parameter
Compare to fractional differencing
cannot reject
non-stationary
continue to ADF with int.
int.
reject
fractional difference stationary
confirm with KPSS
Calculate fractional integration parameter
Compare to first differencing
cannot reject
non-stationary
continue to ADF with int. + trend
int. + trend
reject
de-trend
trend/difference stationary
confirm with KPSS
cannot reject
non-stationary
continue to KPSS
KPSS
int.
cannot reject
fractional difference stationary
confirm with ADF
Calculate fractional integration parameter
Compare to first differencing
reject
non-stationary
continue to KPSS with int. + trend
int. + trend
cannot reject
de-trend
trend/difference stationary
confirm with ADF
reject
non-stationary
continue to step three
Step 1 (a)


Step 1 (a)
De-trended
ADF
KPSS
none
int.
int.
reject
cannot reject
reject
cannot reject
cannot reject
reject
trend stationary
non-stationary
trend stationary
non-stationary
trend stationary
non-stationary
proceed with analyses
continue to ADF with int.
confirm with KPSS
continue to KPSS with int.
confirm with ADF
continue to difference methods
proceed with analyses
proceed with analysis

7. References

ARELLANO, C. & PENTULA, S.G. 1995. Testing for trend stationarity versus difference stationarity. *Journal of Time Series Analysis*, 16:147-164.

ASTERIOU, D. & HALL, S.G. 2007. Applied econometrics: a modern approach. New York: Palgrave Macmillan. 397 p.

BURKE, S.P. & HUNTER, J. 2005. Modelling non-stationary economic time series: A multivariate approach. New York: Palgrave MacMillan. 253 p.

CHOI, I. 1994. Residual-based tests for the null of stationarity with application to US macroeconomic time series. *Econometric Theory*, 10:720-746.

DAVIES, R.B. & HARTE, D.S. 1987. Tests for the Hurst effect. *Biometrika Trust*, 74(1):95-101. Mar.

DICKEY, D.A. & FULLER, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427-431. Jun.

DICKEY, D.A. & FULLER, W.A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057-172. Jul.

DOORNIK, J.A., HENDRY, D.F., ARELLANO, M., BOND, S., BOSWIJK, H.P. & OOMS, M. 2007. Econometric modelling. PcGive, Vol 3. OxMetrics™ 5. London: Timberlake Consultants.

ELDER, J. & SERLETIS, A. 2006. On fractional integrating dynamics in the U.S. stock market. *Chaos, Solitons and Fractals*. 34(1):777-781.

ELLIOTT, G., ROTHENBERG, T.J. & STOCK, J.H. 1996. Efficient tests for an autoregressive unit root. *Econometrica.* 64(1):813-836.

ENDERS, W. 2010. Applied econometric time series. Hoboken: Wiley. 517 p.

ERFANI, A. & SAMIMI, A.J. 2009. Long memory forecasting of stock price index using a fractionally differenced ARMA model. *Journal of Applied Sciences Research*, 5(10):1721-1731.

FANG, H., KON, S.L. & LAI, M. 1994. Fractal structure in currency futures price dynamics. *The Journal of Futures Markets*. 14(2):169-181.

GALLEGATI, M. 2008. Wavelet analysis of share returns and aggregate economic activity. *Computational Statistics and Data Analysis*. 51(6):3061-3074.

GEWEKE, J. & PORTER-HUDAK, A. 1983. The estimation application of long-memory time series models. *Journal of Time Series Analysis*. 4(1):221-238.

GIL-ALANA, L.A. 2006. Fractional integration in daily stock market indexes. *Review of Financial Economics*. 15:28-48.

HOSKING, J.R.M. 1981. Fractional differencing. *Biometrika Trust*, 68(1):1025-1034.

HURST, H.E. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770-799.

KWAITKOWSKI, D., PHILLIPS, P.C.B., SCHMIDT, P. & SHIN, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159-178.

LEYBOURNE, S.J. & MCCABE, B.P.M. 1994. A consistent test for a unit root. *Journal of Business and Economic Statistics*, 12:157-166.

LEYBOURNE, S.J., MCCABE, B.P.M. & TREMAYNE, A.R. 1996. Can economic time series be differenced to stationarity? *Journal of Business & Economic Statistics*, 14(4):435-446. Oct.

LO, A.W. 1991. Long-term memory in stock market prices. *Econometrica*, 59(5):1279-1313. Sep.

MADDALA, G.S. & KIM, I.M. 2000. Unit roots, cointegration, and structural change. Cambridge: Cambridge. 505 p.

MANDELBROT, B. 1972. Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Statistics*, 23(3):1015-1028.
MCCABE, B.P.M. & TREMAYNE, A.R. 1995. Testing time series for difference stationarity. *Annals of Statistics*, 23(3):1015-1028.
NABEYA, S. & TANAKA, K. 1988. Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Annals of Statistics*, 16:218-235.
PARK, J.Y. 1990. Testing for unit roots and cointegration by variable addition. *Advances in Econometrics*, 8:107-133.
PENG, C.K., HAVLIN, S., STANLEY, H.E. & GOLDBERGER, A.L. 1994. Quantification of scaling exponents and crossover phenomena in non-stationary heartbeat time series. *Chaos*, 5(1):82-87.
ROBINSON, P. 1994. Efficient tests of non-stationary hypotheses. *Journal of the American Statistical Association*, 89(428):1420-1437.
ROBINSON, P. 1995. Gaussian semi-parametric estimation of long range dependence. *Annals of Statistics*. 23(5):1630-1661.
SAIKKONEN, P. & LUUKKONEN, R. 1993. Testing for a moving average unit root in ARIMA models. *Journal of the American Statistical Association*, 88:596-601.
SOWELL, F. 1990. The fractional unit root distribution. *Econometrica*. 58(1):495-505.
STYGER, P., VILJOEN, S. & VAN VUUREN, G. 2008. A triptych on the USD-ZAR exchange rate dynamics. (In the Proceeding of the 5th international conference on applied financial economics, Greece, 03-05 July 2008. p. 334-344).
TANAKA, K. 1990. Testing for a moving average unit root. *Econometric Theory*. 6:433-444.

CHAPTER 4 |PAPER 3| THE APPLICATION OF DIFFERENT FORMS OF STATIONARITY IN FINANCIAL TIME SERIES ANALYSIS

The application of different forms of stationarity in financial time series analysis

Jan van Greunen,[36] André Heymans[37] and Gary van Vuuren[38]

Abstract

Since most time series data are non-stationary, the econometrician is required to make the data stationary before embarking on any econometric analysis in order to avoid spurious results. Although there are several different ways to render a non-stationary time series stationary, few econometricians look past the first differencing and log-differencing methods. In order to determine whether this approach is indeed the correct one, we employ a novel process to test whether using the correct form of stationary data enhances the forecasting ability of models such as ARFIMA. The results corroborate this hypothesis in that all the time series that were found to be fractional difference stationary, outperformed their first difference form and *vice versa*.

JEL Classification: C22, G00, G17.

Keywords: Fractionally differenced stationarity, first differenced stationarity, financial time series analysis.

1. Introduction

When working with time series data, it is important to ensure that the data is stationary before proceeding with any type of analysis (Granger & Newbold, 1974:112; Gujurati, 2006:493). However, very few, if indeed any econometricians, emphasise the importance of the form of stationarity needed for a specific test. Most econometricians simply employ the first difference approach, mainly as a result of Nelson and Plosser's (1982) work in which they argued that many macroeconomic time series are difference stationary and not trend stationary. As a result, the popularity of the first difference approach is widespread with countless authors employing the first difference approach (see for example McCabe & Tre-

[36] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: jan.vangreunen@nwu.ac.za.

[37] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: andre.heymans@nwu.ac.za.

[38] School of Economics, North-West University, Potchefstroom Campus, Private bag X6001, Potchefstroom, 2520, South Africa. Email: gary.vanvuuren@nwu.ac.za.

mayne, 1995:1015;[39] Leybourne, McCabe & Tremayne, 1996:45;[40] Burke & Hunter, 2005:22).[41] Besides the first difference approach, another popular method to transform non-stationary data is to log the differences of the time series. The log-difference approach has become the main form of transforming non-stationary financial time series into a stationary returns time series when conducting research on, for example, stock prices and indices. This is evident from papers such as that of Diebold and Yilmaz (2008:4),[42] Lien and Yang (2009:142)[43] and Joshi (2011:2).[44] In some cases, first differencing and log differencing are even used by economists and financial analysts as alternatives to fractional differencing due to the difficulties associated with the latter method (Erfani & Samimi, 2009:1721).

There is, however, a disadvantage to simply assuming that all time series data are first difference stationary. By simply employing the first difference approach, data are often over-differenced and the inherent properties of the data lost (Plosser & Schwert, 1977; De Jong & Whiteman, 1993). The importance of first determining the correct order of integration lies in the fact that data may be reduced to stationarity without the disadvantage of over-differencing, therefore, retaining properties essential to time series modelling. By retaining some of these properties and using such time series in modelling, may enhance the results of an empirical study.

The aim of this paper is, therefore, to test the validity of this 'difference first, ask questions later' approach, by employing a rigorous process in order to determine the form of stationarity in our time series data. By doing this, we are able to discern between those time series that are trend-stationary, those that are fractional difference stationary and those that are in fact first difference stationary. Thereafter, we employ forecasting models while incorporating the different forms of stationarity found in the data.[45] Finally, we compare the results

[39] "The cornerstone of practical time series modelling is the acceptability of the difference stationary assumption".
[40] "Much of modern applied econometric analysis is predicated on the assumption that data series concerned are non-stationary and that … they can be differenced to achieve stationarity".
[41] "... taking the first difference of the non-stationary process has reduced it to stationarity".
[42] "We calculate returns as the change in log price..."
[43] "The futures price series is calculated as a logarithm ... return series for each exchange is calculated as the difference between current futures price and previous futures price."
[44] "Daily returns are identified as the first difference in the natural logarithm of the closing index value..."
[45] Take note that the use of a forecasting model has no particular significance and that any other financial time series model with stationarity as prerequisite could have been used.

found from using different forms of stationary time series in order to determine if there is any significant difference between these results.

In order to achieve the aim, the rest of this paper will be structured as follows. Section 2 will revisit the reason behind the advantages of using stationary time series data and why it is important not to use over-differenced data. The different ways in which a non-stationary time series is rendered stationary are reported in Section 3. The data will be discussed in Section 4 and the method in Section 5. The results from using different forms of stationarity in forecasting are presented in Section 6, and Section 7 will conclude with recommendations.

2. Revisiting the importance of stationarity and of not over-differencing

Granger and Newbold (1974:112) refer to results obtained from regressions that contain non-stationary data as "nonsense" or "spurious" regression results. Spurious regressions is a name given to regressions that were conducted using time series variables that provided very high R^2 estimates and very low Durbin-Watson statistics. The results of the regression may, therefore, very well appear excellent, but are in actual fact useless (Gujurati, 2006:493). The reason for such spurious results is the concept known as non-stationarity. The most common example used in finance to refer to a non-stationary time series is the random walk, since most financial time series are said to follow a random walk (Granger and Newbold 1974:113). The random walk is based on the idea that, for example, the stock price that is observed today is a function of yesterday's stock price plus a random disturbance.[46] A random walk can be represented by the following equation (Granger and Newbold 1974:113):

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

where ε_t is a stationary white noise process. To obtain a stationary time series from Equation 1, it is necessary to take first differences as follows:

$$\Delta x_t = x_t - x_{t-1} = \varepsilon_t. \tag{2}$$

An alternative random walk model also includes a drift and can be presented as:

[46] For an excellent explanation of the random walk model, stationarity, error correction and co-integration, see Murray (1994). This paper will only discuss the essence of the random walk model.

$$x_t = D + x_{t-1} + \varepsilon_t, \tag{3}$$

where D is a constant and the drift parameter. The difference between the two random walk models is that in Equation 1, only the variance changes over time and in Equation 3, the mean and variance change over time (Murray, 1994:37). A non-stationary time series is, therefore, a time series that does not have a constant mean, variance and covariance over time. To conclude the discussion about using non-stationary time series in regression analysis, consider the following two threats posed by non-stationary variables used in regression analysis: 1) When including two non-stationary variables that are not co-integrated in any manner in a regression analysis, the following can occur. The standard errors tend to be large in relation to the coefficients resulting in the rejection of the null-hypothesis more often than what is expected; 2) When running regressions that include two integrated but non-stationary variables, the model will be mis-specified (Granger *et al.*, 1974; Murray, 1994:38). Reducing a non-stationary time series into a stationary time series is, therefore, a central issue in time series analysis.

Before moving to a discussion of why time series should not be over-differenced, it is necessary to formally define a stationary time series. There are two different types of stationarity. Firstly, a time series is considered strictly stationary if the data properties remain unaffected by a shift in the time origin (Maddala & Kim, 2000:10; Montgomery, Jennings & Kulahci, 2008:25). Secondly, a time series is covariance stationary when it has a finite mean and variance (Enders, 2010:54). The next important occurrence to consider is that of over-differencing.

According to Burke and Hunter (2005:23), over-differencing occurs when a non-stationary time series is differenced more times than its order of integration. An example would be to take second differences of a time series that is integrated to the order of one, *I(1)*. The second difference of a time series is taken as follows:

$$\Delta^2 x_t = \varepsilon_t - \varepsilon_{t-1}, \tag{4}$$

and is also a stationary process, but has been over-differenced. It is, therefore, imperative to first discern what the order of integration is before launching into any form of differencing. In order to do so, the econometrician will have to follow a systematic process of determining the statistical properties of the data. The next section will focus on the various methods available to the econometrician to transform non-stationary time series into stationary time series.

3. Achieving stationarity

There are various processes the econometrician can follow in order to render a non-stationary time series stationary. These processes include, among others, de-trending (Section 3.1), differencing with integers (Section 3.2) and fractional differencing (Section 3.3). Since these are the most popular, this paper will highlight the application of these processes next.

3.1 Removing a deterministic trend (de-trending)

Enders (2010:191) explains the process of removing a deterministic trend from a time series by referring to the following representation of a time series:

$$y_t = y_0 + a_1 t + \varepsilon_t, \qquad (5)$$

3.2 where y_0 refers to the initial condition for period zero, $a_1 t$ is the deterministic trend component and ε_t is the pure noise component. The time series, y_t, is de-trended by regressing Equation 5 using an OLS regression and obtaining the values of the series ε_t. The ε_t time series is then obtained by subtracting the estimated values of y_t, obtained from regressing Equation 5,from the original values of y_t (Enders, 2010:191). *Differencing*

Mathematically, the first difference of a time series can be written as $\Delta x_t = x_t - x_{t-1} = \varepsilon_t$, where ε_t is the stationary process obtained from taking first differences. If a time series is rendered stationary after first differences are taken, it is said to be integrated to the order of one, *I(1)*. If a time series needs to be differenced a second time, the process is said to be *I(2)*. However, care should be taken to ensure that an *I(1)* process is not differenced more than once, as this will lead to over-differencing (Burke *et al.*, 2005:23). A time series may also be fractionally integrated, which will require fractional differencing to render the series stationary.

3.3 Fractional differencing

The order of differencing used during the fractional differencing of a financial time series is determined by calculating the specific series' fractional differencing parameter, *d*. Several different methods of determining d have emerged in the empirical literature. Examples of these methods include those employed by Hurst (1951), Mandelbrot (1972), Davies and Hart (1987), Hosking (1981), Geweke and Porter-Hudak (1983), Sowell (1990), Lo (1991), Peng, Havlin, Stanley and Goldberger (1994), and Robinson (1995).

For the purpose of this paper, the Geweke and Porter-Hudak's (1983) (GPH) procedure was used. The selection of the GPH procedure was based on the following: i) GPH is recognised as one of the more frequently used methods to determine d (Elder & Serletis, 2006:778); ii) the fractional integration parameter is calculated directly;[47] and iii) it is robust and non-sensitive to the short-range dependence and variance non-stationarity that are generally found in financial time series data (Fang, Kon & Lai, 1994:170). Once the *d* parameter has been determined, it can be used for fractional differencing, determining stationarity as well as detecting the presence of long memory.

The d parameter can be estimated with the GPH method by running the following ordinary least squares regression:

$$\ln[I(w_j)] = c - d\ln[4\sin^2(w_j/2)] + \eta_j, \quad (6)$$

where $j = 1,...,n$ and n is given as $n = g(T) < T$, $w_j = 2\pi j/T$ and $I(w_j)$ is the periodogram of X at frequency w_j which can be defined as:

$$I(w) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} e^{itw}(X_t - \bar{X})\right|^2. \quad (7)$$

According to Maddala *et al.* (2000:301), the value of n should be set at $n = \sqrt{T}$ in order to estimate the variance of d by using the known variance of $\eta_j, \pi^2/6$. The GPH method can also be used to determine the presence of a unit root. The null-hypothesis of the GPH method is that $d = 0$ and the alternative is that $d \neq 0$. If the null-hypothesis is rejected, the

[47] Unlike the MRS method, which calculates the Hurst exponent that can be used to indirectly calculate *d* (Hosking, 1981:167).

time series contains a unit root and is non-stationary and if the null-hypothesis cannot be rejected the time series is stationary (Elder *et al.*, 2006:780).

When the d parameter is used to determine the type of stationarity, its interpretation can be summarised as follows: a) if d = 0, the time series is covariance stationary; b) if d is between 0 and 0.5, the time series remains covariance stationary; c) if d falls between 0.5 and 1, the time series is no longer considered covariance stationary, but still indicates that the series is mean reverting or has long memory; d) if d ≥ 1, the series is non-stationary and shows no sign of mean reversion; and e) should the value of d be negative, the interpretation of the negative values to the values corresponding to the values above is the same except that the series has been over-differenced (Gil-Alana, 2006:31; Doornik, Hendry, Arellano, Bond, Boswijk & Ooms, 2007:102; Styger, Viljoen & Van Vuuren, 2008:340; Gallegati, 2008:3071). It is, therefore, clear that a d value closer to zero indicates a larger degree of stationarity in a time series (Gil-Alano, 2006:30).

4. Data

The data used are the time series of the following financial and economic variables: i) the daily closing prices of ABSA Bank Ltd. (ASA); ii) the daily closing prices of Nedbank Ltd. (NED); iii) the daily closing values of the JSE all share index (ALSI); iv) the daily values for the South African Consumer Price Index (CPI); and v) the daily realised South African Rand/United States Dollar (ZAR/USD) exchange rate. The sample period chosen for each time series stretches from 9 October 2006 to 7 October 2011, and includes 1253 observations per variable. These time series were selected randomly since the only important factor being that the time series is financial or economic of nature.

5. Method

Contrary to the popular practice, the process of transforming a non-stationary financial time series into a stationary financial time series is more extensive than simply taking first differences and testing for stationarity with a unit root test, before continuing with time series analysis. In order to ensure that the data are truly rendered stationary without over-differencing the data, a novel process is employed.[48] The process commences with the data in its raw form and testing whether the financial time series is stationary using the ADF and

[48] See Figure A.1 in the appendix.

KPSS tests. The time series are tested for stationarity using the three different ADF test equations and the two different KPSS test equations.[49] When estimating the ADF test, while not including an intercept or trend, one of the following may occur. The null-hypothesis could be rejected, in which case the data are deemed stationary in levels around a zero mean. Should this be the case, the data are considered covariance stationary, enabling the analyst to continue modelling without making any further transformations.[50] Alternatively, the null-hypothesis cannot be rejected, indicating the presence of a unit root. In this case, it is necessary to continue using the second ADF regression equation and the first KPSS regression equation.

When conducting the ADF and KPSS tests, while including an intercept, the following results may occur. The null-hypothesis is rejected in the ADF test and not-rejected in the KPSS test. If so, the data are deemed stationary around a constant mean and modelling may continue without any further transformations. Alternatively, the null-hypothesis cannot be rejected in the ADF test and is rejected in the KPSS test, indicating the presence of a unit root. In this case, it is necessary to continue using the third ADF regression equation and the second KPSS regression equation. When conducting the ADF and KPSS tests with an intercept and trend in the regression equations, one of the following results may occur. The null-hypothesis is rejected in the ADF test and not rejected in KPSS test. If so, the data are stationary around a constant mean and a deterministic trend. If this is the case, it is necessary to use the first regression equation of the DF-GLS test. The DF-GLS test offers two regression equations; one including an intercept term and the other including an intercept term and a trend.[51] The inclusion of an intercept term in the test equation indicates whether the time series is stationary around a constant mean (μ) with $\mu \neq 0$. Test results are, therefore, obtained for a de-trended time series. Rejecting the null-hypothesis, while including an intercept term, indicates that the particular time series is trend-stationary. The ADF, KPSS and DF-GLS tests, therefore, confirm that the time series is trend stationary. Before continuing with modelling, it is necessary to de-trend the time series and to repeat the ADF and KPSS tests, while including an intercept term. If the null-hypothesis of the ADF test is rejected and

[49] ADF allows the inclusion of an intercept term, intercept and trend terms or none. KPSS does not allow the third situation.

[50] This is highly unlikely when using financial time series data.

[51] Note that the DF-GLS test regression uses a de-trended dependent variable.

the null-hypothesis of the KPSS test cannot be rejected, the time series is trend-stationary. Alternatively, if the null-hypothesis cannot be rejected in the ADF test but is rejected in the KPSS test, it indicates that non-stationarity persists.

Lastly, an intercept and trend may be included in the regression of the DF-GLS test. Rejecting the null-hypothesis in this instance indicates a trend stationary time series that will require the data to be de-trended first. Should the null-hypothesis not be rejected, the econometrician may move to differencing methods. Therefore, only after following the whole process up to this point, may the econometrician start with the correct form of differencing.

Data can be differenced in two ways with the more popular of the two approaches being that of taking first differences of the data. The second approach would be to fractionally difference the data. Regardless of the data being tested, it is necessary to continue with the process of testing for the type of stationarity by conducting both differencing methods and comparing the fractional differencing parameters of both. Comparing the fractional integration parameters will allow the econometrician to conclude whether the time series should be used in a first difference stationary form or fractional difference stationary form. [52]

The methods discussed above are also used forfractional difference stationarity.Before starting with the process (conducted for each time series individually), first differences and fractional differences must be taken of the same time series. The method used to determine the fractional differencing parameter is the GPH method. In this case, d is used to obtain a fractionally differenced time series before proceeding with the process.

When conducting the ADF test (without an intercept or trend), the following results may occur. The null-hypothesis is rejected and the series deemed first difference/fractional difference stationary around a zero mean and modelling may continue without any further transformations. Alternatively, the null-hypothesis cannot be rejected, indicating the presence of a unit root. In this case, it is necessary to continue using the second ADF regression equation and the first KPSS regression equation. When conducting the ADF and KPSS tests, while including an intercept in the regression equations, one of the following results may occur. The null-hypothesis is rejected in the ADF test and not rejected in KPSS test, there-

[52] Refer to Section 2.2.3 for guidelines on interpreting the fractional integration parameter.

fore indicating that the series is first difference/fractional difference stationary around a constant mean. Before continuing with the analysis, the calculation of the fractional differencing parameter is required. By comparing the fractional differencing parameters of the two stationary time series it can be determined which time series should be used in further modelling.[53] Alternatively, the null-hypothesis cannot be rejected in the ADF test and is rejected in the KPSS test, indicating the presence of a unit root. In this case, it is necessary to continue using the third ADF regression equation and the second KPSS regression equation.

When conducting the ADF and KPSS tests, while including an intercept and trend in the regression equations, one of the following results may occur. The null-hypothesis is rejected while including an intercept and trend after the time series has been first differenced/fractionally differenced. This indicates that the time series contains a trend. In the unlikely event of this occurring, the time series has to be de-trended, re-tested and the results confirmed with the ADF and KPSS tests.

The result of the process discussed above is that the econometrician would now be in a position to report whether the time series under consideration is in fact trend stationary, first difference stationary or fractional difference stationary. Furthermore, analyses to continue after this process could yield more accurate and statistically significant results.

The next step is to determine whether the different forms of stationarity do provide different results after being applied in a time series model. In order to investigate the importance of determining whether a time series is trend-, first or fractional difference stationary and applying it accordingly in modelling, the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model was used (Granger & Joyeux, 1980). The ARFIMA(p,d,q) model for a time series x_t can be written as follows (Doornik & Ooms, 2004:2):

$$\Phi(L)(1-L)^d(x_t-\mu_t)=\Theta(L)\varepsilon_t,\ \ t=1,...,T. \tag{8}$$

The fractional difference operator is represented by $(1-L)^d$. If d is an integer, the model represents an Autoregressive Integrated Moving Average (ARIMA) process, indicating that a stationary ARMA process is obtained after first differencing (Montgomery, *et al.*, 2008:256). Similarly, if fractional differencing is required to achieve a stationary ARMA representation,

[53] Refer to Section 3.3 for the interpretation of the *d* parameter.

it is referred to as an ARFIMA process (Doornik & Ooms, 2004). The autoregressive polynomial and moving average polynomial in the lag operator L of Equation 8 are presented in Equations 9 and 10, respectively, as:

$$\Phi(L) = (1 - \phi_1 L - ... - \phi_p L^p), \tag{9}$$

$$\Theta(L) = (1 + \theta_1 L + ... + \theta_q L^q), \tag{10}$$

where p and q are integers. The orders of p and q are determined using the auto-correlation function and partial auto-correlation function of the specific time series, respectively (Montgomery *et al.*, 2008:255; Erfani *et al.*, 2009:1724). After determining the values of p, d and q, ARFIMA is estimated using the Non-linear Least Squares (NLS) method. Along with the NLS estimation, this paper employs the 'naive' prediction method for forecasting (Doornik *et al.*, 2004:5).

The forecasting process commenced with the inclusion of 1203 in sample observations. The values of p, d and q were estimated for the particular time series and ARFIMA was regressed twice; once while using the fractional d parameter and a second time with $d = 1$. Therefore, once with the time series being fractionally differenced and a second time with the time series being first differenced.[54] Thereafter, a one-step-ahead forecast was made for both AFRIMA processes with the naive forecasting method. The result was a one-step-ahead forecast value for a fractionally differenced time series and for a first differenced time series, which could be compared to determine which is more accurate. This process was repeated fifty times (up to 1252 observations), while including an extra observation with every repetition. As a result, a total of fifty forecasted observations were obtained for both a fractionally and first differenced time series. The two forms of stationarity were compared by estimating two OLS regressions. Both regressions included the observed closing prices as dependent variable and then using the first differenced forecasted values and fractionally differenced forecasted values as independent variable. The regression with the highest R^2 indicates the form of stationarity that outperformed the other in forecasting accuracy.[55] The aim is to determine if a stock that is deemed to be fractional difference stationary will outperform itself in first difference form when comparing forecasting abilities. Alternatively,

[54] The second regression is the equivalent of an ARIMA process.

[55] The R^2 of a model indicates the amount of variation in the dependent variable that is explained by the independent variable (Asteriou & Hall, 2007:38).

will a stock that is deemed to be first difference stationary outperform its fractional difference form in forecasting ability?

6. Results

The empirical results of the process discussed in Section 5, regarding the stationarity of financial time series for the two stocks used, are presented in Table 1. Based on the process explained in Section 5, which is also graphically depicted in Figure A.1 of the appendix, and the interpretation of the d parameters, the following can be concluded: ASA is first difference stationary, NED is fractional difference stationary, the ALSI is first difference stationary, the CPI is fractional difference stationary and the ZAR/USD is fractional difference stationary. The next step is to compare the forecasting performance of each in its first and fractional difference form. This will confirm whether the form of stationarity used for time series forecasting is of any importance or whether first differencing is sufficient to achieve stationarity for this purpose. The forecasted time series are presented in Table A.2 of the appendix and the regression results are presented in Table 2.

Table 1: Empirical results for the stationarity of ASA, NED, ALSI, CPI and ZAR/USD.

	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**m = 71**
NED (level)	0.894 (-2.567)	0.791 (-3.480)	-0.026 (-2.567) [0.674]	-2.188 (-3.435) [0.211]	-2.232 (-3.965) [0.471]	0.894 (0.739)	0.791 (0.216)	0.918 [0]
NED (first diff.)			-36.741 (-2.567) [0]	-36.730 (-3.435) [0]	-36.715 (-3.965) [0]	0.083 (0.739)	0.075 (0.216)	-0.155 [0.151]
NED (frac. diff.) {0.918}			-54.062 (-2.567) [0.0001]	-54.092 (-3.435) [0.0001]	-54.049 (-3.965) [0]	0.424 (0.739)	0.250 (0.216)	**0.141** [0.612]
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**m = 71**
ASA (level)	-1.242 (-2.567)	-2.486 (-3.480)	0.099 (-2.567) [0.714]	-2.646 (-3.435) [0.084]	-2.744 (-3.965) [0.219]	1.096 (0.739)	0.629 (0.216)	0.960 [0]
ASA (first diff.)			-21.334 (-2.567) [0]	-21.334 (-3.435) [0]	-21.326 (-3.965) [0]	0.076 (0.739)	0.081 (0.216)	**0.071** [0.403]
ASA (frac. diff.) {0.960}			-52.606 (-2.567) [0.0001]	-52.599 (-3.435) [0.0001]	-52.552 (-3.965) [0]	0.370 (0.739)	0.201 (0.216)	0.096 [0.045]
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**m =**
ALSI (level)	-0.787 (-2.567)	-1.849 (-3.480)	0.340 (-2.567)	-1.933 (-3.435)	-2.000 (-3.965)	0.873 (0.739)	0.544 (0.216)	1.049 [0]

			[0.783]	[0.317]	[0.600]			
ALSI			-34.208	-34.204	-34.191	0.091	0.094	**0.040**
(first diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.532]
			[0]	[0]	[0]			
ALSI			-72.228	-72.180	-72.099	0.224	0.131	-0.061
(frac. diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.1755]
{1.049}			[0.0001]	[0.0001]	[0.0001]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**m = 71**
CPI	2.977	-0.298	3.348	1.508	-1.362	3.727	1.074	1.026
(level)	(-2.567)	(-3.480)	(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0]
			[0.999]	[0.999]	[0.871]			
CPI			-32.778	-33.016	-33.165	0.689	0.072	-0.169
(first diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.063]
			[0]	[0]	[0]			
CPI			-42.236	-42.397	-42.505	0.189	0.082	**0.013**
(frac. diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.894]
{1.026}			[0.0001]	[0.0001]	[0]			
	DF-GLS		**ADF**			**KPSS**		**GPH - *I(d)***
	intercept	**int. and trend**	**none**	**intercept**	**int. and trend**	**intercept**	**int. and trend**	**m = 71**
ZARUSD	-1.991	-1.995	-0.138	-1.997	-1.977	0.641	0.605	1.028
(level)	(-2.567)	(-3.480)	(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0]
			[0.6359]	[0.2881]	[0.613]			
ZARUSD			-35.405	-35.391	-35.380	0.071	0.076	0.043
(first diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.643]
			[0]	[0]	[0]			
ZARUSD			-83.997	-83.939	-83.853	0.224	0.095	**0.027**
(frac. diff.)			(-2.567)	(-3.435)	(-3.965)	(0.739)	(0.216)	[0.551]
{1.028}			[0]	[0]	[0]			

Note: critical values at 1% in (), p-values in [], test statistics normal and the fractional diff parameter in {}.

The results in Table 2 indicate the following: i) the fractional difference NED series forecast outperformed the first difference forecast by 0.2563 per cent in forecasting accuracy; ii) the first difference ASA series forecast outperformed the fractional difference forecast by 2.0202 per cent; iii) the first difference ALSI series forecast outperformed the fractional difference forecast by 2.4296 per cent; iv) the fractional difference CPI series forecast outperformed the first difference forecast by 0.4233 per cent; and v) the fractional difference ZAR/USD series forecast outperformed the first difference forecast by 1.0444 per cent. These results confirm that the form of stationarity used in time series modelling is indeed of importance and that first differencing to achieve stationarity is not always sufficient. The same ca,n therefore, also be extended to fractional differencing. The econometrician and financial analyst should, therefore, take heed to first determine the data's inherent properties before simply going ahead with any form of differencing. Only once the data's inherent properties are known can the econometrician continue with the correct process to render the data stationary.

Table 2: Comparison of regression results for ASA, NED, ALSI, CPI and ZAR/USD forecasting performance.

NED			ASA		
Dependent variable	Independent variable	R^2	Dependent variable	Independent variable	R^2
Actual	**Fractional difference forecast**	**0.42117**	Actual	Fractional difference forecast	0.54854
Actual	First difference forecast	0.41861	Actual	**First difference forecast**	**0.56874**
ALSI			CPI		
Dependent variable	Independent variable	R^2	Dependent variable	Independent variable	R^2
Actual	Fractional-difference forecast	0.55215	Actual	**Fractional difference forecast**	**0.859811**
Actual	**First difference forecast**	**0.57645**	Actual	First difference forecast	0.855578

ZAR/USD		
Dependent variable	Independent variable	R^2
Actual	**Fractional difference forecast**	0.963267
Actual	First difference forecast	0.952823

7. Conclusion and recommendations

The manner in which a particular time series should be rendered stationary must be led by the statistical properties of the data, and not by blindly following popular techniques. The statistical properties of a non-stationary time series is the only criteria that should be considered when deciding on whether the data must be de-trended, first differenced or fractionally differenced in order to render it stationary. To this end, we propose a novel process (Appendix: Figure A.1) of determining the correct manner in which a non-stationary time series should be rendered stationary. We also applied this process to the closing prices of two stock prices, a stock index, consumer price index and an exchange rate in order to test the hypothesis that using the correct form of stationarity will improve forecasting results. The time series used for this purpose consisted of the daily closing prices of ABSA Bank Ltd. (ASA), the daily closing prices of Nedbank Ltd. (NED), the daily closing values of the JSE all share index (ALSI), the daily values for the South African Consumer Price Index (CPI) and the daily realised ZAR/USD exchange rate. The period selected wasfrom 9 October 2006 to 7 October 2011 and included 1253 observations per time series. In order to test whether the correct form of stationarity did indeed improve forecasting results, we used the ARFIMA model to forecast future prices.

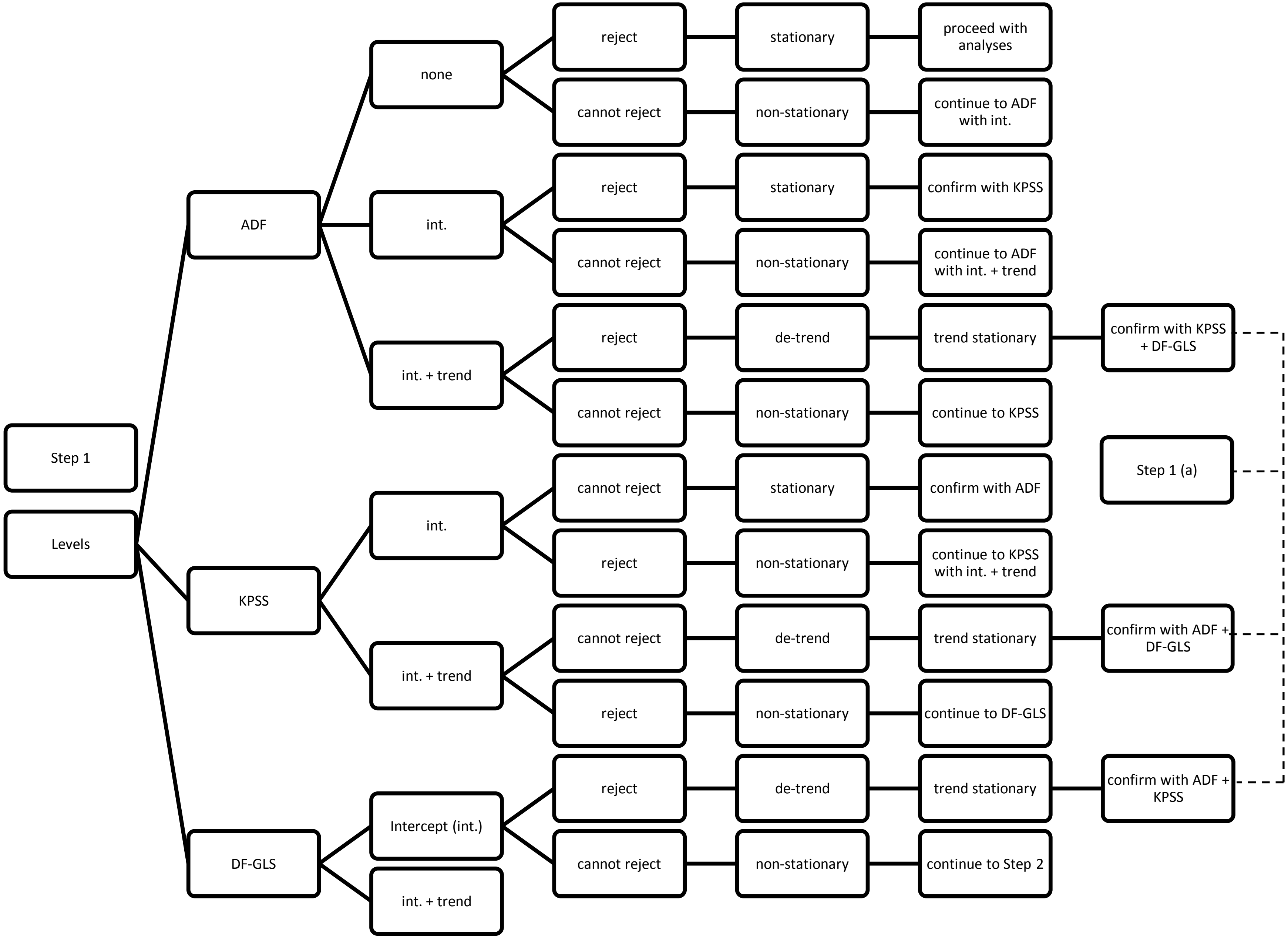
After applying our process to the non-stationary ASA, NED, ALSI, CPI and ZAR/USD time series, we found that ASA and ALSI were first difference stationary, while NED, CPI and

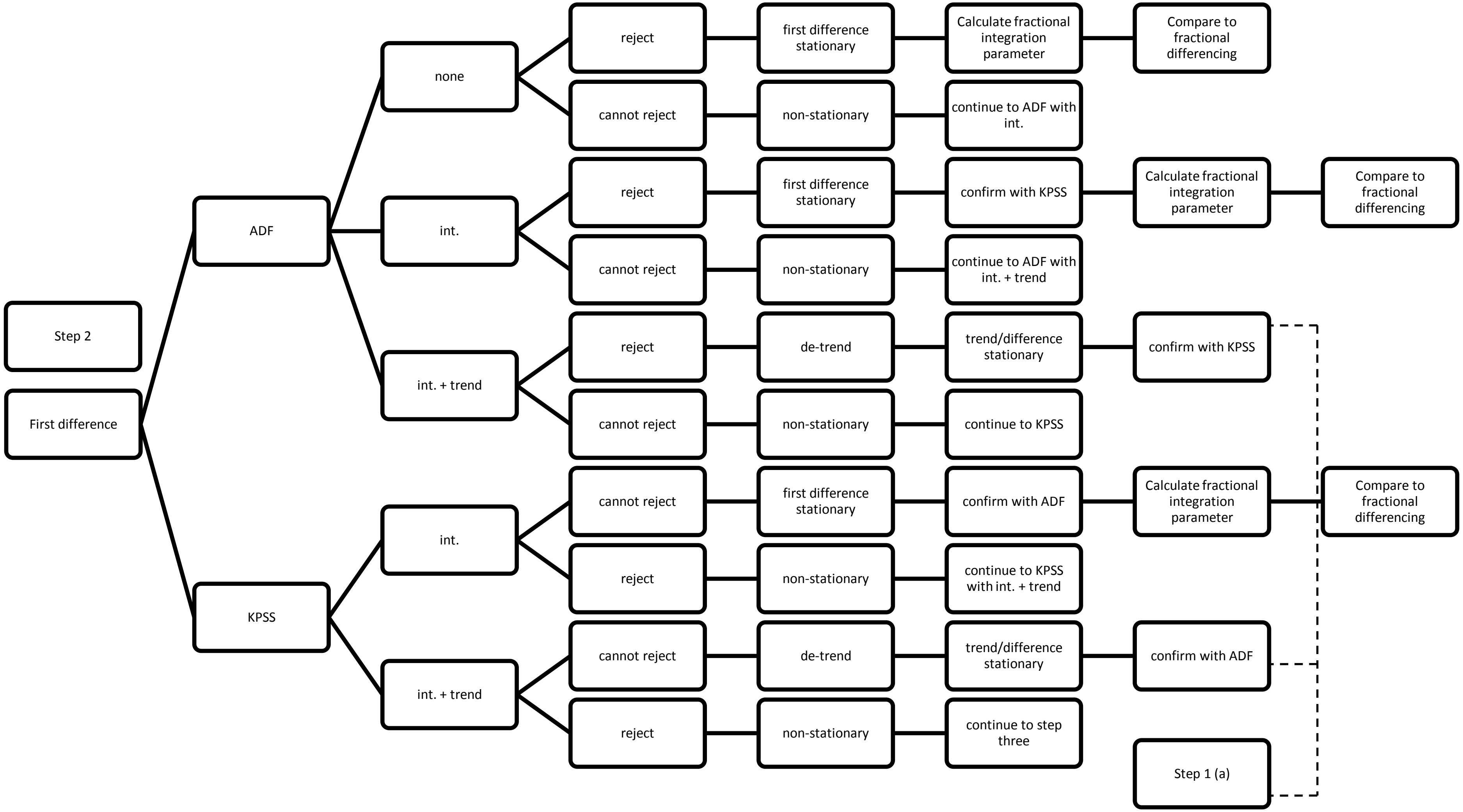
ZAR/USD were fractional difference stationary. We then proceeded by forecasting fifty one-step-ahead values, by using both the first and fractional difference time series of each time series. These forecasted time series were used as independent variables in OLS regressions with the actual observed time series as the dependent variables. The regression with the highest R^2 will indicate whether the forecasting time series is superior to its alternative. Results from the regressions indicated that all the time series that were found to be fractional difference stationary, outperformed their first difference forms and *vice versa*.

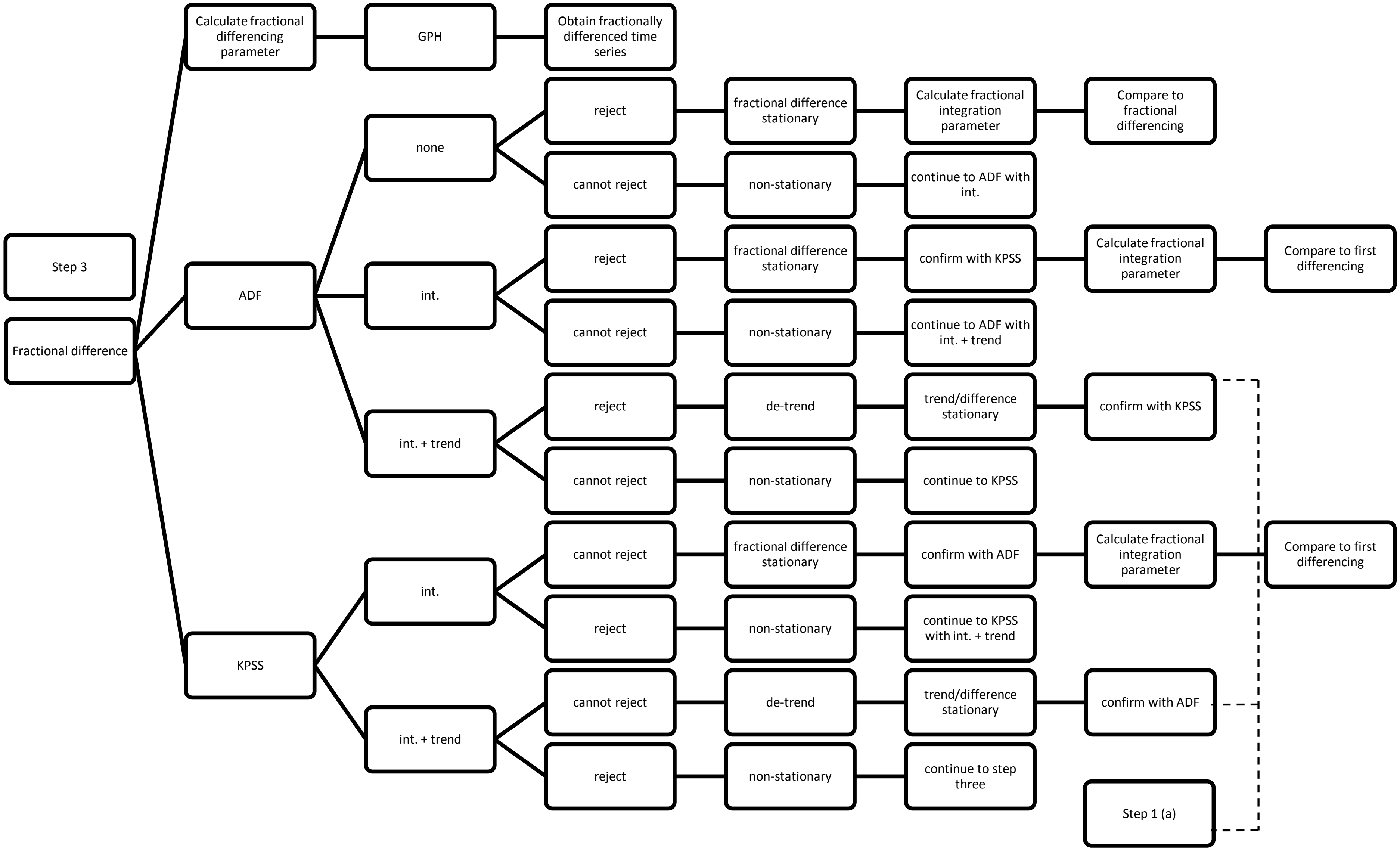
The results confirm that it is indeed reckless to difference first, and ask questions later. First differencing is not the only method available for rendering a time series stationary, and it is imperative that econometricians start looking at the properties of the data and not blindly follow a process suggested for a different dataset in the literature.

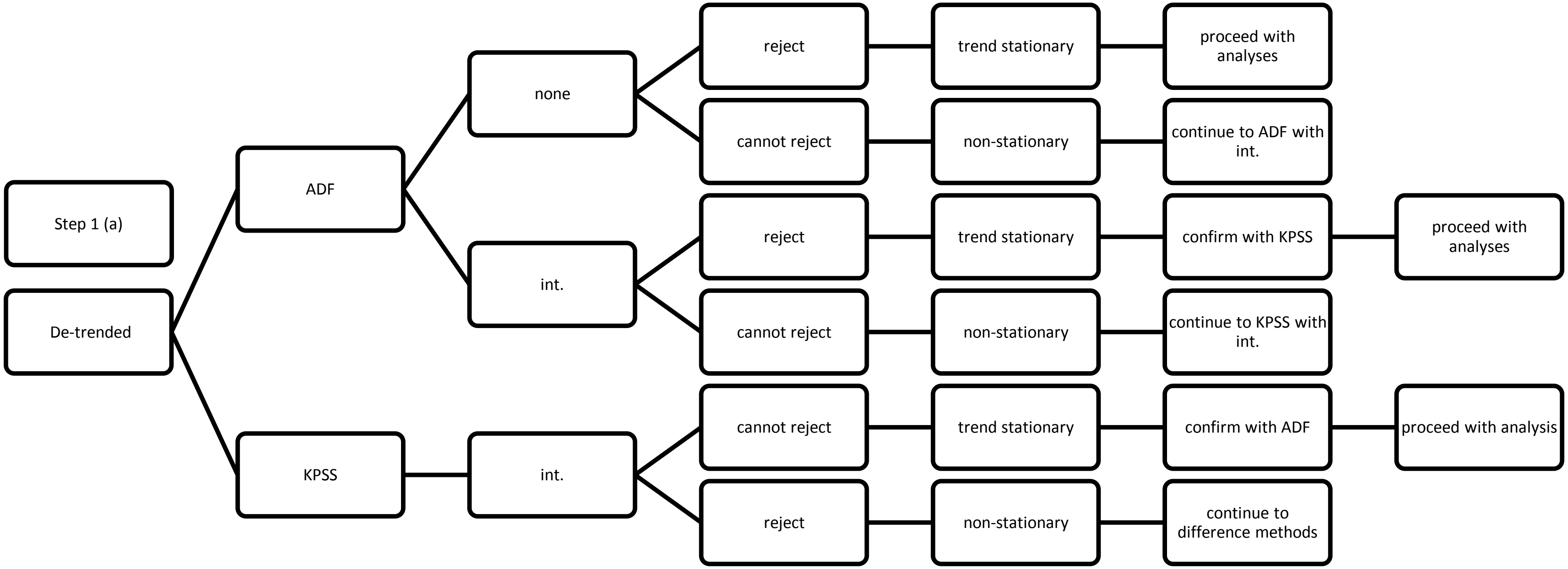
8. Appendix

Figure A.1: The process of rendering a non-stationary time series stationary (Step1, Step 1.a, Step 2 and Step 3)




Step 2
First difference
ADF
KPSS
none
int.
int. + trend
int.
int. + trend
reject
cannot reject
reject
cannot reject
reject
cannot reject
cannot reject
reject
cannot reject
reject
first difference stationary
non-stationary
first difference stationary
non-stationary
de-trend
non-stationary
first difference stationary
non-stationary
de-trend
non-stationary
Calculate fractional integration parameter
continue to ADF with int.
confirm with KPSS
continue to ADF with int. + trend
trend/difference stationary
continue to KPSS
confirm with ADF
continue to KPSS with int. + trend
trend/difference stationary
continue to step three
Compare to fractional differencing
Calculate fractional integration parameter
confirm with KPSS
Calculate fractional integration parameter
confirm with ADF
Step 1 (a)
Compare to fractional differencing
Compare to fractional differencing


Step 3
Fractional difference
Calculate fractional differencing parameter
GPH
Obtain fractionally differenced time series
ADF
none
reject
fractional difference stationary
Calculate fractional integration parameter
Compare to fractional differencing
cannot reject
non-stationary
continue to ADF with int.
int.
reject
fractional difference stationary
confirm with KPSS
Calculate fractional integration parameter
Compare to first differencing
cannot reject
non-stationary
continue to ADF with int. + trend
int. + trend
reject
de-trend
trend/difference stationary
confirm with KPSS
cannot reject
non-stationary
continue to KPSS
KPSS
int.
cannot reject
fractional difference stationary
confirm with ADF
Calculate fractional integration parameter
Compare to first differencing
reject
non-stationary
continue to KPSS with int. + trend
int. + trend
cannot reject
de-trend
trend/difference stationary
confirm with ADF
reject
non-stationary
continue to step three
Step 1 (a)


Step 1 (a)
De-trended
ADF
KPSS
none
int.
int.
reject
cannot reject
reject
cannot reject
cannot reject
reject
trend stationary
non-stationary
trend stationary
non-stationary
trend stationary
non-stationary
proceed with analyses
continue to ADF with int.
confirm with KPSS
continue to KPSS with int.
confirm with ADF
continue to difference methods
proceed with analyses
proceed with analysis

ASA, NED, ALSI, CPI and ZAR/USD.

NED				ASA			
		Forecast				Forecast	
Date	Actual	Frac. Diff.	First Diff.	Date	Actual	Frac. Diff.	First Diff.
29-Jul-11	14050	14194	14203	29-Jul-11	13150	13191	13193
01-Aug-11	14250	14073	14076	01-Aug-11	13295	13161	13162
02-Aug-11	13899	14260	14260	02-Aug-11	13401	13295	13301
03-Aug-11	13745	13912	13927	03-Aug-11	13290	13393	13394
04-Aug-11	13297	13777	13777	04-Aug-11	12725	13292	13277
05-Aug-11	13150	13341	13347	05-Aug-11	12755	12772	12730
08-Aug-11	13033	13203	13198	08-Aug-11	12355	12788	12794
10-Aug-11	12675	13084	13078	10-Aug-11	12541	12414	12409
11-Aug-11	13151	12733	12729	11-Aug-11	13200	12571	12606
12-Aug-11	13524	13190	13172	12-Aug-11	13420	13169	13235
15-Aug-11	13790	13532	13522	15-Aug-11	13500	13405	13396
16-Aug-11	13614	13777	13776	16-Aug-11	13325	13478	13441
17-Aug-11	13940	13603	13611	17-Aug-11	13500	13322	13259
18-Aug-11	13600	13929	13921	18-Aug-11	13101	13500	13450
19-Aug-11	13685	13594	13602	19-Aug-11	13500	13089	13067
22-Aug-11	13900	13691	13682	22-Aug-11	13100	13513	13491
23-Aug-11	13811	13892	13886	23-Aug-11	13150	13085	13095
24-Aug-11	14090	13802	13804	24-Aug-11	13325	13170	13165
25-Aug-11	13800	14077	14069	25-Aug-11	13055	13317	13340
26-Aug-11	13903	13792	13798	26-Aug-11	13250	13087	13068
29-Aug-11	14491	13904	13896	29-Aug-11	13810	13234	13265
30-Aug-11	14092	14461	14453	30-Aug-11	13568	13781	13799
31-Aug-11	14225	14065	14082	31-Aug-11	13949	13551	13531
01-Sep-11	14350	14226	14209	01-Sep-11	14150	13901	13904
02-Sep-11	14147	14329	14330	02-Sep-11	13665	14096	14088
05-Sep-11	13765	14137	14142	05-Sep-11	13580	13664	13613
06-Sep-11	13250	13778	13782	06-Sep-11	13400	13583	13561
07-Sep-11	13865	13289	13291	07-Sep-11	13975	13417	13411
08-Sep-11	14111	13885	13866	08-Sep-11	14494	13932	13989
09-Sep-11	13740	14078	14098	09-Sep-11	13951	14449	14472
12-Sep-11	13530	13745	13751	12-Sep-11	13831	13935	13906
13-Sep-11	13705	13552	13551	13-Sep-11	13800	13853	13810
14-Sep-11	13715	13714	13712	14-Sep-11	13900	13787	13801
15-Sep-11	13900	13712	13720	15-Sep-11	14251	13900	13910
16-Sep-11	13674	13897	13894	16-Sep-11	14150	14227	14248
19-Sep-11	13690	13671	13682	19-Sep-11	13775	14121	14130
20-Sep-11	13520	13703	13696	20-Sep-11	13993	13782	13766
21-Sep-11	13650	13525	13535	21-Sep-11	13953	14000	14000
22-Sep-11	13421	13661	13654	22-Sep-11	13510	13931	13959
23-Sep-11	13078	13425	13438	23-Sep-11	13395	13538	13531
26-Sep-11	13195	13111	13112	26-Sep-11	13301	13408	13441
27-Sep-11	13855	13213	13216	27-Sep-11	13766	13336	13358
28-Sep-11	13727	13839	13833	28-Sep-11	13770	13753	13806
29-Sep-11	13280	13719	13717	29-Sep-11	13375	13764	13775
30-Sep-11	13674	13308	13299	30-Sep-11	13434	13405	13374
03-Oct-11	13750	13661	13665	03-Oct-11	13412	13452	13447
04-Oct-11	13300	13738	13738	04-Oct-11	13315	13428	13427
05-Oct-11	13485	13327	13318	05-Oct-11	13474	13337	13332
06-Oct-11	13450	13486	13488	06-Oct-11	13510	13476	13487
07-Oct-11	13770	13445	13455	07-Oct-11	13895	13510	13511

Table A.2: ARFIMA forecast results for ASA, NED, ALSI, CPI and ZAR/USD. (Cont.)

ALSI				CPI			
		Forecast				Forecast	
Date	Actual	Frac. Diff.	First Diff.	Date	Actual	Frac. Diff.	First Diff.
29-Jul-11	31257.31	31226	31216	29-Jul-11	18301	18118	18107
01-Aug-11	30871.78	31278	31294	01-Aug-11	18160	18301	18301
02-Aug-11	30520.43	30892	30892	02-Aug-11	18211	18168	18167
03-Aug-11	29601.61	30548	30506	03-Aug-11	18000	18209	18223
04-Aug-11	29256.75	29632	29544	04-Aug-11	17250	18003	17997
05-Aug-11	28391.18	29304	29216	05-Aug-11	16900	17237	17221
08-Aug-11	28658.57	28433	28376	08-Aug-11	16851	16857	16824
10-Aug-11	29490.2	28724	28710	10-Aug-11	17000	16823	16796
11-Aug-11	29826.4	29554	29577	11-Aug-11	17150	16991	16970
12-Aug-11	30159.42	29852	29869	12-Aug-11	17175	17157	17144
15-Aug-11	29987.27	30162	30128	15-Aug-11	17375	17176	17196
16-Aug-11	30178.72	29966	29899	16-Aug-11	17521	17377	17388
17-Aug-11	29288.86	30169	30124	17-Aug-11	17400	17531	17545
18-Aug-11	29378.69	29271	29276	18-Aug-11	17214	17406	17469
19-Aug-11	29433.12	29387	29442	19-Aug-11	17330	17206	17269
22-Aug-11	29259.86	29441	29510	22-Aug-11	17202	17322	17316
23-Aug-11	29604.59	29266	29281	23-Aug-11	17490	17201	17204
24-Aug-11	29349.45	29611	29567	24-Aug-11	17380	17506	17505
25-Aug-11	29425.05	29350	29274	25-Aug-11	17296	17380	17400
26-Aug-11	30292.76	29427	29386	26-Aug-11	17650	17292	17299
29-Aug-11	30365.13	30290	30305	29-Aug-11	17750	17655	17650
30-Aug-11	31005.5	30346	30423	30-Aug-11	18845	17761	17780
31-Aug-11	31088.12	30983	31067	31-Aug-11	18725	18919	18856
01-Sep-11	30518.92	31059	31037	01-Sep-11	18500	18758	18789
02-Sep-11	29888.14	30499	30419	02-Sep-11	18200	18510	18458
05-Sep-11	29525.83	29885	29795	05-Sep-11	18500	18189	18205
06-Sep-11	30514.6	29532	29547	06-Sep-11	18500	18507	18539
07-Sep-11	30918	30529	30643	07-Sep-11	18697	18708	18492
08-Sep-11	30440.87	30906	30984	08-Sep-11	18697	18704	18655
09-Sep-11	29855.75	30423	30456	09-Sep-11	18498	18493	18712
12-Sep-11	30032.71	29856	29721	12-Sep-11	18412	18405	18530
13-Sep-11	30459.82	30037	29903	13-Sep-11	18755	18406	18373
14-Sep-11	30998.05	30472	30447	14-Sep-11	19150	18767	18744
15-Sep-11	31051.35	31011	31102	15-Sep-11	19200	19176	19171
16-Sep-11	30959.59	31046	31154	16-Sep-11	19120	19219	19212
19-Sep-11	31343.9	30956	30953	19-Sep-11	19080	19128	19083
20-Sep-11	31311.17	31345	31240	20-Sep-11	19008	19084	19101
21-Sep-11	30323.04	31295	31183	21-Sep-11	18601	19008	19028
22-Sep-11	30061.21	30330	30263	22-Sep-11	18300	18583	18582
23-Sep-11	29718.56	30074	30120	23-Sep-11	18380	18275	18260
26-Sep-11	30752.4	29737	29857	26-Sep-11	18600	18372	18390
27-Sep-11	30339.03	30756	30878	27-Sep-11	18780	18607	18630
28-Sep-11	29688.89	30341	30356	28-Sep-11	18850	18793	18779
29-Sep-11	29674.2	29700	29584	29-Sep-11	19200	18861	18845
30-Sep-11	29877.9	29698	29562	30-Sep-11	19200	19225	19216
03-Oct-11	29178.27	29900	29876	03-Oct-11	18700	19213	19223
04-Oct-11	29468.56	29194	29290	04-Oct-11	18775	18683	18646
05-Oct-11	29888.87	29499	29505	05-Oct-11	19038	18772	18785
06-Oct-11	30244.9	29912	29954	06-Oct-11	18880	19050	19042
07-Oct-11	30884.7	30248	30246	07-Oct-11	19200	18879	18876

Table A.2: ARFIMA forecast results for ASA, NED, ALSI, CPI and ZAR/USD. (Cont.)

ZAR/USD			
		Forecast	
Date	Actual	Frac. Diff.	First Diff.
29-Jul-11	6.6466	6.71	6.7229
01-Aug-11	6.6425	6.65	6.6361
02-Aug-11	6.7708	6.6461	6.6573
03-Aug-11	6.6771	6.7725	6.7575
04-Aug-11	6.7228	6.6787	6.6845
05-Aug-11	6.7808	6.7241	6.7081
08-Aug-11	6.7998	6.7808	6.7906
10-Aug-11	6.948	6.7989	6.7931
11-Aug-11	6.9307	6.9453	6.9423
12-Aug-11	7.1558	6.9266	6.9378
15-Aug-11	7.1815	7.1503	7.1306
16-Aug-11	7.2318	7.1733	7.1975
17-Aug-11	7.137	7.2231	7.202
18-Aug-11	7.1646	7.1279	7.1619
19-Aug-11	7.1295	7.1582	7.1376
22-Aug-11	7.126	7.1245	7.1482
23-Aug-11	7.2207	7.1227	7.1114
24-Aug-11	7.1971	7.2173	7.2207
25-Aug-11	7.1685	7.1939	7.2021
26-Aug-11	7.2109	7.1663	7.155
29-Aug-11	7.2464	7.2085	7.2213
30-Aug-11	7.2105	7.2436	7.2261
31-Aug-11	7.0912	7.2087	7.2294
01-Sep-11	7.0575	7.092	7.0778
02-Sep-11	7.022	7.0589	7.0695
05-Sep-11	7.0067	7.0239	7.0182
06-Sep-11	7.0016	7.0087	7.0006
07-Sep-11	7.0805	7.0036	7.007
08-Sep-11	7.081	7.0812	7.0782
09-Sep-11	7.093	7.0814	7.0846
12-Sep-11	7.1315	7.0931	7.0893
13-Sep-11	7.2444	7.131	7.1338
14-Sep-11	7.3321	7.2426	7.2276
15-Sep-11	7.4272	7.3287	7.3387
16-Sep-11	7.3989	7.4222	7.4206
19-Sep-11	7.4081	7.3936	7.402
20-Sep-11	7.4261	7.4038	7.4033
21-Sep-11	7.5486	7.4223	7.4222
22-Sep-11	7.6659	7.5431	7.5459
23-Sep-11	7.7977	7.6735	7.6668
26-Sep-11	8.1412	8.121	7.7871
27-Sep-11	8.2825	8.1303	8.1378
28-Sep-11	8.1599	8.2738	8.3068
29-Sep-11	7.9481	8.1546	8.1348
30-Sep-11	7.8612	7.9456	7.9465
03-Oct-11	7.8728	7.8601	7.8507
04-Oct-11	8.0268	7.8727	7.8761
05-Oct-11	8.0783	8.0229	8.0259
06-Oct-11	8.2609	8.0744	8.0771
07-Oct-11	8.1359	8.2564	8.2534

8. References

ASTERIOU, D. & HALL, S.G. 2007. Applied econometrics: a modern approach. New York: Palgrave Macmillan. 397 p.

BURKE, S.P. & HUNTER, J. 2005. Modelling non-stationary economic time series: A multivariate approach. New York: Palgrave MacMillan. 253 p.

DAVIES, R.B. & HARTE, D.S. 1987. Tests for the Hurst effect. *Biometrika Trust*, 74(1):95-101. Mar.

DIEBOLD, F.X. & YILMAZ, K. 2008. Measuring financial asset return and volatility spillovers, with application to global equity markets. *Working paper no. 8-16*. Reasearch department, Federal reserve Bank of Philidelphia.

DOORNIK, J.A., HENDRY, D.F., ARELLANO, M., BOND, S., BOSWIJK, H.P. & OOMS, M. 2007. Econometric modelling. PcGive, Vol 3. OxMetrics™ 5. London: Timberlake Consultants.

DOORNIK, J.A. & OOMS, M. 2004. Inference and forecasting for ARFIMA models with an application to US and UK inflation. *Studies in Nonlinear Dynamics & Econometrics*, 8(2):1-23.

ELDER, J. & SERLETIS, A. 2006. On fractional integrating dynamics in the U.S. stock market. *Chaos, Solitons and Fractals*. 34(1):777-781.

ENDERS, W. 2010. Applied econometric time series. Hoboken: Wiley. 517 p.

ERFANI, A. & SAMIMI, A.J. 2009. Long memory forecasting of stock price index using a fractionally differenced ARMA model. *Journal of Applied Sciences Research*, 5(10):1721-1731.

FANG, H., KON, S.L. & LAI, M. 1994. Fractal structure in currency futures price dynamics. *The Journal of Futures Markets*. 14(2):169-181.

GALLEGATI, M. 2008. Wavelet analysis of share returns and aggregate economic activity. *Computational Statistics and Data Analysis*. 51(6):3061-3074.

GEWEKE, J. & PORTER-HUDAK, A. 1983. The estimation application of long-memory time series models. *Journal of Time Series Analysis*. 4(1):221-238.

GIL-ALANA, L.A. 2006. Fractional integration in daily stock market indexes. *Review of Financial Economics*. 15:28-48.

GRANGER, C.W.J. & JOYEUX, R. 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15-29. Jan.

GRANGER, C.W.J. & NEWBOLD, P. 1974. Spurious regressions in econometrics. *Journal of Econometrics*, 2:111-120.

GUJARATI, D.N. 2006 Essentials of econometrics. New York: McGraw Hill. 553 p.

HOSKING, J.R.M. 1981. Fractional differencing. *Biometrika Trust*, 68(1):1025-1034.

HURST, H.E. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770-799.

JOSHI, P. 2011. Return and volatility spillovers among Asian stock markets. *SAGE Open*. http://sgo.sagepub.com/content/1/1/2158244011413474. Date of access: 25 October 2011.

LEYBOURNE, S.J., MCCABE, B.P.M. & TREMAYNE, A.R. 1996. Can economic time series be differenced to stationarity? *Journal of Business & Economic Statistics*, 14(4):435-446. Oct.

LIEN, D. & YANG, L. 2009. Intraday return and volatility spillover across international copper futures markets. *International Journal of Managerial Finance*, 5(1):135-149.

LO, A.W. 1991. Long-term memory in stock market prices. *Econometrica*, 59(5):1279-1313. Sep.

MADDALA, G.S. & KIM, I.M. 2000. Unit roots, co-integration, and structural change. Cambridge: Cambridge. 505 p.

MANDELBROT, B. 1972. Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Statistics*, 23(3):1015-1028.

MCCABE, B.P.M. & TREMAYNE, A.R. 1995. Testing time series for difference stationarity. *Annals of Statistics*, 23(3):1015-1028.
MONTGOMERY, D.C., JENNINGS, C.L. & KULAHCI, M. 2008. Introduction to time series analysis and forecasting. Hoboken: Wiley. 445 p.
MURRAY, M.P. 1994. A drunk and her dog: An illustration of co-integration and error correction. *The American Statistician*, 48(1):37-39.
NELSON, C. & PLOSSER, C. 1982. Trends and random walks in macroeconomic time series models: some evidence and implications. *Journal of Monetary Economics*, 10(2):139-162.
PENG, C.K., HAVLIN, S., STANLEY, H.E. & GOLDBERGER, A.L. 1994. Quantification of scaling exponents and crossover phenomena in non-stationary heartbeat time series. *Chaos*, 5(1):82-87.
ROBINSON, P. 1995. Gaussian semi-parametric estimation of long range dependence. *Annals of Statistics*. 23(5):1630-1661.
SOWELL, F. 1990. The fractional unit root distribution. *Econometrica*. 58(1):495-505.
STYGER, P., VILJOEN, S. & VAN VUUREN, G. 2008. A triptych on the USD-ZAR exchange rate dynamics. (In the Proceeding of the 5th international conference on applied financial economics, Greece, 03-05 July 2008. p. 334-344).

CHAPTER 5 – CONCLUSIONS AND RECOMMENDATIONS

5.1 INTRODUCTION

First differencing a financial time series has become a commonly used method to render non-stationary time series data stationary. This approach of first differencing has been scrutinised by comparing it to other methods that can be used to render time series data stationary. The aim of this study was to determine the impact of different forms of stationarity on financial time series analysis to gain insight into how applicable other forms of stationarity are in the field of financial time series analysis. The following sub-sections provide an overview of the conclusions (Section 5.2) reached and recommendations (Section 5.3) made based on the research contained in this work.

5.2 CONCLUSIONS

In Chapter 2, it was found that the stationarity of time series is an essential characteristic required to be achieved before analysing time series data. Stationarity has been shown to be a prerequisite to perform applied econometric analysis. Due to this prerequisite, it has become common practice to difference financial time series data in order to achieve stationary. In this chapter, the common practice of differencing data indiscriminately was challenged by presenting an overview of what has been established through other methods of achieving stationarity.

A variety of different forms of stationarity have been reviewed. These include: 1) difference stationarity; 2) fractionally differenced stationarity; 3) trend stationarity; 4) cyclical stationarity; and 5) seasonal stationarity. An important factor in rendering a time series stationary is to determine the presence of a stochastic trend, deterministic trend, unit root, cyclical component or seasonal component. In the case of determining the presence of a stochastic trend or unit root, the econometrician and financial analyst must follow the differencing method for achieving stationarity. However, should a time series display a deterministic trend, the data must be de-trended and then tested for stationarity before any differencing can take place. A time series containing a cyclical component can be de-trended by using the HP filter or BK filter to determine whether the series is cyclically stationary. A seasonally-stationary time series can be achieved by removing a seasonal component from that time series by means of seasonal differencing. After a financial time series has been reduced to stationarity, it is necessary to confirm stationarity through tests. The popular tests available that have a unit root as null-hypothesis with stationarity as an alternative are the DF, ADF and DF-GLS tests. Econometricians and financial analysts can also use the KPSS test with the null-hypothesis being that of stationarity. It has been suggested that econometricians and financial analysts should use the KPSS test in conjunction with one of the unit root tests as a confirmatory analysis for the presence of stationarity.

The discussion above shows the econometrician and financial analyst that reducing a financial time series to stationarity is a much more intricate process than simply taking first differences, testing for stationarity with the ADF test, and proceeding with, for example, forecasting models. Therefore, when conducting applied econometric analysis to time series, it is necessary to investigate the properties of the time series before launching into differencing the data. Only thereafter should it be decided which method must be applied to reduce a financial time series to the correct form of stationarity.

Chapter 3 continued challenging the assumption that most financial time series are first difference stationary. The common 'difference first, ask questions later' approach was revisited by taking a more systematic approach when analysing the statistical properties of financial time series data. To this end, a novel systematic process was developed to determine which method best suited the specific time series in order to render the data stationary. This process therefore focuses on, and is led by, the underlying statistical properties of a particular time series.

After applying the relevant process[56] and methods[57] of reducing a time series to stationarity, the following results were reported: i) one time series was found to be trend stationary; ii) four of the time series were first difference stationary; iii) five time series were over-differenced after being first differenced; and iv) more than half of the time series tested were fractional difference stationary.

The results confirmed that it is imperative that the statistical properties of the specific data series be determined before deciding on the actions to be taken. The assumption of most time series being first difference stationary cannot hold any longer due to the risk of over-differencing a time series. If researchers practice due diligence and allow the statistical

[56] Testing for stationarity by employing DF-GLS, ADF and KPSS.
[57] De-trending, first differencing and fractional differencing.

properties of a time series to lead them in rendering the series stationary, the occurrence of over-differencing can be avoided and improved empirical results may be achieved.

In Chapter 4, it was shown that the statistical properties of a non-stationary time series are the only criteria that should be considered when deciding on whether the data must be de-trended, first differenced or fractionally differenced in order to render them stationary. To this end, a novel systematic process of determining the correct manner in which a non-stationary time series should be rendered stationary, is proposed. This process was applied to the daily closing prices of ABSA Bank Ltd. (ASA), the daily closing prices of Nedbank Ltd. (NED), the daily closing prices of the JSE all share index (ALSI), the daily values for the South African Consumer Price Index (CPI) and the daily realised ZAR/USD exchange rate. The period selected was from 9 October 2006 to 7 October 2011, and included 1253 observations per time series. After applying this process to the non-stationary ASA, NED, ALSI, CPI and ZAR/USD time series, ASA and ALSI were found to be first difference stationary, while NED, CPI and ZAR/USD were fractional difference stationary. Forecasting 50 one-step-ahead values then followed, using both the first and fractionally differenced time series of each time series. These forecasted time series were used as independent variables in OLS regressions with the actual observed time series as dependent variables. The regression with the highest R^2 indicated whether the forecasting time series is superior to its alternative. Results from the regressions indicated the following: i) the fractional difference NED series forecast outperformed the first difference forecast by 0.2563 per cent in forecasting accuracy; ii) the first difference ASA series forecast outperformed the fractional difference forecast by 2.0202 per cent; iii) the first difference ALSI series forecast outperformed the fractional difference forecast by 2.4296 per cent; iv) the fractional difference CPI series forecast outperformed the first difference forecast by 0.4233 per cent; and v) the fractional difference ZAR/USD series forecast outperformed the first difference forecast by 1.0444 per cent.

The results confirm that it is indeed reckless to difference first and ask questions later. First differencing is not the only method that should be used to render a time series stationary, and it is imperative that econometricians and financial analysts begin exploring the data properties and do not blindly follow processes suggested for different datasets in the literature.

Based on the research conducted in this dissertation, and the results obtained from applying different empirical analysis, the following can be concluded: 1) different forms of stationary financial time series do have different impacts on time series analysis. This was shown when the different forms of stationarity found in a single time series were compared to each other in their forecasting accuracy[58]; and 2) methods other than simply first differencing to render a time series stationary do have significance and should be considered in financial time series analysis.[59]

5.3 CLOSING STATEMENT AND RECOMMENDATIONS

Stock markets, traders, speculators, hedgers and financial risk managers have, in the last six years,[60] been subject to booms, recessions, recoveries and sovereign debt crises. These events have precipitated new opportunities, new regulations, new risks and levels of volatility that were previously considered inconceivable. These new levels of volatility could have led to a fundamental shift in the properties of financial time series data, thereby leading us to question the widely accepted belief that most time series are first difference stationary.[61]

This study set out to determine whether different forms of stationarity have an impact on financial time series analysis. From these findings it is clear that further research is potentially endless. There are countless articles on countless subjects that make use of the first or log-difference approach to render data stationary. Since this is the case, it may very well be worthwhile to redo these econometric tests after first discerning whether the data used are in fact first difference stationary.

In terms of further research on the processes followed, there is still a void in testing for other forms of stationarity. The current stationarity tests available are limited to unit root tests and tests with stationarity as null-hypothesis. The tests that are available have also been criticised as having low power (Maddala *et al.*, 2000:100). Furthermore, there are no tests that are aimed directly at other forms of stationarity. For example, there are no tests developed with the sole purpose of testing whether a fractionally differenced time series has been rendered stationary. Lastly, there is also no research dedicated towards determin-

[58] See Chapter 4.
[59] See Chapters 2 and 3.
[60] 2006-2011.
[61] See Chapter 2.

ing whether unit root tests are suitable when determining whether a fractionally differenced time series is stationary.

REFERENCES

AADLAND, D. 2005. De-trending time-aggregated data. *Economic Letters*, 89:287-93.

ARELLANO, C. & PENTULA, S.G. 1995. Testing for trend stationarity versus difference stationarity. *Journal of Time Series Analysis*, 16:147-164

ASHLEY, R. & VERBRUGGE, R.J. Comments on "A critical investigation on detrending procedures for non-linear processes". *Virginia Tech Economics Department, Working paper.* http://ashleymac.econ.vt.edu/working_papers/ashley_verbrugge_comment.pdf. Date of Access: 4 May 2012.

ASTERIOU, D. & HALL, S.G. 2007. Applied econometrics: a modern approach. New York: Palgrave Macmillan. 397 p.

BAXTER, M. & KING, R.G., 1999. Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81:575-593.

BOX, G.E.P. & JENKINS, G.E.M. 1976. Time series analysis: Forecasting and control. San Francisco: Holden Day. 575 p.

BURKE, S.P. & HUNTER, J. 2005. Modelling non-stationary economic time series: A multivariate approach. New York: Palgrave MacMillan. 253 p.

CACCIA, D.C., PERCIVAL, D., CANNON, M.J., RAYMOND, G. & BASSINGWAIGHTE, J.B. 1997. Analyzing exact fractal time series: Evaluating dispersional analysis and rescaled range methods. *Physica A*, 246(3-4):609-632.

CHEUNG, Y. 1993. Long memory in foreign exchange rates. *Journal of Business and Economic Statistics*. 11(1):93-101.

CHOI, I. 1994. Residual-based tests for the null of stationarity with application to US macroeconomic time series. *Econometric Theory*, 10:720-746.

DAVIES, R.B. & HARTE, D.S. 1987. Tests for the Hurst effect. *Biometricka Trust*, 74(1):95-101. Mar.

DE JONG, D.N. & WHITEMAN, C.H. 1993. Estimating moving average parameters: Classical pileups and Bayesian posteriors. *Journal of Business and Economic Statistics*, 11(3):311-317.

DICKEY, D.A. & FULLER, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427-431. Jun.

DICKEY, D.A. & FULLER, W.A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057-172. Jul.

DIEBOLD, F.X. & YILMAZ, K. 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *Working paper no. 8-16*. Reasearch department, Federal reserve Bank of Philidelphia.

DOORNIK, J.A. & OOMS, M. 2004. Inference and forecasting for ARFIMA models with an application to US and UK inflation. *Studies in Nonlinear Dynamics & Econometrics*, 8(2):1-23.

DOORNIK, J.A., HENDRY, D.F., ARELLANO, M., BOND, S., BOSWIJK, H.P. & OOMS, M. 2007. Econometric modelling. PcGive, Vol 3. OxMetrics™ 5. London: Timberlake Consultants.

ELDER, J. & SERLETIS, A. 2006. On fractional integrating dynamics in the U.S. stock market. *Chaos, Solitons and Fractals*. 34(1):777-781.

ELLIOTT, G., ROTHENBERG, T.J. & STOCK, J.H. 1996. Efficient tests for an autoregressive unit root. *Econometrica.* 64(1):813-836.

ENDERS, W. 2010. Applied econometric time series. Hoboken: Wiley. 517 p.

ERFANI, A. & SAMIMI, A.J. 2009. Long memory forecasting of stock price index using a fractionally differenced ARMA model. *Journal of Applied Sciences Research*, 5(10):1721-1731.

FAMA. E.F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383-417. May.
FANG, H., KON, S.L. & LAI, M. 1994. Fractal structure in currency futures price dynamics. *The Journal of Futures Markets*. 14(2):169-181.
FIELITZ, B.D. 1971. Stationarity of random data: Some implications for the distribution of stock price changes. *Journal of Financial and Quantitative Analysis*, 6(3):1025-1034.
GALLEGATI, M. 2008. Wavelet analysis of share returns and aggregate economic activity. *Computational statistics and data analysis*. 51(6):3061-3074.
GEWEKE, J. & PORTER-HUDAK, A. 1983. The estimation application of long-memory time series models. *Journal of Time Series Analysis*. 4(1):221-238.
GIL-ALANA, L.A. 2006. Fractional integration in daily stock market indexes. *Review of Financial Economics*. 15:28-48.
GRANGER, C.W.J. & JOYEUX, R. 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15-29. Jan.
GRANGER, C.W.J. & NEWBOLD, P. 1974. Spurious regressions in econometrics. *Journal of Econometrics*, 2:111-120.
GUJARATI, D.N. 2006 Essentials of econometrics. New York: McGraw Hill. 553 p.
HODRICK, J.R. & PRESCOTT, E.C. 1997. Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1-16. Feb.
HOSKING, J.R.M. 1981. Fractional differencing. *Biometrika Trust*, 68(1):1025-1034.
HURST, H.E. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770-799.
JOSHI, P. 2011. Return and volatility spillovers among Asian stock markets. *SAGE Open*. http://sgo.sagepub.com/content/1/1/2158244011413474. Date of access: 25 October 2011.
KALMAN, R.E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1): 35-45.
KING, R.G. & REBELO, S.T. 1993. Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17:207-231.
KWAITKOWSKI, D. PHILLIPS, P.C.B., SCHMIDT, P. & SHIN, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159-178.
LEYBOURNE, S.J. & MCCABE, B.P.M. 1994. A consistent test for a unit root. *Journal of Business and Economic Statistics*, 12:157-166.
LEYBOURNE, S.J., MCCABE, B.P.M. & TREMAYNE, A.R. 1996. Can economic time series be differenced to stationarity? *Journal of Business & Economic Statistics*, 14(4):435-446. Oct.
LIEN, D. & YANG, L. 2009. Intraday return and volatility spillover across international copper futures markets. *International Journal of Managerial Finance*, 5(1):135-149.
LO, A.W. 1991. Long-term memory in stock market prices. *Econometrica*, 59(5):1279-1313. Sep.
MADDALA, G.S. & KIM, I.M. 2000. Unit roots, cointegration, and structural change. Cambridge: Cambridge. 505 p.
MANDELBROT, B. & TAQQU, M. 1979. Robust R/S analysis of long-run serial correlation. *Bulletin of the International Statistical Institute*, 42nd Conference Proceedings, Manila. 2:59-104.
MANDELBROT, B. & WALLIS, J. 1969a. Computer experiments with fractional Gaussian noises. *Water Resources Research*, 5:228-267. Feb.
MANDELBROT, B. 1972. Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Statistics*, 23(3):1015-1028.

MCCABE, B.P.M. & TREMAYNE, A.R. 1995. Testing time series for difference stationarity. *Annals of Statistics*, 23(3):1015-1028.
MONTGOMERY, D.C., JENNINGS, C.L. & KULAHCI, M. 2008. Introduction to time series analysis and forecasting. Hoboken: Wiley. 445 p.
MURRAY, M.P. 1994. A drunk and her dog: An illustration of cointegration and error correction. *The American Statistician*, 48(1):37-39.
NABEYA, S. & TANAKA, K. 1988. Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Annals of Statistics*, 16:218-235.
NELSON, C. & PLOSSER, C. 1982. Trends and random walks in macroeconomic time series models: some evidence and implications. *Journal of Monetary Economics*, 10(2):139-162.
PARK, J.Y. 1990. Testing for unit roots and cointegration by variable addition. *Advances in Econometrics*, 8:107-133.
PENG, C.K., HAVLIN, S., STANLEY, H.E. & GOLDBERGER, A.L. 1994. Quantification of scaling exponents and crossover phenomena in non-stationary heartbeat time series. *Chaos*, 5(1):82-87.
PLOSSER, C. I. & SCHWERT, G. W. 1977. Estimation of a noninvertible moving average process: The case of over-differencing. *Journal of Econometrics*, 6:199-224.
ROBINSON, P. 1994. Efficient tests of non-stationary hypotheses. *Journal of the American Statistical Association*, 89(428):1420-1437.
ROBINSON, P. 1995. Gaussian semi-parametric estimation of long range dependence. *The Annals of Statistics*. 23(5):1630-1661.
SAIKKONEN, P. & LUUKKONEN, R. 1993. Testing for a moving average unit root in ARIMA models. *Journal of the American Statistical Association*, 88:596-601.
SOWELL, F. 1990. The fractional unit root distribution. *Econometrica*. 58(1):495-505.
STYGER, P., VILJOEN, S. & VAN VUUREN, G. 2008. A triptych on the USD-ZAR exchange rate dynamics. (In the Proceeding of the 5^{th} international conference on applied financial economics, Greece, 03-05 July 2008. p. 334-344).
TANAKA, K. 1990. Testing for a moving average unit root. *Econometric Theory*. 6:433-444.