ECONOMETRIC FORECASTING OF FINANCIAL ASSETS USING NON-LINEAR SMOOTH TRANSITION AUTOREGRESSIVE MODELS

Maya Clayton

A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews

2011

Econometric forecasting of financial assets using non-linear smooth transition autoregressive models

Maya Clayton

Thesis submitted for the degree of

Doctor of Philosophy

School of Management

University of St. Andrews

March 2010

I, Maya Clayton, hereby certify that this thesis, which is approximately 70,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in February 2006 and as a candidate for the degree of PhD in May 2007; the higher study for which this is a record was carried out in the University of St. Andrews between 2006 and 2010.

date ……. signature of candidate …….....

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date ……..... signature of supervisor …….....

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis: Access to printed copy and electronic publication of thesis through the University of St. Andrews.

date ……. signature of candidate ……

date ……..... signature of supervisor …….....

Abstract

Following the debate by empirical finance research on the presence of non-linear predictability in stock market returns, this study examines forecasting abilities of non-linear STAR-type models. A non-linear model methodology is applied to daily returns of FTSE, S&P, DAX and Nikkei indices. The research is then extended to long-horizon forecastability of the four series including monthly returns and a buy-and-sell strategy for a three, six and twelve month holding period using non-linear error-correction framework. The recursive out-of-sample forecast is performed using the present value model equilibrium methodology, whereby stock returns are forecasted using macroeconomic variables, in particular the dividend yield and price-earnings ratio. The forecasting exercise revealed the presence of non-linear predictability for all data periods considered, and confirmed an improvement of predictability for long-horizon data. Finally, the present value model approach is applied to the housing market, whereby the house price returns are forecasted using a price-earnings ratio as a measure of fundamental levels of prices. Findings revealed that the UK housing market appears to be characterised with asymmetric non-linear dynamics, and a clear preference for the asymmetric ESTAR model in terms of forecasting accuracy.

Acknowledgements

I am extremely grateful to my supervisor Prof. David McMillan for his invaluable professional advice, and for his great support throughout the whole process. You were a friendly guide to help me through the scary woods of the exciting world of econometrics.

I am also tremendously grateful to my beloved husband for his support, his endless love and patience; you were always there for me. You naïvely agreed to proof read my work, and so you don't have to read it again, just for you – the butler did it...

Table of contents

Chapter 1
Introduction

The main objective of this research is to investigate methods of econometric forecasting and to assess whether non-linear approach can improve forecasting accuracy of financial asset returns compared to traditional linear models. Time-series modelling and forecasting are important for a wide range of disciplines. A number of researchers have demonstrated the importance of accurate time-series forecasts for market participants and policy-makers (Granato and Suzuki, 1996; Montgomery et al., 1998; Alexander, 1999; McMillan, 2002). Thus, McMillan (2002) highlights the importance of understanding dynamics within financial markets, especially if these are characterised by non-linear adjustments.

The inability of linear models to successfully explain certain financial phenomena, such as leptokurtosis, volatility clustering and the leverage effect (Brooks, 2002) supports the application of non-linear methodologies to financial modelling. Furthermore, unexpected dramatic changes in the stock market price in the late 1990s and early 2000s where the prices significantly diverged from their fundamental values, have influenced research to re-examine the standard present value model and the topic of stock market predictability. This study will examine forecasting abilities of non-linear models, namely smooth transition autoregressive (STAR) models, compared to linear alternatives in the form of a random walk model and a linear regression using daily stock returns.

The assumption of predictability of stock market returns is inconsistent with the efficient market hypothesis, however an ample number of research studies have confirmed presence of predictability across various financial assets (Fama and French, 1998; Campbell and Shiller, 2001; Lewellen, 2004; Torous et al., 2004; Campbell and Yogo, 2006). In addition, the failure of linear models to validate the present value model encouraged the assumption of the presence of non-linear dynamics within the relationship between stock prices and their determinants, in particular dividend yield (McMillan, 2004; Kanas, 2005; Rapach and Wohar, 2005; Bali et al., 2008). The non-linear approach confirmed the apparent mean reversion behaviour of stock prices characterised with non-linear adjustments to the long-run equilibrium. The presence of these non-linear adjustments were attributed to the presence of market frictions, including transaction costs and limit to arbitrage, the presence of speculative bubbles, and interaction between noise traders and informed arbitrageurs. While it has been challenging to prove the presence of bubbles in the financial market due to difficulties involved in identifying the bubbles, the market frictions and traders' interaction have been successfully modelled using non-linear models. In particular, McMillan (2004) suggests that the exponential STAR (ESTAR) model is able to capture different dynamics following the different magnitude of divergences, thus accounting for market frictions where arbitrageurs will only engage in trade when a price deviation exceeds a certain cost barrier. Similarly, the logistic STAR (LSTAR) accounts for different dynamics arising from the sign of disequilibrium, thus capturing traders' behaviour in bullish and bearish markets (McMillan, 2001).

Thus, the current study intends to apply non-linear models to an error-correction framework in order to examine out-of-sample forecasting performance of STAR-type

models in the context of monthly stock returns time-series using dividend yield and price-earnings ratio. Furthermore, it has been suggested that the stock market predictability increases with the horizon, thus forecasts performed on long-horizon data suggested to offer more accurate forecasts (Fair and Shiller, 1990; Montgomery et al., 1998). Hence, further to the investigation of monthly returns predictability, the research will consider long-horizon forecasting in the form of three, six, and twelve month periods. While previous studies have concentrated on an in-sample long-horizon stock return predictability, this investigation will extend the limited research into an out-of-sample predictability of stock returns.

Furthermore, extending the type of financial assets examined in this study, the non-linear error-correction methodology is applied to the housing market. Whereby the present value equilibrium framework is applied to the forecasting of house price returns, using a real income as a measure of fundamental price levels. The research into non-linear forecasting of house prices is somewhat limited, compared to an overwhelming amount of research into financial market predictability. However, the housing market dynamics are of an immense importance for policy-makers as the effects of housing market changes might have severe consequences on the economy as a whole (Muellbauer and Murphy, 1997; Crawford and Fratatoni, 2003; Fraser et al., 2008; Miles, 2008). Thus, Case et al. (2001) found changes in housing market to have a greater effect on consumption than changes in the stock market. In addition, Koetter and Poghosyan (2009) pointed out that imbalances in the housing market might lead to instability in the financial and banking sector, thus highlighting the importance of understanding the dynamics of the housing market for policy makers.

The approach to forecasting house prices used in this study is based on the methodology suggested by Black et al., (2005). However, different from previous research (Black et al., 2005; Black et al., 2006; Goodman and Thibideau, 2008), the current study employs non-linear tests of stationarity in the addition of non-linear STAR-type models and performs an out-of-sample forecast, as opposed to in-sample examination.

The structure of the thesis is as follows: Chapter 2 offers an extensive review of time-series modelling and forecasting literature (Section 2.2), with an overview of linear and non-linear models and forecasting methodology applied in further empirical chapters (Section 2.3). Chapter 3 is an empirical study of daily stock returns predictability in the context of non-linear modelling. Chapter 4 applies a non-linear error-correction model to examine predictability of monthly stock using dividend yield and price-earnings ratio (Section 4.4), extending the research further by considering long-horizon out-of-sample forecasting (Section 4.5). Chapter 5 extends the examination to a different type of financial assets, and applies a non-linear approach to forecasting UK house prices using a price-income ratio. Chapter 6 summarises the empirical results and concludes.

Chapter 2
Review of literature and non-linear empirical forecasting techniques

2.1. Introduction

This chapter provides an overview of econometric forecasting with emphasis on non-linear modelling, followed by a detailed discussion of a methodology which will be applied in empirical chapters of the thesis. The literature review will provide an evaluation of an informative basis using existing concepts and theories within the subject of a non-linear forecasting approach.

The chapter is organised as follows: Section 2.2 provides a review of time-series modelling and forecasting literature, including an overview of non-linear models, topics of stationarity and stock return predictability, and issues involved in econometric forecasting and assessment of forecasting accuracy. The methodology is included in Section 2.3. Section 2.4 concludes.

2.2. Literature review

Introduction to econometric modelling

Time-series modelling and accurate forecasts are important for a wide range of disciplines. Thus, Granato and Suzuki (1996) demonstrated the use of econometric forecasting in political science by applying econometric modelling to voting behaviour. Similarly, Montgomery et al. (1998) examined the US unemployment rate and emphasised the importance of accurate forecasting of the series for the economy as a whole. Correspondingly, McMillan (2002) suggests that non-linear adjustment within financial markets presents an important issue for market and policy makers. Thus, while small deviations from the fundamental asset pricing equilibrium might remain uncorrected by the market participants, significantly larger variations in fundamental equilibrium, on the other hand, put an increasing pressure for both market participants and policy makers to intervene in order to correct disequilibrium.

Importantly, Chatfield (1977) opposes the notion of a true model on the basis that any econometric model that has been fitted to the data is merely an approximation to the truth, and some models are simply more robust to deviations from the selected model over time than others. Chatfield (1977) also proposed that models allowing parameters and structure to vary over time would have an advantage in suiting a real data over constant approximations.

Overview of non-linear models

Introduction to non-linear modelling

Time-series modelling, in particular cointegration methodology, has number of practical applications to financial markets. Some of these include spot-futures arbitrage, yield curve modelling, and index tracking (Alexander, 1999). Furthermore, the interest in non-linear models emerged from empirical observations of financial markets and the inability of linear models to explain some frequently occurring phenomena in financial data. Such phenomena include leptokurtosis, which is tendency of financial data to display fat tails and excess peakedness at the mean in its distribution. Volatility clustering is another common occurrence in financial assets returns where volatility has the tendency to appear in bunches in such way that large returns regardless of the sign follow large returns, whereas small returns follow small returns. In addition, linear models cannot account for the tendency of volatility to rise more following large price falls than following price rises of the same magnitude, which is known as leverage effect. On the contrary, non-linear models can capture these phenomena and successfully model financial series behaviour for the further use in forecasting. However, non-linear models require different estimation techniques to linear structure models, hence a number of researchers disregard the use of non-linear models due to their complexity and lack of appropriately valid tools of analysis[1]. For instance, Feige and Pearce (1976) point out the optimality of autoregressive moving average (ARMA) models in the use of forecasting is due to their low marginal cost which is outperformed

[1] For a detailed review of earlier work and development of non-linear modelling and a full list of references refer to Tong (1990). Tong (1990) explores the development of non-linear modelling through the first introduction of certain non-linear concepts and models to further development such as introduction of special cases and applications to various data sets.

by the high marginal benefit of the generated forecast. Evidently, there is a predicament of costs and complexity of implementing and interpreting non-linear models over their usefulness in modelling financial time-series. Chatfield (1997) suggests that while the principals of multivariate models, where a forecasting model of a variable includes explanatory variables, are theoretically appealing, there is a danger of inclusion of unnecessary explanatory variables which in turn might lead to a reduced forecasting ability of the model. Chatfield (1997) mentions that in many case studies simple univariate models appear to be more robust to model misspesification than more complex models are.

Further interest in non-linear behaviour in financial markets followed from numerous discussions and tests of whether the purchasing power parity (PPP) holds. According to Brooks (2002), the theory behind PPP is that the long-run exchange rate between two countries equals the ratio of their relative price levels. PPP implies stationarity of the real exchange rate. One method of testing PPP is through cointegration. According to the theory, the log of the exchange rate between two countries and the logs of the price levels in these countries should be cointegrated with the cointegrating vector [1 -1] (Brooks, 2002). In addition, the validity of PPP can be assessed by testing whether the real exchange rates are mean-reverting (Chortareas et al., 2002). However, the PPP hypothesis does not seem to hold when the standard Dickey-Fuller (DF) unit root test is applied. Whereas, the PPP hypothesis is supported when alternative panel unit root tests are used (MacDonald, 1996). Hence many researchers have suggested that this could be due to the fact that exchange rates follow a non-linear process which in fact is stationary.

Chortareas et al. (2002) also note an increased interest in applying non-linear models to modelling real exchange rates (Michael et al., 1997; Sarantis, 1999; Baum et al., 2001). These non-linear models include the threshold autoregressive (TAR) and smooth transition autoregressive (STAR) models. However, Kapetanios et al. (2003) point out that research literature lacks any investigations and attempts to distinguish between non-stationarity linear systems and stationary non-linear STAR models.

However, as pointed out by Abhyankar et al. (1995), the presence of non-linear structure in financial markets time-series data will be inconsistent with the statement of efficient market hypothesis (EMH).

Non-linear dynamics in financial time-series

The presence of non-linear dynamics in financial time-series is well documented with an ample number of studies confirming the presence of non-linearities across different types of financial time-series data. Thus, Abhyankar et al. (1995) found clear evidence of non-linear dependence in FTSE 100 returns using high-frequency data. Lekkos and Milas (2004) applied the STAR model to analyse excess returns predictability of the UK government bonds using various risk factors, including the forward premium, the slope of the term structure, excess FTSE stock returns and the FTSE index dividend yield. The results revealed regime-switching behaviour within the returns and time-varying structure of the expected excess returns. Consequently, while the linear autoregressive moving average (ARMA) model was the most commonly used model for time-series analysis and forecasting since the early 1970s, as De Gooijer and Kumar (1992) point out, that occasionally the preference for non-linear models was suggested by theory or

data, as linear models seemed to be unable to explain certain phenomena observed in financial time-series.

One of the proposed explanations for such non-linear mean-reverting adjustments in real exchange rates, in particular, was the presence of transaction costs. Thus, Kapetanios et al. (2003) point out that in the presence of transaction costs in the asset market, the profitability of arbitrage when there is a differential between the risk adjusted returns on two assets depends on whether this differential is greater than the transaction costs involved. Hence, Kapetanios et al. (2003) proposed that there is an inverse relationship between the speed of reversion to equilibrium and the size of the differential between returns, i.e. the larger the differential between the assets returns, the stronger the tendency to reverse back to the equilibrium. This can be explained due to the fact that owing to the presence of transaction costs, small deviations from the equilibrium price will not be corrected. Consequently, this will be reflected in non-linear behaviour of speed of reversion to the equilibrium as it will increase with the size of the deviation (McMillan, 2001). In other words, the speed of reversion will be close to zero in the case of small imbalances of the price hence indicating traders' inactivity. However, the speed of reversion will be increasing rapidly as the price deviations become larger creating profitable arbitrage opportunities. In addition, McMillan (2001) suggests other market frictions, such as short selling and borrowing constraints, to be the cause of non-linear behaviour. Effectively, deviations caused by these factors will differ in magnitude and will result in asymmetric dynamics of returns.

McMillan (2005b), nevertheless, argues that even though many studies in this area recognise non-linear dynamics caused by the presence of transaction costs, the speed of reversion, however, is modelled to be the same regardless of the sign of the

disequilibrium. This is based on the grounds of the behavioural finance approach considering the interactions between noise traders and arbitrageurs. McMillan (2001) explains the presence of asymmetries in financial market returns due to the interaction of informed traders and noise or uninformed traders, whose presence on the market ensures profitable arbitrage opportunities. As opposite to noise traders, informed traders will only engage in trading activities when price movements around the equilibrium price are significantly large for an arbitrage profit to be made exceeding any transaction costs. Martens et al. (1998) also pointed out that index-futures arbitrageurs will react in a similar way by not entering the market when the price deviation is not sufficient enough to compensate for the costs of transaction, thus creating a band of inactivity for arbitrage traders around the equilibrium. Martens et al. (1998) demonstrate the effects of the magnitude and the sign of mispricing, where the impact of mispricing increases with its size and the information effect of negative mispricing errors having a greater impact compared to the positive errors, by applying a threshold error-correction approach.

Another suggested possible explanation of such phenomenon can be explained by the presence of bubbles. Conversely, according to Evans (1991), temporary speculative deviations in price time series may occur due to periodically collapsing bubbles. There have been a number of attempts to model such non-linear dynamics consistent with a bubble component (van Norden and Vigfusson, 1998; Bohl and Siklos, 2004). However, Campbell et al. (1997) point out the difficulty of identifying and testing bubbles empirically. Moreover, there is a lack of theoretical support for the explanation of bubbles. McMillan (2009) points out that models based on the bubble approach do not contain any information about the dynamics that take place in the period leading to the start of the bubble. This issue would also be very important when using models in

forecasting as the bubble might be difficult to predict, particularly due to difficulties involving identification of initial dynamics leading to the occurrence of the bubble.

Furthermore, it appears that the theories of behavioural finance offer adequate explanation of the initial formation of market bubbles. According to behavioural finance theories, traders behave differently in rising and falling markets. Hence, their actions will endure different speeds of reversion depending on whether the change in the market was positive or negative in the sign. In other words, it is expected that noise or uninformed traders have tendency to overreact in a response to good news, i.e. positive disequilibrium (deviation from equilibrium). On the other hand, in the case of negative news, i.e. negative disequilibrium, noise traders seem to exhibit conservative behaviour (Shleifer, 2000). Bullish markets lead to overconfidence, trend-chasing and overreaction, whereas bearish markets are characterised by more conservative behaviour of traders as they are influenced by fundamental news (McMillan, 2006). If this empirical observation holds, then it is apparent that the speed of reversion will depend not only on the size but also on the sign of deviation from equilibrium. As a suggestion considering the discussed issue, McMillan (2005b) proposes the ESTR model, which allows for asymmetry in the sign of the disequilibrium.

However, West (1988) found little direct evidence of noise trading to have a significant effect on stock price determination in the late 1980s. Moreover, West (1988) defined a rational bubble as an extraneous event that has an effect on stock prices because it is expected to do so by the market participants, and also pointed out that different researchers may interpret the term bubble differently. Thus, for instance, bubble might be referred to as the explosive process, or it can be seen as any deviation from fundamental values due to speculation.

However, non-linear modelling poses a number of computational challenges. Thus, there is the additional issue of testing for presence of non-linearity, or in other words deciding whether a linear specification is sufficient enough to model particular financial data and thus which non-linear framework will resolve such matter. De Gooijer and Kumar (1992) pointed out that in terms of practical use the main requirement of a non-linear model is to be general enough in order to be able to capture a wide range of non-linearities. This criterion also applies to tests of non-linearity, model diagnostic and evaluation. Brooks (2002) suggests that the initial choice of linear or non-linear type of models considers whether there are any suggestions from financial theory that particular variables may have a non-linear relationship. Similarly, Teräsvirta et al. (2005) expressed their concerns that incorrect specification of a non-linear model at the model building stage could lead to the model producing an inferior forecast. In addition, Marmol and Velasco (2004) expressed their concerns about the presence of the spurious regression which may occur when applying cointegration analysis. The problem of spurious regression arises due to the presence of non-stationarity that can induce significant correlations between non-stationary series despite the absence of theoretical groundings or justification for any relationship between these series. However, despite the difficulties involved in non-linear modelling, the use of non-linear models is developing fairly rapidly.

Regime-switching models

There is a vast number of various non-linear models, however, this paper will concentrate on regime-switching type of non-linear models including threshold

autoregressive (TAR), standard and extended versions of smooth transition autoregressive (STAR), and error-correction model.

TAR and STAR models

Threshold autoregressive models (TAR) are a class of non-linear time-series autoregressive models, which, unlike standard autoregressive models, allow for locally linear approximation over different states (Brooks, 2002). TAR models were first proposed by Tong (1978) and later developed further by Tong and Lim (1980) and Tong (1983). It was initially suggested as an alternative model for describing periodic time-series (Tsay, 1989). Tong (1990) describes TAR models as a simplified way of presenting a complex stochastic system in terms of decomposing it into a set of smaller sub-systems. Tsay (1989) identifies the main features of threshold type models which include limit cycles, amplitude dependent frequencies, and jump phenomena. Generally, linear time-series models are unable to capture such characteristics of financial time series data. The main difference between the TAR type of models and Markov switching models is that the state variable, i.e. the variable determining the behaviour of the series under a particular state, is assumed to be known or observable, whereas it is prone to variation under conditions of Markov switching regimes.

The threshold autoregressive process is able to capture asymmetric limit cycles, as the main motivation for these models was to describe limit cycles of cyclical time-series (Tsay, 1989). Applications of TAR models include modelling exchange rates and modelling arbitrage opportunities implied by the difference between the spot and futures prices for a given market. For instance, Teräsvirta and Anderson (1992) raised an issue

of non-linearity in business cycles, as linear models such as ARMA can only be used in the case of symmetrical business cycles. Teräsvirta and Anderson (1992), presented strong evidence of presence of non-linearity in business cycles which confirmed that business cycles exhibit asymmetric behaviour.

Further, Teräsvirta and Anderson (1992) suggested that any non-linear time-series can be represented by a smooth transition autoregressive (STAR) model. Unlike standard TAR models, STAR models allow for more gradual transition of the dependent variable between regimes. The regime indicator in these models is a continuous function rather than an abrupt on-off switch of TAR models (Brooks, 2002). STAR models were first proposed by Chan and Tong (1986) as a generalisation of a non-linear two-regime univariate self-exciting threshold autoregressive (SETAR) model. Self-exciting threshold autoregressive (SETAR) models are a special case of general univariate TAR models, where the state-dependent variable is the dependent variable itself.

As pointed out by McMillan (2001), STAR models are able to capture two types of asymmetric adjustments such as the direction and size of the disequilibrium. In other words, these models allow for different dynamics depending on whether the value of the variable is above or below the threshold parameter, and between periods when the variable takes a large or small value. The logistic smooth transition autoregressive (LSTAR) model has a logic distribution that approximates to the normal distribution and also has an advantage in terms of being able to estimate its parameter using analytical derivatives. Luukkonen et al. (1988) also note the LSTAR model as having distinct computational advantages over standard TAR. However, the most important feature of the LSTAR process is that the model allows for smooth transition when the threshold is set to differentiate dynamics between positive and negative values of the

dependent variable. Furthermore, exponential STAR (ESTAR), on the other hand, is used when modelling the magnitude of the dynamics of the data as the model allows to account for different behaviour of the time-series depending on the size of values of the dependent variable. These differences in dynamics initiate non-linear adjustments of the data and follow theoretical explanations of the presence of non-linearities in the financial markets due to market frictions and interaction between informed and noise traders.

In addition, STAR models, by definition, offer a smooth transition between regimes as opposed to abrupt switch of TAR and Markov switching models, which seems to be more a plausible response in stock markets characterised by a large number of participants engaging in trading activities at slightly different times (Sarantis, 2001; McMillan, 2002). Moreover, Sarantis (2001) suggested that the differences in timing of market participants' reactions are due to heterogeneous beliefs of individual traders, variations in learning speeds and different investment horizons.

Equilibrium-correction systems

Equilibrium-correction econometric systems have emerged from cointegrating analysis. The error-correction mechanism was first introduced by Sargan (1964) and then further developed by Engle and Granger (1987). The hypothesis of cointegration is based on a notion that certain economic variables do not diverge from each other greatly in the long-run. Such variables might drift apart in a short-run due to various reasons, for instance seasonality; however, economic forces will intervene to bring them back to the equilibrium. Such economic forces include market mechanisms, such as arbitrage

trading, and government intervention. The concept of cointegration is closely linked to the existence of an error-correction model. In other words, cointegrating variables belong to an economic system which converges over time into a long-run equilibrium.

Cointegration was developed in order to investigate common trends in financial time-series and had proved to be a compelling technique of modelling long-run and short-run dynamics in multivariate economic systems. Furthermore, Alexander (1999) pointed out that portfolio risk management assessment techniques involve correlation analysis of returns, whereas cointegration analysis is based on raw price data. Hence, when the price data is differenced for standard risk-return models, vital information about long-term trends in the data might be removed. Alexander (1999) highlights the difference between the notions of cointegration and correlation which are related, however, are different concepts. Correlation mirrors co-movements in returns, whereas cointegration measures long-run co-movements in prices. A cointegrating relationship may still be present even when correlation between series is low. Hence, Alexander (1999) suggests that cointegration methodology generates more effective long-term hedging techniques. Similarly, investment management strategies benefit from being based on a cointegration approach rather than on standard correlation techniques which are unable to account for the presence of long-term trends in the data.

Harris and Sollis (2003) point out that differencing the variables when estimating dynamic models in order to achieve stationarity might result in vital long-run information to be lost. Hence, an error correction model is a more suitable approach since the model will incorporate both the short-run and long-run characteristics, where disequilibrium is in fact a process of adjustment to the long-run equilibrium model.

Cointegration techniques are widely used to test the asset pricing model, including testing the validity of the present value model, according to which current stock prices are discounted values of future dividends with the discount rate being equal to the required rate of return. Early studies on the present value model assumed dividends to be trend-stationarity, however found prices to be too volatile and inconsistent with the theory where rationally expected future dividends are discounted by a constant real interest rate (Caporale and Gil-Alana, 2004). Hence, later studies had suggested invalidity of trend-stationarity. However, it must be pointed out that the early tests were based on standard unit root tests for determining the order of integration. Hence, it can be argued that results of previous studies which were inconsistent with the present value model might be due to the low power of standard unit root tests. In addition, Caporale and Gil-Alana (2004) pointed out that failure to find cointegration could signify the presence of speculative bubbles rather than invalidity of the present value model.

Moreover, Campbell and Shiller (1988a) point out that our understanding of long-run equilibrium of cointegrating variables is more efficient in explaining long-run tendencies rather than short-run deviations. As a result, long-run equilibrium models, such as the error-correction model (ECM), are valid for describing long-run relationships between variables while having limited ability to explain slow adjustments to the equilibrium after a short-run random shock. Campbell and Shiller (1988a) suggest the following factors in an attempt to explain the lack of instantaneous adjustment back to the equilibrium including sticky prices, long-term contracts, or costs of adjustments. Consequently, there have been a number of various research studies attempting to develop an econometric model able to fit the long-run properties of the data as well as accommodate the type of short-run deviations.

Further extensions of the standard error-correction model include fractional cointegration and the Markov error-correction model. Thus, while the standard cointegration testing procedure often relies on standard unit-root tests which assume the order of integration to be an integer, fractional cointegration methodology allows the order of cointegration to be other than an integer. Moreover, it has been suggested that slow mean reversion might not be captured by the standard cointegration analysis as opposed to the fractional integration. Baillie and Bollerslev (1994) pointed out that short-run deviations seem to be highly persistent as a result of the error-correction term to react slowly to shocks. Thus, the deviations from the cointegrating relationship can be described as following a long memory process, or in other words, the effect of a random shock dies out at a slower rate comparing to exponential decay of autocorrelation functions, such as the ARMA process. Consequently, Caporale and Gil-Alana (2004) point out that standard cointegration analysis restricts the equilibrium error to be an *I(0)* process, which might not be consistent with highly persistent deviations from equilibrium where errors respond more slowly to shocks.

A fractionally integrated process first proposed by Granger (1980) is specifically intended to capture such long memory-type behaviour. The process allows a fractionally integrated process to describe a wider range of mean-reversion behaviour of financial variables that are beyond the capabilities of standard cointegration analysis. Similarly, Caporale and Gil-Alana (2004) suggest that the reason for empirical evidence surrounding studies of the present value model being inconclusive is the use of the discrete options *I(1)* and *I(0)* applied in a classical cointegration approach, which can be argued to be a restrictive condition. Hence, the researchers propose that the process of adjustment to the equilibrium might be expressed through a fractional integration *I(d)*.

According to Caporale and Gil-Alana (2004), standard cointegrating tests fail to recognise slow adjustments of deviations occurring as a result of shocks, thus producing results contradicting the present value models theory. Caporale and Gil-Alana (2004) provide evidence of the presence of fractional cointegration in relationship between stock prices and dividends, thus supporting the validity of the present value models over a long horizon. In addition, Cheung and Lai (1993) argue that fractionally integrated error-correction terms generate a flexible and parsimonious model that is able to capture low-frequency dynamics of short-term disequilibrium movements.

Another type of error-correction models are known as Markov error-correction (MEC) models. These models are characterised by being able to model the different rates of adjustment of deviations from the long-run equilibrium (Psaradakis et al., 2004). The main advantage of the MEC model is its flexibility allowing for non-stationary behaviour of deviations from the long-run equilibrium. This assumption seems to follow empirical observational evidence, as adjustments of an economic system after, for instance, a dramatic market crash are unlikely to be similar to adjustment following normal recession. Psaradakis et al. (2004) state that motivation for this type of investigation has emerged from historic observation of the US stock prices in certain periods when theory struggled to explain their behaviour in terms of their underlying fundamentals. As an attempt to provide reasonable explanations for such phenomenon, some researchers have proposed incorporating a time-varying discount factor, while others explained it due to the presence of intrinsic rational bubbles. Thus, Psaradakis et al. (2004) applied MEC methodology to US stock prices, and demonstrated that MEC models are able to identify periods of a short-term disequilibrium which is not corrected as a result of presence of either an intrinsic bubble or a time-varying discount factor. In

addition, the MEC model can also account for adjustment that are might be non-continuous or not constant in their strength. In addition, Psaradakis et al. (2004) point out that the MEC model is most suitable for cases where the change in the regime is caused by a sudden shock, which cannot be modelled by smooth transition or threshold models.

Psaradakis et al. (2004) suggest that while cointegrated relationships between stock prices and dividends seems to hold in the long-run, prices may deviate from the underlying fundamentals in the short-run. It appears that cointegration relationship fails in the short-run. However, it can be suggested that cointegration is still present in the short-run, but the adjustments are characterised by different rates, or speeds of adjustment. Therefore, further extensions of equilibrium-correction models allow for non-linear disequilibrium adjustments (Granger and Swanson, 1996; Balke and Fomby, 1997; Michael et al., 1997; Siklos and Granger, 1997; Peel and Davidson, 1998)

Applications of non-linear models

Guidolin and Timmermann (2006) demonstrated that standard linear models were unable to capture regime-switching dynamics of joint distribution to US stock and bond returns fully, whereas non-linear models provided a much thorough appreciation of the complexity of the data series. On the contrary, Brooks (2002) points out that despite the fact that switching models are able to fit the data sufficiently, these models do not seem to generate superior forecasts than linear models or random walk model. Dacco and Satchell (1999) suggest that poor forecasting results are due to the difficulties involved in forecasting the actual regime that the time-series will be in. Thus, Clements and

Smith (1997) point out that a number of researchers (Diebold and Nason, 1990; De Gooijer and Kumar, 1992) describe existing evidence on whether non-linear forecasts are superior to linear ones as irregular and are rather unconvincing.

Moreover, there seems to be much debate regarding forecasting performance of non-linear models as a whole. Thus, De Gooijer and Kumar (1992) carried out a review of development in non-linear time-series forecasting and concluded that there was no uniformity in literature to whether non-linear models provide forecasts superior to linear alternatives. Despite the disagreement in literature regarding non-linear time-series modelling, De Gooijer and Kumar (1992) are optimistic on the subject and suggest that non-linear models can be useful in modelling and forecasting certain financial phenomena when linear models fail to do so. Clements and Smith (1999), on the other hand, carried out an empirical research comparing multi-step forecast performances of SETAR and linear autoregressive models, in an attempt to surpass the conclusion made by De Gooijer and Kumar (1992) of no uniformity regarding evidence of forecasting ability of non-linear models. While Clements and Smith (1999) added to the empirical research of multi-step forecasting techniques, their results indicated non-linear models to have a forecasting advantage over linear alternatives depending on the regime of serial dependence. Overall, Clements and Smith (1999) concluded that non-linear models hold a significant potential improvement over linear models in terms of forecasting performance.

Thus, Abhyankar et al. (1995) found evidence of the presence of non-linearity in high frequency minute-to-minute FTSE returns, however also found the series to be adequately explained by a simple GARCH process. Clements and Smith (1997) compared different methods of multi-step ahead forecasting using self-exciting

autoregressive (SETAR) models and AR models as a comparative benchmark. The study analysed the US gross national product (GNP) data over two time periods, namely 1973:01 – 1990:04 and 1991:01 – 1994:04. In the first time period the SETAR models were found to be outperforming the AR model by 10% at four-steps ahead forecast. However, it did not demonstrate any clear preference over longer horizons. Moreover, the results from the second period were unclear in terms of the preferred forecasting model. Montgomery et al. (1998) compared forecasting performance of a number of linear and non-linear models using the US quarterly unemployment rate in order to capture asymmetric cyclical behaviour of the data during economic expansions and contractions. The univariate linear models including ARIMA were unable to efficiently describe cyclical asymmetries of the data, while non-linear models were found to produce significantly improved multi-step out-of-sample forecasts during economic contractions. Non-linear models used in the study included the threshold autoregressive (TAR) model and Markov switching autoregressive models.

Martens et al. (1998) applies an error-correction model to futures and spot prices where futures and index returns are explained by past futures and index returns, with the error-correction term being represented by mispricing error where deviations from the equilibrium are not arbitraged away immediately. The approach by Martens et al. (1998) presents the error-correction term as a reflection of the effects of arbitrage, whereby the traders' actions divert prices back to the equilibrium level thus causing the error-correction term to revert to zero. Moreover, Martens et al. (1998) suggested that mean reversion could also be caused by the concept of infrequent trading where mispricing as a result of new market information is not followed by a lagged reaction of market participants, thus not every trader will engage in the correction of mispricing over the

same short period of time. Thus, the effects of arbitrage and infrequent trading create somewhat similar patterns in futures prices, causing the mean reversion in a non-linear error-correction system, with infrequent trading producing more gradual reversal.

McMillan (2004) applied a non-linear error-correction model to short- and long-term UK interest rates and found the logistic STAR (LSTAR) model to outperform linear and non-linear alternatives considered in the study. The presence of market frictions, such as transaction costs, borrowing constraints and short-selling, induces non-linear equilibrium adjustment, whereby the speed of adjustment varies depending on the magnitude of deviation. In addition, actions by policy-makers cause the speed of adjustment to vary between positive and negative movements in the inflation rate, where the expectation of rising inflation is characterised by a quicker response than falling inflation. Similarly, Teräsvirta et al. (2005) found LSTAR models to generate more accurate forecasts comparing to linear AR models. Moreover, research found that combining forecasts improved the overall accuracy of forecasts. However, Teräsvirta et al. (2005) pointed out that it was unclear from the results of an investigation whether the difference in forecasting gain was substantial enough to justify application of non-linear models and the complex model building required in estimation of non-linear models.

As Alexander (1999) pointed out, based on a random walk approach, the best forecast for the future would be the current value plus a random shock or a drift. Nevertheless, since cointegration models contain information about the long-term equilibrium of the system, these models might be considered as potentially valid forecasting tools.

Stationarity and non-linearity

Testing for the presence of non-linearity

Naturally, implementation of non-linear models raises an issue of testing for the presence of non-linearity, that is, deciding whether a linear specification is sufficient enough to model particular financial data and thus which non-linear framework will resolve such matter. Brooks (2002) suggests that the initial choice of linear or non-linear types of models should consider whether there are any suggestions from financial theory that particular variables may have a non-linear relationship. Consequently, Luukkonen et al. (1988) recommended testing the presence of non-linearity as a first step in practical model building before applying a complex non-linear model to the data.

Early studies concentrated on portmanteau tests to detect non-linearity in time-series data. Thus, Davies and Petruccelli (1986) compared two promising tests for time-series non-linearity in anticipation that some time-series previously assumed to be a linear process might in fact contain non-linearity, and thus benefit from non-linear modelling. The tests included variations of a portmanteau test for non-linearity. However, the study found no definitive results in support of either of the tests considered. Moreover, Davies and Petruccelli (1986) argued that generation of a general statistic able to detect global non-linearity will be highly implausible.

Luukkonen et al. (1988) proposed a STAR model non-linearity test with the test statistic following Chi-squared distribution which, according to researchers, compares well to the CUSUM test used for testing against SETAR non-linearity. Similar to the Box-Jenkins procedure, Luukkonen et al. (1988) suggest that the model selection process, starting with specifying the correct order of the linear autoregressive model component,

could be performed by using an information criterion, such as SBIC (Schwarz, 1978). Luukkonen et al. (1988) proposes a number of tests that are more practical than the ones suggested by Chan and Tong (1986) and are not restricted to LSTAR non-linearity, and find an augmented first-order test procedure to be the most successful in providing a good alternative to the CUSUM test.

Kapetanios et al. (2003) point out that the earlier literature had mainly concentrated on assessing linear models, while not focusing on possible existence of non-linear dynamics in financial time-series. However, more recent literature has shifted the interest to the presence of non-linearities in financial market dynamics. Available empirical evidence supports the notion of presence of non-linear behaviour in financial variables. For instance, Chortareas et al. (2002) found evidence of non-linear mean-reversion in real exchange rates for the G7 countries. In addition, Caner and Hansen (2001) found the US unemployment rate to follow a stationary threshold autoregressive process.

Abhyankar et al. (1995) identified two main questions to be investigated in non-linear econometric modelling. One of these questions is inevitable when addressing the issues of the presence of non-linearity. First, whether it is possible to effectively identify the presence of non-linearity in financial time-series; and second, if the presence of non-linearity was detected, is there a suitable time-series model able to explain such non-linear behaviour. The discussion in this section of the chapter is intended to identify a reasonable number of studies that had attempted to answer these questions, however, it seems that the literature is unable to provide a clear answer to both of these subjects.

Importance of stationarity and the concept of non-linear stationarity

Many financial series are believed to be non-stationary. However, some researchers suggest that in many cases these series can be considered stationary in terms of a non-linear fashion. Hence, in the presence of non-linear behaviour in the series it is logical and reasonable to apply a specially designed non-linear test to detect the presence of stationarity. Moreover, De Gooijer and Kumar (1992) pointed out that the stationarity of the data should be established and corrected prior to model identification and estimation, as non-stationarity overwhelms genuine features and dynamics of the data and is not a source of non-linearity on itself. Davies and Petruccelli (1986) also pointed out that the early non-linearity tests required time-series data to be stationary prior to the testing.

Caner and Hansen (2001) suggest that previous studies on non-linear time-series, including TAR models, have assumed stationarity of the data used, which made it difficult to distinguish between non-stationarity and non-linearity. They also claim that early statistical methods are unable to discriminate non-stationarity from non-linearity due to the problem of the joint modelling of unit roots and thresholds. Moreover, many researchers were faced with a problem of examining such series since most well known methods of analysis are mainly developed exclusively for linear series, making these impossible to apply to non-linear dynamics. Consequently, there was an incentive to develop new techniques and frameworks suitable for time-series of non-linear nature. As a result, a large proportion of literature on non-linear dynamics is concentrating on testing for the presence of non-linear behaviours in financial markets. On the other hand, researchers who already are convinced that financial markets exhibit non-linear

behaviour concentrate on testing specific non-linear models in terms of their ability to explain market movements.

Henry (2006) points out that the non-stationary nature of financial markets is prone to structural breaks makes the equilibrium-correction a fairly risky forecasting tool. Hendry (2000) showed that the presence of local shifts in the data produced invalid forecasts from vector equilibrium-correction models (VEqCMs). Hence, it is evident that while VEqCMs generates significant forecast for a stationary series, it becomes unreliable should the location shifts occur.

It is apparent that applying standard linear Dickey-Fuller (DF) unit root test to a non-linear stationary process can lead to modelling misspecification and, hence, incorrect results. As a result there were a number of alternative unit root tests proposed by various researchers.

Tests of non-linear stationarity

Most studies on alternative stationarity tests were motivated by the fact that the standard DF test persistently fails to reject the null hypothesis of a presence of a unit root. In addition, following numerous studies on purchasing power parity (PPP) (MacDonald, 1996; Edison and Kloveland, 1987; Chortareas et al., 2002; Lo and Wong, 2006) where data exhibited some regime changes, it was found that the unit root can be rejected for such data after certain adjustments for regime changing shocks were made. All of these factors motivated studies aimed at finding an alternative procedure to standard unit root tests, namely the DF test.

Balke and Fomby (1997) have applied Monte Carlo simulations to the threshold autoregressive model with three regimes, thus intending to analyse non-stationarity and non-linearity jointly in terms of threshold cointegration. As a result, Balke and Fomby (1997) found that for threshold parameters the power of the DF test falls considerably. Similarly, Michael et al. (1997) suggest that standard unit root test and cointegration do not account for the effects of STAR non-linearity and hence can lead to biased results. Enders and Granger (1998) also developed a unit root test with the alternative hypothesis of stationarity with asymmetric adjustment. These tests were based on the threshold autoregressive (TAR) and the momentum threshold autoregressive (M-TAR) models. According to Enders and Granger (1998), M-TAR models are able to capture sharp or deep movements in the time-series sequence. Enders and Granger (1998) applied their test to term structure of interest rates and found that in the case of approximate symmetric adjustment the standard DF test is more powerful compared to the TAR and M-TAR models. However, the results are contrary when adjustment is asymmetric. In this case TAR and M-TAR models are significantly more powerful over the DF test.

Caner and Hansen (2001) developed a new asymptotic theory for an unrestricted two-regime threshold autoregressive (TAR) model with a possible unit root, which allows to distinguish a non-linear process from a non-stationary one. The methodology involved using asymptotic and bootstrap-based tests. Caner and Hansen (2001) found that in the cases where the true process is non-linear the standard DF test and ADF test are much less powerful than the suggested alternative test based on the TAR model. Furthermore, Chortareas et al. (2002) have proposed a unit-root test procedure against a stationary non-linear STAR building on a combination of works by Kapetanios et al. (2001) and

Schmidt and Phillips (1992). Chortareas et al. (2002) applied the test to the real exchange rates and found that for the majority of cases studied the null of unit root was rejected against the non-linear stationary STAR model. The DF test, on the other hand, was unable to reject the null. These results confirm that there is strong evidence of non-linear mean-reversion in the real exchange rates.

Kapetanios et al. (2003) argue that the presence of transaction costs in financial assets markets results in non-linear adjustments of rates of return to equilibrium thus exhibiting apparent non-stationarity. In other words, processes might only appear non-stationary when in fact they are stationary but non-linear. As a result, the standard DF and augmented DF (ADF) tests are not powerful enough against such dynamics. Kapetanios et al. (2003) introduced an easy to apply procedure for testing the presence of non-stationarity in time-series data using exponential smooth transition autoregressive (ESTAR) processes, applied their testing procedure to the real interest rates and real exchange rates from the 11 major OECD countries, and as a result have developed unit root test framework resistant against the ESTAR stationary process. The proposed test was found to have better power comparing to the standard DF test. Kapetanios et al. (2003) also found evidence of non-linear mean-reversion in both series.

Furthermore, on the basis of the test by Kapetanios et al. (2003), Sollis et al. (2002) proposed an asymmetric STAR-type unit root test by introducing asymmetry to mean reversion adjustment of real exchange rates. The test was further developed by Sollis (2009) as a stationarity test against asymmetric STAR process non-linearity. Similarly, Pascalau (2007) proposed a stationarity test framework, which allows to test the LSTAR process non-linear stationarity as well as testing for general STAR-type stationarity.

The group of non-linear stationarity tests based on works of Kapetanios et al. (2003) have proved to be robust and are characterised by the ease of application. These tests will be applied in empirical chapters of this paper and will be discussed in more technical detail in Section 2.3.

Stock returns predictability

As indicated by numerous researchers, the main application of econometric modelling is its application in forecasting (Chong and Hendry, 1986; Granger and Newbold, 1986; Diebold and Mariano, 1995; Montgomery et al., 1998; Pindyck and Rubinfeld, 1998; Brooks, 2002). The forecasting of financial variables, such as price series, based on the detection of patterns in the past values of the variable is usually referred to as technical analysis. Nevertheless, the research into stock market returns predictability produced rather extensive debate whether it is feasible to predict stock market behaviour using statistical measures of econometric modelling and as to whether technical analysis has any solid theoretical grounds and, hence, whether it is viable enough to use in practice. The issue received newly deserved attention when the hypothesis of market efficiency came under examination, since a number of studies suggested that the stock market returns do not fully reflect the market risks as proposed by the efficient market concept. Hence, as pointed out by Brock et al. (1992), the presence of predictability in stock market returns could be explained either by market inefficiency or time-varying equilibrium returns. Brock et al. (1992) found that the simplest trading rules used in

their study have confirmed their predictive power, however, the researchers have warned against data snooping and also suggested that the returns generating process for the stocks might be more complex than is anticipated by the linear models. This statement might be considered as an implication of non-linear dynamics of stock prices. Allen and Karjalainen (1999) avoid the problem of data snooping that occurs due to the ex post specification by using a learning generic algorithm which is generated using the data prior to the start of the test period. However, their results are consistent with the view that markets are efficient in the sense that the technical trading rule implemented in their study was unable to generate profit after transaction costs.

However, contrary to the implication of the efficient market hypothesis, evidence from numerous empirical investigations provides sufficient evidence of stock market returns predictability (Pesaran and Timmermann, 1995, 2000; McMillan, 2001; Rapach et al., 2005). Thus, Pesaran and Timmermann (1995, 2000) supported the presence of predictability of US and UK stock returns using a linear recursive forecasting approach. Moreover, Pesaran and Timmermann (1995) suggest that stock returns predictability seems to hold across international markets as well as different time horizons. Moreover, Abhyankar et al. (1997) suggested that early studies have doubted the possibility of stock market prices being described by a deterministic process due to the market movements being mainly triggered by the random flow of news. However, as the researchers pointed out, further profound understanding of non-linear systems and development of non-linearity detection tests supplemented further enquiry into the forecastability of market returns.

Pesaran and Timmermann (1995) pointed out that stock returns predictability could be explained maintaining the validity of an efficient market by the time-varying expected

returns. The researchers also suggested that while returns predictability could be a supporting evidence of market efficiency on the condition of constant expected returns, the predictability of excess returns, nonetheless, does not imply that the stock market is inefficient. Pesaran and Timmermann (1995) reference the concept of an intertemporal equilibrium model of the economy, which can explain stock predictability in conjunction with market efficiency, however attempts to substantiate the theory were proven to be inconclusive.

Rapach et al. (2005) suggests that the preference for macroeconomic variables comes from the fact that these are most likely to influence investments, consumption levels and expected cash flows, and hence are important variables of asset-pricing models. In addition, Rapach et al. (2005) also point out that due to mixed results of various empirical investigations there is no clear conformity to a particular macro variable as the most reliable in terms of stock returns predictability. Rapach et al. (2005) found interest rates to be the most reliable and consistent macro variable for forecasting stock returns.

Pesaran and Timmermann (1995) carried out an investigation into predictability of US stock returns and found evidence of stock predictability, however did not succeed in establishing a robust forecasting model. The results also revealed that the level of predictability is related to the patterns of business cycles and magnitude of the shocks. Moreover, periods of high levels of excess returns predictability seemed to correspond with periods of high volatility. Thus the study found that predictability of excess returns was higher during volatile period in the 1970s, compared to calmer periods of the 1960s and 1980s which were characterised by much smaller forecasting gains. Pesaran and

Timmermann (2000) also found presence of predictability when repeating their approach using UK stock returns data.

Fang and Xu (2003) carried out an empirical study into predictability of asset returns using daily data on the Dow Jones Industrial Average by combining a technical analysis approach and time-series forecasts. They claimed that since the asset returns were correlated, it was possible to capture predictability of the data, and as a result confirmed their suspicions. Moreover, Fang and Xu (2003), using a rolling out-of-sample forecasting technique, found that while both the technical trading approach and time-series models could both be successful in predicting the series, these two approaches seem to predict different components of the data. Thus, the trading rule approach captures positive movements in returns and performs well in a bull market, while the time-series approach identifies the negative movements performing well in a bear market. Hence, they suggest that the combination of the two approaches is far superior to either technical trading rules or time series modelling forecasts when used on their own.

As Brock et al. (1992) point out, the presence of stock returns predictability could be explained by either market inefficiency where the market prices deviate from their fundamental values, or by time-varying equilibrium returns in efficient markets. However, there is a lack of evidence support either of these two theories. Thus, Shively (2007) supports the efficiency of the stock market by providing evidence of time-variation in expected returns using the link between excess volatility and asymmetric volatility in stock prices, where the latter is explained by the leverage effect.

Non-linear dynamics in stock returns

Initial interest in the application of non-linear models has emerged from characterising cyclical behaviour of many of economic time-series (Sarantis, 2001). Furthermore, ample research revealed presence of non-linearities in stock prices (Tong 1990; De Gooijer et al., 1992; Abhyankar et al., 1997; McMillan, 2001; McMillan, 2002; Guidolin et al., 2008)[2]. Thus, Abhyankar et al. (1997) offer an extensive review of empirical studies testing the presence of non-linear dependence in real-time returns, with an overwhelming number of studies providing supporting evidence of presence of non-linearity. Abhyankar et al. (1997) analysed real stock data of four major indices including S&P 500, DAX, Nikkei 225 and FTSE 100, and was unable to reject the hypothesis of independence, thus providing evidence supporting non-linear structure of the considered data. The researchers, however, noted that some degree of observed non-linear dependence could be attributed to volatility clustering, which nevertheless is unlikely to explain non-linearity entirely. Similarly, Sarantis (2001) investigated presence of non-linearities in stock prices of the G-7 countries using the STAR model, and found linearity to be rejected for all time-series considered in the study and for the data to exhibit asymmetric cycle patterns. Moreover, the results of out-of-sample STAR model forecasts, namely logistic STAR (LSTAR) and exponential STAR (ESTAR), proved to be favourable in terms of forecasting gains. He and Modest (1995) point out that asset pricing models in general are based on the principal of fundamental equilibrium where the current asset price equals to its fundamental value. Linear models fail to satisfy restrictions imposed by the equilibrium condition. Moreover, Abhyankar

[2] For further references to earlier literature on non-linear modelling refer to Taylor et al. (2001).

et al. (1997) suggests that predictability of market returns could be consistent with market efficiency on the condition of short-term occurrence thus not allowing for speculative profit opportunities.

While extensive research provides supporting evidence of the presence of non-linear dynamics in the financial markets, further examinations of stock market behaviour attempt to offer adequate explanations of such phenomena. Thus, market frictions, such as transaction costs, limits to arbitrage, short selling and borrowing constraints were found to cause asymmetric adjustments to the fundamental equilibrium of asset pricing and thus causing non-linearities within the financial market (He and Modest, 1995; McMillan, 2002). In addition, there is a number of studies suggesting that temporary deviations of stock prices from their fundamentals, or in other words, from the long-run equilibrium relationship between stock prices and dividends, may be caused by the presence of speculative bubbles. Psaradakis et al. (2004) points out that there are two distinguishable types of bubble including periodically collapsing bubbles (Evans, 1991) and intrinsic bubbles (Froot and Obstfeld, 1991). According to Evans (1991), periodically collapsing bubbles are characterised by explosive conditional means, however do appear to follow a stationary process when tested using standard unit-root procedures. Froot and Obstfeld (1991), on the other hand, argued that intrinsic bubbles are responsible for short-term deviations from the long-run equilibrium. Moreover, an interaction between informed and noise traders is suggested as one of the reasons for observed non-linearity of financial markets (McMillan, 2001; McMillan, 2002).

Montgomery et al. (1998) compared forecasting performance on a number of linear and non-linear models using the US quarterly unemployment rate in order to capture asymmetric cyclical behaviour of the data during economic expansions and

contractions. The univariate linear models including ARIMA were unable to efficiently describe cyclical asymmetries of the data, while non-linear models were found to produce significantly improved multi-step out-of-sample forecasts during economic contractions. Abhyankar et al. (1995), on the other hand, found evidence of the presence of non-linearity in high frequency minute-to minute FTSE returns, however also found the series to be adequately explained by a simple GARCH process. Similarly, McMillan (2001) found presence of non-linearity in S&P monthly index returns and a non-linear smooth transition threshold type model to outperform the linear regression alternative in producing an out-of-sample forecast by only a marginal, however, nonetheless statistically significant difference.

Ready (2002) carried out an investigation into numerous research, often with contradicting conclusions on the predictability of daily returns using the example of US stock indices. In spite of arguments of the traditional view in finance literature, Ready (2002) claims that the actions of financial companies such as investment and financing are not responsive enough to short-term changes in the market as their activities are too cumbersome in order to react and adopt quickly. Hence, generally small market imperfections might be too costly to consider and take advantage of after taking into account transaction and processing costs. However, as Leitch and Tanner (1991) pointed out, in practice numerous companies pay extensive fees to professional forecasters in an attempt to account for those changes in the market. Moreover, Baker and Wurgler (2002) suggest that an activity of any company, and most importantly its capital structure, is a reflection of its cumulative attempts to design their actions in time with the equity market, whereby firms issue equity when their market values are high, and repurchase it when the market values are low. Equity market timing and its

significant effect on capital structure are strongly supported by Baker and Wurgler (2002). Moreover, they point out that the earnings forecasts related to the issue of equity generally tend to have an effect on investors anticipating prospective earnings. This evidence is one of a few examples of the importance and effect of forecasting in financial markets. In addition, Baker and Wurgler (2002) provide supporting evidence of predictability of the capital structure based on past values, such that they find a strong link between the current capital structure and the variation in the market-to-book ratio from previous years, going as far as ten years, maintaining the view that fluctuations in the market have a long term effect on capital structure. The basis for this evidence provided by Baker and Wurgler (2002) were drawn from analysis of actual financial decisions, analysis of equity issues following equity repurchases over a considerable period of time, and analysis of realised and forecasted equity earnings, as well as qualitative supporting evidence based on managerial surveys. There is no doubt that forecasting procedures play an important role in capital structure managing.

While Brock, Lakonishok and LeBaron (1992) found that the simple trading rules based on moving averages performed well in achieving a realistic profit, Ready (2002) criticised findings of the study for being a result of data snooping. Ready (2002) argues that the fact that such trading rule approach could generate after transaction costs profit, would either be an indication of market inefficiency or that there is a presence of time variation in stock returns. This is due to the fact that Brock et al. (1992) have employed the moving average style trading rules on the basis of their popular use amongst practitioners in the late 1980's. This popularity clearly came from the effectiveness of such approach in practice, however, in terms of academic theory and financial modelling methodology these findings can be considered spurious. Instead, Ready

(2002) supports the procedure proposed by Allen and Karjalainen (1999) which is resistant to the spurious data snooping results as shown in Brock et al. (1992) as it is based purely on the patterns of past data. Furthermore, Ready (2002) suggests that even in the case of presence of predictability in stock return data, methodology by Brock et al. (1992) does not seem sufficient enough to exploit this possibility to its full potential. However, when compared in an empirical study, all promising approach by Allen and Karjalainen (1999) fails to outperform the moving average trading rule by Brock et al. (1992). Since both studies used a different data set and different time periods, Ready (2002) suggests a possibility that the success of one approach and failure of the other might be explained due to the difference in market behaviour, and in particular the presence on non-linearity, during different time periods. This is further supported by Ready's (2002) own attempt to use Brock's et al. (1992) approach using a different time set and finding it to perform poorly. Finally, after providing empirical evidence against usefulness of Brock's et al. (1992) trading rule based on simple moving average in predicting daily returns, Ready (2002) is still unable to confirm that earlier results were due to data snooping as well as to reject the null hypothesis of no predictability. This study anticipates that the consideration of STAR-type non-linearity in comparison with linear ARIMA models, as well as the simple random walk, will shed light on possible improvement in assessing presence of predictability in daily returns.

Econometric forecasting and forecasting accuracy

Introduction to econometric forecasting

Forecasting in econometrics can be described as an attempt to predict the future, with an intention to improve the effectiveness of decision-making mechanisms (Holden et al., 1990). Forecasts are relevant and required due to the uncertainty of the future. The effect of uncertainty is especially evident in the case of decisions taken at a present time but the impact of which is experienced later in the future. In other words, the essential reason for forecasts is their usefulness. This is important as financial decisions often require long-term investment of resources; the outcome of which will heavily depend on future events. Hence, the accuracy of the forecast is directly linked with the overall utility, effectiveness and in most cases profitability. Doran (1999) describes forecasts as predictions, sometimes expressed as a probability, based on the knowledge of past behaviour where the extent of past trends consistency in future depends upon the forecaster's judgement. Hence, any forecast is in fact a statement of likelihood that an outcome will occur.

Montgomery et al. (1998) prompt that forecasting is the main application of many econometric models. Diebold and Mariano (1995) pointed out the importance of forecasts and, hence, forecasting accuracy in practical uses in science, which includes economics and finance. In addition, time-series modelling and forecasting accuracy are important for a wide range of disciplines including political science. Thus, Granato and Suzuki (1996) discuss the use of forecast encompassing methodologies in assessing issues of political behaviour and in particular voting behaviour. The researchers also highlighted the importance of explanatory accuracy of time-series models in the context

of political behaviour. Montgomery et al. (1998) carried out research into forecasting the US unemployment rate with displayed asymmetric cyclical behaviour during periods of economic contractions, highlighting the importance of accurate forecasts of this important social and political element of the economy. In the financial industry forecasts are needed for financial and investment planning, control of companies in terms of operational procedures, and other aspects of day-to-day activities of a company. Holden et al. (1990) point out that the careful planning of a company's operation depends on the accuracy of the forecast of the economy and related industry, thus denoting the importance of forecasting. Financial agents and industry members as well as governments extensively use macroeconomic forecasts in an attempt to make an informed, and thus the most acceptable, decision in a particular situation. Participants in the financial markets use forecasts in profit maximising activities when determining differences between the present and future values of assets. Practical uses of forecasting in financial markets include forecasting returns on various assets, risk assessment techniques such as value at risk (VaR)[3], volatility of returns or correlation between different stock market movements. Forecasting techniques are also heavily used in trading and hedging. Diebold and Mariano (1995) also suggested the forecasts should be used to guide decision makers rather than to rely on the results solely, since the test statistics of any forecast does not fully reflect its economic loss.

[3] Value at Risk (VaR) is usually used in portfolio risk assessment in an attempt to summarise the total risk of a portfolio in a single number to assist senior management. VaR is expressed as a percentage and represents a potential loss that will not exceed a specified level of confidence over a specified time period. For a more detailed description see Hull (2003).

Types of forecasting models

There are two main types of forecasting models: univariate time-series models and causal or structural models. Time-series models evaluate historical data of the underlying variable using statistical analysis. Whereas, causal models involve statistical examination of other explanatory variables that constitute an economic model used to explain the behaviour of the variable under consideration. More specifically, a univariate time-series model forecast is an attempt to model and predict financial variables based on the information contained within their own past values and past values of an error term. Structural models, on the other hand, are multivariate in nature and attempt to predict the behaviour of financial variables based on movements in the current and past values of the other explanatory variables. Or to put it differently, structural forecasting models simply attempt to predict future values by relating a dependent variable to one or more independent variables. Unlike univariate time series models, structural models have an underlying theoretical explanation of the variable's behaviour. Whereas time-series models intend to capture and model empirically observed features in the behaviour of the variable.

Pindyck and Rubinfeld (1998) state that the primary purpose of single-equation regression models, such as univariate time series, is forecasting, which allows the making of inferences about the likelihood of future events based on current and past observations. Univariate time series models are considered to be fast, cost effective and simple to apply. The information input required for this approach includes historical data of the underlying variable obtained at equal intervals. Although this type of model does not offer an explanation of the behaviour of the variable, it eliminates problems associated with the complexity of causal models. Moreover, univariate time series

models are considered to be the most popular type of economic model used for forecasting. These include such models as autoregressive (AR) and autoregressive moving average (ARMA) models, which can be used to produce forecasts by applying the standard Box and Jenkins procedure (Box and Jenkins, 1970).

As an alternative to univariate time-series models, there is another popular time-series forecasting technique known as a method of exponential smoothing and forecasting. According to Wagle (1965), the exponential smoothing method uses a weighted average of actual values from the current period and the previous periods in order to forecast expected values of the variable in the next period. It also uses the weighted sum of squared errors in order to award less weight to more distant values in the past periods since these will have less effect on the current and future values of the variable. The method of exponential smoothing can also account for seasonality and the presence of trend in the series. The weights used when implementing the exponential approach are selected using the mean square error (MSE) of prediction. Thus, the less the value of the MSE the better the weight is assigned. However, Wagle (1965) states simplicity as one of the major disadvantages of the exponential smoothing forecasting methodology, suggesting that critical economic variables that may explain the behaviour of the series are omitted from the model. In addition, Harvey et al. (1998) defines the exponential smoothing forecasting technique as *ad hoc*, claiming that these models are implemented without reverence to a defined statistical model therefore not taking into account any economic or historical issues involved in the formation of the series. In other words, the method of exponential smoothing can be criticised for its poor explanatory power, since it does not have any explicit statistical foundation.

On the contrary to the exponential smoothing methodology, Granger and Newbold (1986) pointed out the main advantages of the univariate forecast as these being quick and inexpensive to apply. They also advised that the forecasting errors, which are in essence indicators of the accuracy of the forecast, should be balanced with the costs of producing the forecast. In other words, it is not worth spending large resources, such as time and financial costs, on a payoff which is only a small increase in forecasting accuracy. Hence, the payoff has to be more than only marginally beneficial to justify a higher cost of the forecast. In addition, Harris and Sollis (2003) draw attention to the fact that forecasts made using simple linear univariate models are often sufficiently accurate, with more complex models being only marginally more accurate. Moreover, Brooks (2002) points out that time-series models can be used in situations when structural models are inappropriate. This is due to a number of reasons and can be viewed as a few advantages of time-series approach over structural models. These points include the possibility that the explanatory variable that is thought to influence the movements of the underlying variable might be unobservable or unmeasurable. In addition, the data for the explanatory variable and the underlying variable could be measured at a different frequency of observations. This is often the case when for the financial series of daily frequency the possible explanatory variable is thought to be a macroeconomic variable which is usually measured on a monthly basis. Additionally, Harris and Sollis (2003) point out further concerns regarding different issues involved in analysing financial time-series and macroeconomic time-series. One of these matters involves the differences in data frequency. Thus, financial time-series tend to have higher frequencies compared to the macroeconomic data. The financial time-series are also characterised by so-called 'long-memory' (Harris and Sollis, 2003) which implies

the dependency of the variables on the past observations over the long time horizon. In addition, the financial time-series is more prone to time-varying volatility than the macroeconomic series, creating volatility clusters in the series when presented graphically. If the data contains changing variance it is said to be heteroscedastic (Pindyck and Rubinfeld, 1998). There are various tests and methods for correction of heteroscedasticity of the data, however, these will not be considered in this chapter.

Moreover, the concern of differences between linear and non-linear models in forecasting encouraged examination into comparative performance of linear and non-linear forecasts. Thus, Montgomery et al. (1998) claimed non-linear models to outperform linear alternatives in predicting quarterly US unemployment rate on the basis of values of MSE, whereby the most sufficient statistical reduction in MSE was an indication of the most improved forecasting result. Montgomery et al. (1998) carried out research into the application of non-linear forecasting models performance during economic expansions and contractions, or in other words, an investigation whether non-linear models react to conditions during the economic boom and recession, using the example of the quarterly US unemployment rate. The results of the study show that the use of non-linear models significantly improve forecasting performance for the data series during economic contraction.

Clements and Smith (2001) carried out research into assessing the forecasting performance of SETAR models in comparison to the linear random walk model. The researchers are accepting the fact that there is clear evidence of the presence of non-linearities in the market variables, and apply a non-linear SETAR model to exchange rate forecasting. As a result, Clements and Smith (2001) find significant improvement in forecasts produced by the SETAR models over the simple random walk, however,

they point out that the use of traditional forecasting accuracy measures, such as root mean squared forecast errors, may significantly diminish these differences in models' performance and thus the superiority of a non-linear approach.

In addition, as a possible reason for frequent failure of linear forecasts Doran (1999) suggests the possibility of a non-linearity break, or in other words, discontinuity from the past trend. With some types of dynamic analysis it is possible to predict that certain non-linearity will occur, however such analysis cannot forecast when it will occur.

In-sample and out-of-sample forecasts

In order to avoid the problem of overfitting, Rapach et al. (2005) employ a variation of the forecast encompassing test in order to determine the best in-sample model before applying to the out-of-sample forecasting exercise, which is then subjected to the same test procedure. However, it has to be pointed out that in-sample fit, as well as an in-sample predictive ability of a model, does not necessarily imply out-of-sample predictive ability. Clark and McCracken (2005) suggest structural breaks as a possible explanation of differences between in-sample and out-of-sample forecasting ability. As a result, Clark and McCracken (2005) suggested in- and out-of-sample tests of forecasting ability in the presence of structural breaks. In addition, Clements and Smith (2001) advocate the use of non-parametric modelling in forecasting as this illuminates the possibility of model failure due to incorrect function form specification. In addition, Clements and Hendry (1998) suggest the assumption of constant parameters, as opposed to time varying, as one of the main reasons for macroeconomic forecasts to be characterised by a good in-sample fit while producing poor out-of-sample forecasts.

Thus, the researchers suggest that when the assumption of constancy of parameters fails, the in-sample fit provides a poor guide to out-of-sample forecasting performance, consequently recommending empirical models being able to account for structural breaks. Van Dijk and Franses (2003), on the other hand, advised that these results could be due to non-linearity detected in the data series being spurious, suggesting that other features of the data, such as heteroscedasticity, structural breaks and outliers, could have been mistaken for the presence of non-linearity. Moreover, non-linear models might demonstrate less successful forecasting results due to pure coincidental possibility that the forecasting period is not described by the non-linear regime. However, Van Dijk and Franses (2003) suggest that the main reason for poor performance of out-of-sample non-linear forecasts is an inappropriate forecast evaluation criteria.

Combining forecasts

Newbold and Granger (1974) carried out an investigation into univariate forecasting methods and found while all the models performed reasonably well, a combination of these generated far more accurate forecasting results. However, the researchers warned that a poorly constructed combined forecast might result in worse output than that of an individual forecast, hence, the performance of these should be carefully monitored. The study used the mean squared error as a measure of forecasting performance. The paper offered a view of combining forecasts as a useful tool used to increase the efficiency of a set of forecasts. However, the equal weighting approach used in the study was criticised for treating all the forecasts included in the combination as having equal informational content. Since the very early stages of the introduction of the equal

weighting method used in combining forecasts, heavy criticism was expressed by academics due to having major flaws in its theoretical grounds, however the method remains the most commonly used approach for combination of forecasts.

To the contrary of equal weighting technique, Guidolin and Timmermann (2009) applied a flexible forecast combination to US interest rates where the methodology allows for variable weights to be assigned to different models included in the forecast. The study confirmed combined regime-switching forecasts to outperform individual univariate forecasts on the basis of the RMSE statistic predominantly at short horizons.

Fang and Xu (2003) carried out a research into asset returns predictability by combining technical trade rule analysis and time-series forecasts. Using the rolling out-of-sample approach the study demonstrated the combination approach of two methodologies to produce a superior forecasting result compared to forecasts achieved by each approach individually. This effect, according to the researcher, is achieved due to both approaches being asymmetric during buy and sell periods, thus being able to capture different dynamics of the data's predictability. Thus, the trading rule approach performed better during the rising market, whereas time-series models performed better in falling markets.

Forecasting accuracy

The usefulness of a forecast is determined by such factors as its accuracy, or in other words, the difference between the forecasted and actual values, ease of application of the output and time required to produce the forecast as well as the cost of implementation. Consequently, forecasting accuracy is also assessed as means of a

comparison between competing forecasting models. Chatfield (1997) argues that the best method of comparison of forecasting accuracy depends on a range of factors including the context of the forecast, type of data considered and availability of analytical expertise. Moreover, Chatfield (1997) points out that the meaning of forecast superiority could be assessed differently, which might not necessarily be the least forecasting errors or forecast's ability to generate profit.

Doran (1999) also highlighted the importance of noise as a limit of forecasting. Noise in this instance is defined as a measure of error which in effect is the variance around the trend-line. Consequently, the greater the noise, the greater is the uncertainty in the accuracy of the forecast. There are various techniques of assessing the accuracy of forecasts, including measures of minimising the mean of forecasted errors, such as mean square error (MSE); Akaike's information criteria (AIC); Schwartz criteria; Diebold and Mariano (1995) test of equal forecast accuracy, and many others. However, none of these methods are considered to be the solely preferred technique, thus enforcing researchers to use a few different tests when comparing forecasting outputs.

Chatfield (1997) suggests that comparing competing forecasts should be performed on the results of out-of-sample forecasts as opposed to in-sample estimations. Moreover, Diebold and Mariano (1995) pointed out that the forecasting superiority of any model over alternatives on the basis of statistical measures does not necessarily imply that other models do not contain any additional information. Hence, Diebold and Mariano (1995) proposed a forecasting accuracy comparison test which is based on predictive performance unlike tests that assess the deviation between the forecasting model and the data. Their test of equal forecast accuracy proved to be applicable to a wide range of forecasting models. However, the researchers advised for the test to be used in

conjunction with other statistical measures and diagnostics for comparing models' forecasting performance.

Moreover, Van Dijk and Franses (2003) explain the apparent good in-sample fit combined with unsatisfactory out-of-sample forecast of non-linear models compared to linear models due to unsuitable model selection criteria and forecast evaluation techniques. As a result researchers recommend the weighted Diebold-Mariano test of equal predictive accuracy, which is based on a concept that different observations have different weights of importance within the dynamics of the time-series. Subsequently, according to Van Dijk and Franses (2003), an accurate forecast of extreme observations, or outliers, is essential as these data points could be important indicators of major changes in economic behaviour. Linear models are likely to forego these changes, while non-linear models are able to capture these extreme data points.

Makridakis et al. (1979) carried out an extensive research into forecasting accuracy measurement techniques. In order to minimise any bias potentially arising from using a single data set, the researchers employed 111 time-series obtained from different sources across different countries and industries, and different time periods as well as at different data frequencies. However, researchers themselves pointed out that since the majority of the time-series data was monthly, during the 1970s and came from French sources, it was not a random data set. After generating twelve points of forecast for each series, Makridakis et al. (1979) found that when employing a single forecasting method the accuracy results differ depending on the choice of the loss function. Makridakis et al. (1979) differentiate between model fitting and forecasting, suggesting that these will require different loss functions. Thus, while naïve and exponential smoothing model forecasts result in smaller forecasting errors due to these methods hedging forecasts

towards the mean, ARMA models follow data patterns more closely, however resulting in larger forecasting errors when the forecast fails.

Overall, Martens et al. (1998) pointed out the lack of appropriate evaluation criteria for non-linear time-series models. Clements and Smith (1999) mention that the majority of the studies comparing forecasting performances of linear and non-linear models base their conclusions on the results of forecasting error magnitude style tests, such as mean squared error (MSE). Extensive literature of assessment of forecasting accuracy is somewhat limited in considering non-linear forecasting techniques and appropriate tests of accuracy for such models.

Quantitative and subjective forecasting approaches

Makridakis et al. (1979) provide a comprehensive review of literature on the subject of comparative accuracy of quantitative methods of forecasting against judgemental forecasts. Subjective forecasts sometimes might outperform the statistical type forecasts, nonetheless, such forecasts are rare and less accessible than econometric forecasting. While there is a lack of studies supporting superiority of judgemental forecasts over quantitative methods, with an overwhelming number of studies supporting the latter approach, there is still a debate over a single preferred quantitative forecasting approach. Moreover, Andersen (1977) argues that despite the fact of numerous techniques being used for forecasting, including various computer packages, it is evident that sometimes a subjective forecast by a dealer or a trader may be more accurate than that produced by a statistical method. Anderson (1977) contributes this phenomenon to the fact that a dealer (salesman, broker, entrepreneur, etc.) will usually

include additional information when carrying out a forecast. This information might be vital in a particular case, however, it might be too specific to an individual situation or industry to be known to a forecaster when carrying out a theoretical prediction. It also might be the case that it is too difficult to include such information in the mathematical model due to, for instance, measurement difficulties. Wagle (1965) also suggests that statistical forecasting models should be considered as a supplementary aid to policy-makers, rather than a sole tool, given that personal experience and subjective judgement have proved to be good methods of forecasting outperforming purely mathematical models. Moreover, Clements and Hendry (1998) point out that the failure to correctly predict major dramatic economic changes in the UK followed by the recession in the 1990s, consumer boom in the late 1980s and patterns of post-war consumption, has lead to overall reduced levels of confidence in macro-economic forecasting methods.

The supporting argument of subjective forecasts over statistical forecasts is consistent with the fundamental approach as opposed to pure technical analysis of financial markets. Technical analysis is based on the assumption that stock markets move in persistent trends, and thus examine past market data with the purpose of estimating future trends. Fundamental analysis, on the other hand, uses economic data rather than financial market values, and hence will base the predictions of market movements using a subjective form of forecasts. Most criticism of these techniques come from studies considering trading rules in order to examine the presence of predictability in stock market prices. This section of the chapter will offer a glance at the literature concerned with stock market predictability and, thereof, uses and applications of non-linear models in these investigations.

Newbold and Granger (1974) advised that more complex forecasting models with abilities to incorporate qualitative and quantitative information about the data as well as the past and current values are preferred theoretically, however these might not be readily available and it might still be a difficult task to incorporate such information into a forecasting model. Univariate time-series forecasting models, on the other hand, are quick and inexpensive to implement and, besides being often used as a benchmark, may produce forecasts of sufficient accuracy. Allen and Karjalainen (1999) suggest that the majority of empirical investigations into technical trade rules found this technique unable to generate profit. The researchers arrive at the same conclusion. After employing a genetic learning algorithm using the daily S&P 500 index data and accounting for transaction costs, the results of the study suggest that it is not possible to achieve after transaction costs profits by means of technical trading approach thus implying efficiency of the stock market.

Leitch and Tanner (1991) questioned the reasoning behind many profit-maximising firms investing in economic forecasts while conventional error average measures indicate naïve forecasts to perform as well as or better than professional forecasts. The researchers carried out an investigation into the relationship between profitability of forecasts and conventional error measurements of forecasting performance using interest rate data. In addition, Leitch and Tanner (1991) pointed out that there is a possibility that the firms employing professional forecasters might not be using the lowest mean of absolute forecasting error as an indication of the preferred forecast, which is most likely due to the absence of a consistent relationship between forecast profitability and statistical error-magnitude measurement. However, researchers did find

a strong statistical association between the directional accuracy and profitability of a forecast.

Makridakis et al. (1979) state that besides research proving that quantitative forecasts are less cost and time consuming than the judgemental approach, the subjective forecasts pose a number of difficulties. These include the lack of application of valid principles and solid theoretical bases, which in turn translate into basing forecasts on irrelevant information, anchoring effect, where decision makers base their evaluations on pre-existing perceptions instead of logical relevance and facts, and perception biases, where specific cases tend to be generalised. All of these factors contribute towards judgemental forecasts to be highly unreliable. Chatfield (1977) suggested that in practical terms successful forecasting implementation could be achieved when forecasting is considered in coalition with the management process and be constantly revised and corrected.

2.3. Methodology

Introduction

Time-series data represents a sequence of observations on a single variable obtained over time. In the case of time-series analysis the order of the data becomes an important issue. Granger and Newbold (1986) argue that the main objective of time-series analysis

involves construction of a model which exhibits similar properties to the observed series allowing the researcher to make inferences about the behaviour of a series for the main purpose of hypothesis testing and forecasting. The main practical application of time-series models is their use in forecasting as well as explaining the behaviour of the underlying variable.

As a first step in time-series analysis Chatfield (1977) recommended to plot the data against time, since this could provide a useful visual confirmation of certain features of the data such as trend, seasonality, discontinuities and outliers. Moreover, Chatfield (1977) pointed out that different technical approaches are required to analyse different types of data, thus suggesting that analysis for short-run non-stationary data financial time-series, for instance, would be different to analysis required for long-run stationary economic series. Finally, Chatfield (1977) emphasised the use of common sense when applying time-series analysis, since a considerable degree of subjective judgement is invaluable in statistic investigations. Similarly, in a later paper, Chatfield (1997) advises clarification of objectives and potential purpose of the forecast as starting points of any forecasting exercise, followed by plotting the time-series data against time as the time plot might assist in choosing an appropriate model for fitting and forecasting the data.

This section of the chapter will concentrate on presenting methodologies involved in time-series analysis applied in the subsequent empirical chapters: Chapter 3, Chapter 4 and Chapter 5. Thus this chapter will consider approaches involved in estimating linear and non-linear models including an error-correction model, testing for stationarity using the standard unit root tests as well as stationarity tests in the context of non-linearity. Furthermore, econometric forecasting methodology will be discussed, followed by tests of time-series forecasting accuracy including statistical and economic loss functions.

Linear models

Linear models discussed in this paper include a random walk process, time-series regression and ARMA-type processes.

Random walk model

The notion of a stochastic process is an important issue in time-series analysis. Gujarati (2003) defines random, or stochastic, process as a collection of random variables ordered in time. An example of a random process is a random walk model, which can be described by the following equation:

$$y_t = y_{t-1} + \varepsilon_t \tag{2.1}$$

or by a random walk model with a drift (δ):

$$y_t = \delta + y_{t-1} + \varepsilon_t \tag{2.2}$$

where, y_t and y_{t-1} are the dependent variables at time t and $t-1$ respectfully, and ε_t is a random disturbance term. In the context of financial markets the random walk model simply states that a price of a stock, for instance, today (y_t) is equal to its price

yesterday (y_{t-1}) plus a random shock today (ε_t). It is also worth noting that a random walk model is an example of a non-stationary stochastic process.

Linear regression

Simple linear regression models the degree of linear association between variables (Brooks, 2002), which can be extended to use in time-series:

$$y_t = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \cdots + \alpha_i x_{it} + u_t \tag{2.3}$$

where the dependent variable (y_t) is regressed on the explanatory variable (x_t), with i number of observations, α_0 ... α_i are regression coefficients and u_t is the error term.

ARMA process

According to Pindyck and Rubinfeld (1998), time-series models, such as autoregressive process, AR(p), and moving average process, MA(q), where p and q represent lag lengths, are designed to describe the movement of a time-series by relating the series to its own past values while attempting to minimise the weighted sum of current and lagged random disturbance terms.

Thus, the moving average process is described in terms of weighted sum of current and lagged random disturbances, where each observation y_t of the moving average process

of order q, MA(q), is generated by weighted average of random disturbance terms going back q periods:

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} \tag{2.4}$$

where $\theta_1, \ldots, \theta_q$ are the parameters of the moving average model and u_t is a disturbance term assumed to be identically independently distributed (*i.i.d.*), in other words, the disturbances follow the white noise process. A moving average process, MA(q), has a memory only of length q, thus limiting the time horizon of a forecast up to the step q, as all forecasts of more than q steps ahead have a tendency to collapse to the intercept or to zero in the case of no constant in the moving average process.

Distinct from the moving average process, an autoregressive process has infinite memory, thus allowing for the forecasts to be made for long-time horizons. In the form of an equation the autoregressive process of order p, AR(p), can be described as a process where the current observation y_t is the result of a weighted average of past observations going back p periods including a current random disturbance term, u_t:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + u_t \tag{2.5}$$

where μ is a constant term accounting for the mean of the stochastic process. The autoregressive models are also based on the assumption of the disturbance terms being a white noise process.

The autoregressive moving average (ARMA) process can be described as a combination of moving average and autoregressive processes as it combines deterministic characteristics of both processes. Hence, an ARMA process is a function of its past values and lagged random disturbance terms, thus incorporating the AR process component, as well as a current disturbance term, thus including the MA process component. The general form of the ARMA (p, q) model is as follows (Brooks, 2002):

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} + u_t \tag{2.6}$$

The integrated autoregressive moving average (ARIMA) models are used in order to model non-stationary time-series. The order of the integration in integrated autoregressive moving average models specifies the number of times the series should be differenced in order to achieve stationarity. ARIMA models are extensively used when analysing time-series due to its relative flexibility.

Box and Jenkins (1970) proposed an approach to time-series analysis whereby the procedure is developed for univariate forecasting based on the ARMA process. Chatfield (1977) points out that while AR, MA and ARMA models have been studied extensively, the Box-Jenkins procedure provided a systematic approach to modelling and forecasting these types of models, offering a comprehensive methodology of model identification and checking with a possibility to extend the approach to non-stationarity and seasonal data. The ease of application and reliability of the methodology secured the Box-Jenkins approach as the most widely accepted method of ARMA modelling.

The procedure involves three steps determining the of order of the model (p, d, q), where p is the order of the AR component and indicates the number of autoregressive parameters (ϕ), d is the number of times the data series is differenced in order to achieve stationarity, and q is the MA order indicating the number of parameters of the moving average component (θ). The three stages involve identification, where values of p, d, and q are chosen; estimation, where coefficients of the model are obtained by employing standard statistical methods; and diagnostic checking of model adequacy, where the residuals of the model that was estimated at stage two of the procedure are tested for significance. A requirement of an estimation of a correct model is complete when the analysis of the residuals certifies that errors of the estimated model are independent and identically distributed, or, in other words, the error term is random and follows a white noise process. Should the diagnostic check reveal inadequacy of the estimated model, the whole procedure of model building is reiterated starting from the first stage until an adequate model is estimated.

However, while the Box-Jenkins approach allows a degree of flexibility in the choice of a model, Chatfield (1977) suggested that the flexibility also allows for a possibility to choose a misspecified model. Moreover, while original procedure required analysis of an autocorrelation function (ACF) and a partial autocorrelation function (PACF) at the identification stage, in practice it appeared to be difficult to identify the behaviour of ACF and PACF of the series by comparing these plots to theoretical functions. Cho (2002) pointed out that parameters estimated by observing the ACF and PACF can be subjective and hence lead to an unreliable and inaccurate estimation. Similarly, early studies, such as Wagle (1965) considered ARMA modelling a poor forecasting tool due to a complex estimation procedure. However, significant improvements were made ever

since in order to improve and expand the original Box-Jenkins methodology. Thus, the coefficients at the identification and estimation stages of the procedure are estimated using the Akaike's information criteria (AIC) or Schwarz's Bayesian criterion (SBC), which provides more reliable statistical reference and avoids the subjectivity of the ACF and PACF interpretation. According to Brooks (2002), information criteria are a function of the residual sum of squares and accounts for the loss of degrees of freedom that occurs when extra parameters are added to the model. In the context of ARMA models specification, parameters which minimise the value of the information criteria are considered to be the correctly specified.

$$AIC = ln(\hat{\sigma}^2) + \frac{2k}{T} \tag{2.7}$$

$$SBIC = ln(\hat{\sigma}^2) + \frac{k}{T} lnT \tag{2.8}$$

where $\hat{\sigma}^2$ is the residual variance, k is the total number of parameters estimated, which in the context of ARMA model is the sum of lag lengths for the AR and MA components and unity ($k = p + q + 1$), and T is the sample size.

Models described above are examples of the most commonly used linear models. However, according to Campbell, Lo and MacKinlay (1997), the payoffs to options as well as investors' willingness to trade off returns and risk are characterised by non-linear functions. Similarly, it can be argued that most financial data can be described by non-linear functions rather than linear models.

Non-linear models

According to Gujarati (2003), non-linear models are such models that are non-linear in parameters regardless of whether the variables are linear or not. In addition, genuine non-linear models cannot be linearised in its parameters unlike most linear models that only appear to be non-linear. Hence, Gujarati (2003) entitles such models as intrinsically non-linear regression models. Gujarati (2003) points out that estimation of non-linear regression models is often an interactive process or, to put it differently, involves a trial-and-error method. In other words, initial estimation of values for model parameters are based on prior experiences or prior empirical work as opposed to simple fitting of a linear model using OLS.

TAR model

Threshold autoregressive (TAR) models are a class of non-linear time-series autoregressive models. Unlike standard autoregressive models, TAR models allow for locally linear approximation over different states (Brooks, 2002). The TAR model contains a first order autoregressive process in each of the specified regimes. The number of thresholds for a model will always be the number of regimes minus unity. For instance, a model containing only one threshold will have two regimes. Naturally, general TAR models allow for more than two regimes and more than one lag in the autoregressive process. However, for the ease of illustration of the process of TAR models, one threshold TAR will be considered here.

Tsay (1989) explains the TAR model as a process with at least two regimes with different linear autoregressive models under each regime. In this model, the threshold value r acts as a point of reference whereby the state-determining variable lagged k periods and denoted s_{t-k} can take values that are either below or above the threshold value. Thus, the dependent variable y_t is specified to follow a first order autoregressive process with an intercept coefficient μ_1 and autoregressive coefficient ϕ_1 if the value of the state-determining variable is below the threshold value. If the value of the state-determining variable is greater or equal to the value of the threshold, then y_t is specified to follow an autoregressive process with the intercept μ_2 and the autoregressive coefficient ϕ_2 (Brooks, 2002).

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} & if\ s_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} & if\ s_{t-k} \geq r \end{cases} \tag{2.9}$$

where u_{1t} and u_{2t} are the error terms for each autoregressive process.

The state-dependent variable (s_{t-k}) is the variable that is thought to influence the dependent variable (y_t) to shift from one type of behaviour to another, i.e. from one regime to another. This variable is determined by considering issues of financial and economic theory.

SETAR, or self-exciting TAR model, is the case where the state-determining variable is the variable under consideration, i.e. the dependent variable itself, $s_{t-k} = y_{t-k}$. In this case, it is the lag of y_t itself that determines the current regime this variable is in.

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} & if\ y_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} & if\ y_{t-k} \geq r \end{cases} \tag{2.10}$$

In general, threshold models can be extended to models with higher number of lags of the dependent variable as well as the number of states. Also, the number of lags in each regime can be different. Hence, the general formula for the TAR model is as follows (Brooks, 2002):

$$y_t = \sum_{j=1}^{J} I_t^{(j)} \left(\phi_0^{(i)} + \sum_{i=1}^{P_i} \phi_i^{(j)} y_{t-i} + u_t^{(j)} \right) \quad , \quad r_{j-1} \leq s_{t-k} \leq r_j \tag{2.11}$$

where

$I_t^{(j)}$ = indicator function for the *j-th* regime. The indicator takes the value of unity if the underlying variable is in state *j,* and the value of zero otherwise.

s_{t-k} = observed variable that determines the switching point.

$u_t^{(j)}$ = a zero mean *i.i.d.* error process.

TAR models are characterised by discrete transitions between regimes. In other words, under a TAR model the dependent variable is either in one regime or the other. This is on contrary to Markov switching models where the dependent variable is in all of the states with different probability of being in either one at each point in time.

When applying model building in practice, the data should be tested for the presence of non-linearity, or in other words, testing linear models against a simple non-linear alternative before applying more complex non-linear models. Chan (1990) points out that the main discussion involving TAR models is built around limited availability to test whether a TAR model fits a given data better than linear autoregressive (AR) model. There are a number of tests for detecting the presence of threshold autoregressive non-linearity in time-series data. Petruccelli and Davies (1986) introduced a portmanteau test to detect a specific class of state dependent models, namely a self-exciting threshold autoregressive structure. A portmanteau test is based on examining autocorrelations of squared residuals from a linear fit, where the levels of significance are based on the asymptotic Chi-squared (χ^2) distribution of the test statistic for the linear process (McLeod and Li, 1983). Tsay (1989) also proposed a test for threshold non-linearity which is based on arranged autoregression and predictive residuals, similar to the portmanteau test by Petruccelli and Davies (1986). In essence, the test is a combination of non-linearity tests by Keenan (1985), Tsay (1986), and Petruccelli and Davies (1986). In the advantage of the proposed test it is a very simple procedure allowing wide practical applications, as its asymptotic distribution approximates to the *F*-distribution. Tsay (1989) found the *F*-statistic more dominant than the portmanteau test in assessing data for non-linearity in most of the cases, but not universally. Tsay (1989) points out that there is still a debate over finding the optimal test.

Estimation of TAR models involves estimation of model parameters (ϕ_i, r_j, d, p_j). This is a difficult process mainly due to the fact that the parameters cannot be determined simultaneously, hence the values chosen for one of the parameters are most likely to

influence the estimates for the other parameters. Initially, Tong (1983; 1990) proposed a complex non-parametric lag regression procedure for estimating thresholds (r_j) and a delay parameter *d*. The most valid method of estimation is to estimate threshold values using the non-linear least squares (NLS) optimisation procedure. Nevertheless, this method is not feasible as the relationships between the variables are discontinuous in the thresholds and, hence, cannot be estimated at the same time with other parameters of the model. On the contrary, Tsay (1989) suggests using ordinary least squares techniques since TAR models on the whole consist of a set of linear models. Hence, modelling procedure for threshold models proposed by Tsay (1989) involves four steps and is based on simple linear regression techniques. A further method suggests using a grid research procedure that intends to minimise residual sum of squares over a range of values of the threshold(s) for the model under consideration.

Nevertheless, Tsay (1989) points out that the TAR model was not widely applied due to the issues concerning modelling procedure and difficulties in identifying the threshold variable and estimation of the threshold values. The model estimation procedure drawn by Tong and Lim (1980) is very complex involving intense computing stages. This procedure also does not provide the diagnostic statistic necessary to ensure the need of a threshold model for a given data set. Hence, Tsay (1989) proposed a model-building procedure for TAR models that could be applied in practice, including a test statistic for testing threshold non-linearity which is derived by a simple linear regression. Tsay (1989) also used supplementary graphic devices for an identification of the number of potential thresholds.

Further, estimation of TAR models required determining the threshold order, i.e. the lag length. The simplest method of determining the lag length for the autoregressive

component for each of the regimes is to assume that the same number of lags for each of the regimes. The lag length itself in this case will be chosen using a standard approach for determining the lag length for a linear autoregressive model. However, despite the simplicity of implication, it is unlikely that the same number of lags for each regime would be sufficient in describing the data which is drawn from different regimes. Moreover, this method undermines the whole concept of threshold models whereby the data has different behaviour in different states. An alternative method involves simultaneous selection of the lag length for each regime using the information criterion. However, Franses and van Dijk (2000) pointed out that in practice it is likely that the model will be resident in one particular regime for a considerably longer period of time compared to other regimes. For this reason the information criterion will be able to accommodate and consider such behaviour. Consecutively, Tong (1990) suggested modified Akaike's information criteria (AIC) that weights the residual variance for each regime by the number of observations in that regime hence avoiding the dilemma described above.

In addition, the delay parameter, d, can be determined in the same principal as the lag length for each regime using an information criterion. However, the addition of this new dimension to parameters estimation will result in the increased number of potential models to be estimated. Hence, in practice the value of the delay parameter is normally set to unity due to theoretical explanations. Thus, Kräger and Kugler (1993) suggested that in financial markets the recent past value of the state-dependent variable is more likely to influence the current state than the value from two, three, etc. periods ago.

STAR models

Teräsvirta and Anderson (1992) in their paper assumed that any non-linear time-series can be represented by a smooth transition autoregressive (STAR) model. STAR models, unlike standard TAR models, allow for a more gradual transition of the dependent variable between regimes. The regime indicator in these models is a continuous function rather than an on-off switch (Brooks, 2002). Extensions of the STAR model considered in this paper include logistic STAR (LSTAR), exponential STAR (ESTAR) and asymmetric exponential STAR (AESTAR) models. The general STAR model for the dependent variable r_t and y_{t-i} as an explanatory variable is represented as follows:

$$r_t = \pi_0 + \sum_{i=1}^{p} \pi_i y_{t-i} + \left(\theta_0 + \sum_{i=1}^{p} \theta_i y_{t-i}\right) F(s_{t-d}) + \varepsilon_t \tag{2.12}$$

where $F(s_{t-d})$ is the transition function with s_{t-d} as the transition variable which determines the switching point, d is the delay parameter and ε_t is an error term. π_i and θ_i are the autoregressive components of the model.

The exponential STAR (ESTAR) is an extension of the standard STAR model which allows the differentiation between dynamics of the time-series caused by the different magnitude of the explanatory variable utilising the following transition function:

$$F(s_{t-d}) = 1 - exp(-\gamma(s_{t-d} - c)^2/\sigma^2(s_{t-d}))\ , \qquad \gamma > 0 \tag{2.13}$$

where γ is the smoothing parameter, c is the transition parameter and σ is the variance of the transition variable.

Further extension of the STAR model, the logistic STAR (LSTAR) model, on the other hand, captures time-series dynamics that occur as a result of different signs of the determinant, i.e. positive or negative values of the explanatory variable:

$$F(s_{t-d}) = \left(1 + exp(-\gamma\,(s_{t-d} - c)/\sigma(s_{t-d}))\right)^{-1}, \qquad \gamma > 0 \tag{2.14}$$

With regard to the LSTAR model the delay parameter d is assumed to be unknown, whereas the autoregressive order is known (Luukkonen et al, 1988). However, in practice, the order of the autoregressive part of the model is often unknown and has to be estimated. Thus it is suggested that when attempting model building the appropriate order of the linear AR model should be specified first. This can be achieved by applying usual model selection techniques, such as information criteria, including AIC and SBIC.

Similarly, asymmetric ESTAR (AESTAR) process proposed by Sollis et al. (2002) models different speeds of adjustment within the mean reversion system, so that the AESTAR model for time-series variable y_t, at T number of observations, reverts to the mean μ, which can be described in the form of variables' deviations from the mean ($z_t = y_t - \mu$) as follows:

$$\Delta z_t = \alpha S_t(\gamma_1, \gamma_2, z_{t-1}) + \sum_{i=1}^{k} \beta_i \Delta z_{t-i} + \varepsilon_t \tag{2.15}$$

where the asymmetry is described by the logistic transition function $S_t(\gamma_1, \gamma_2, z_{t-1})$, which allows for different speeds of mean reversion, γ_1 and γ_2, using the Heaviside indicator, I_t, with k lagged differences:

$$S_t(\gamma_1, \gamma_2, z_{t-1}) = [1 + exp\{-\gamma_1^2 z_{t-1}^2 I_t - \gamma_2^2 z_{t-1}^2 (1 - I_t)\}]^{-1} - 0.5 \tag{2.16}$$

$$I_t = 1 \; if \; z_{t-1} > 0 \tag{2.17}$$

$$I_t = 0 \; if \; z_{t-1} \leq 0$$

Naturally, the asymmetric function (2.16) collapses to the symmetric model when $\gamma_1^2 = \gamma_2^2$.

Davies et al. (1988) outlines two main methods of identifying and fitting STAR models. One method is based on CUSUM (cumulative sum) tests, whereas the other method is based on a likelihood ratio test (LRT). The analysis procedure for identifying non-linearity based on the CUSUM test contains four steps (Petruccelli and Davies, 1986). The first step of the procedure involves carrying out the actual CUSUM test in order to select values of k (lag length) and d (delay parameter). The series will be said to be linear if none of the values are selected. *Vmask* and *runs* tests are used in the second stage of the procedure to locate initial threshold estimates. The third step involves fitting the selected models and computing their SBIC. The last stage of the process

involves assessing the threshold values and returning to the third stage if the local minimum of SBIC has not been achieved.

Test procedures based on a likelihood ratio test are more time consuming compared to the CUSUM test. Due to such complication Davies et al. (1988) suggest to use only one threshold with this procedure. The procedure consists of three stages and involves obtaining the least squares estimate of the threshold value at first. The first stage of the process also involves fitting the model and calculating the value of its mean squared error (MSE), i.e. MSE (k, d). The second step involves choosing the parameter estimates for k_0 and d_0 such that MSE (k_0, d_0) is the minimum MSE for all k and d found in stage one. Once estimates are found the likelihood ratio statistic can be determined for this model against the null model, which in this case is AR (k_0). The last stage of the procedure consists of assessing the significance of the likelihood ratio test statistic by simulating n number of observations from the null AR (k_0) model and repeating the process from stage one.

However, none of these proposed procedures seem to be flawless and a number of researchers have suggested various alternatives. For instance, Luukkonen et al. (1988) proposed three tests for testing linearity against STAR models which seem to be more powerful than the CUSUM tests, especially in the case of testing against SETAR models. Chan and Tong (1986) suggest a likelihood ratio test statistic for testing linearity of SETAR models. However, as pointed out by Luukkonen et al. (1988), due to irregularity of the likelihood function the statistic should be determined separately for each application. As a suggestion to this problem, Luukkonen et al. (1988) proposed a set of tests which allows to test for non-linearity in a whole class of STAR models.

These tests are based on the asymptotic χ^2 distribution and seem to have realistic power.

Petruccelli and Davies (1986) used a portmanteau test, which is a CUSUM-type test, for testing threshold autoregressive non-linearity. This type of test is based on the predictive residuals of ordered autoregressions. However, Luukkonen et al. (1988) have compared three tests developed in their paper to the CUSUM test by Petriccelli and Davies (1986). As a result the so called third-order test procedure proposed in the paper was found to be a reasonable alternative with significant computational advantage. This test can also be applied to STAR models in general, including LSTAR. In addition, these tests were found to be more practical compared to the one developed by Chan and Tong (1986). Luukkonen et al. (1988) also proposed designing a simulation experiment in order to observe behaviour in small samples.

Luukkonen et al. (1988) suggested Lagrange multiplier (LM) tests that can be used when considering non-linear smooth transition autoregressive models. However, Chan (1990) pointed out that the Lagrange multiplier test cannot be used in the case of TAR models due to the discontinuous nature of its autoregressive function.

Chan (1990) proposes a test statistic λ which approximates to the (conditional) likelihood ratio test when the noise term follows normal distribution. In essence, the λ statistic is the normalised reduction in the sum of squares due to the partial linearity of the set of autoregressive functions. Chan (1990) has found the λ statistic in general to be more powerful comparing to Petruccelli and Davies' (1986) portmanteau test.

Error-correction model

The concept of cointegration is based on the fact that certain economic variables appear to move together and do not diverge from each other dramatically in the long-term thus forming a cointegration relationship. These variables may drift apart in the short-run, however will be pulled back to the long-term equilibrium by the economic forces within the market mechanism. The concepts of the cointegration and error-correction model are very closely linked, whereby cointegrating variables belong to an economic system with a long-run equilibrium, which in its turn can be described by the error-correction model, so that the model must exist if two variables are cointegrated. Similarly, the ECM generates series and is used in the testing stage of the Engle-Granger cointegration procedure. The definition by Engle and Granger (1987) states that two variables are cointegrated if each of these individual variables have the same order of integration, i.e. they need to be differenced the same number of times to achieve stationarity, and a linear combination of these variables is stationary, *I(0)*. Whenever two cointegrating variables diverge from each other the economic forces will tend to correct the equilibrium error, and the adjustment back to the equilibrium is described by the error-correction model. The non-linear error-correction model is able to describe the different dynamics that are characteristic for the long-run and short-run horizons, or in other words the process of adjustment to the long-run equilibrium. The original paper by Engle and Granger (1987) introduced the definition of a cointegrating process, and since then was extensively referenced and extended.

The standard error-correction model (ECM) models long-run equilibrium relationship between first differenced and lagged cointegrating variables in the following form:

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 (y_{t-1} - \gamma x_{t-1}) + u_t \tag{2.18}$$

where $(y_{t-1} - \gamma x_{t-1})$ is the error-correction term, which should follow the *I(0)* process if y_t and x_t are cointegrated with cointegrating coefficient γ. In other words, in the presence of a valid long-run equilibrium relationship the error-correction term will be stationary. In the above equation the cointegrating coefficient γ in fact defines the long-term relationship between x and y. β_1 describes the short-run relationship between changes in the x and changes in the y, whereas β_2 represents the speed of adjustment back to the equilibrium. The model can be estimated using the OLS procedure and can have an intercept in the cointegrating term and/or in the model. In addition, Brooks (2002) reminds that ECM can be estimated for more than two variables.

The definition of the cointegration process given in the seminal paper by Engle and Granger (1987) is far more complex, however, definition given above is adequate for understanding the basics of the concept and sufficient for the purpose of the present study.

As it was demonstrated earlier in the chapter, Section 2.2, cointegration techniques are used in context of the present value model. Moreover, this study intends to apply non-linear error-correction model techniques to the forecasting exercise of the stock returns using the dividend yield and the price-earnings ratio as determinant variables. The simple error-correction model would take the following form (McMillan, 2004):

$$r_t = \alpha_1 (p_{t-1} - \beta_0 - \beta_1 d_{t-1}) + \varepsilon_t \tag{2.19}$$

where r_t, p_t, d_t represent returns, prices and dividend yield respectively, $(p_{t-1} - \beta_0 - \beta_1 d_{t-1}) = z_{t-1}$ is the error-correction term, coefficient α_1 is the speed of adjustment to the equilibrium and ε_t is an error term.

A threshold autoregressive (TAR) error-correction model is a non-linear extension of the original Engle and Granger's (1987) model used in order to capture a non-linear adjustment mechanism (Enders and Granger, 1998; Enders and Siklos, 2001; McMillan 2004).

$$\Delta r_t = I_t \rho_1 r_{t-1} + (1 - I_t)\rho_2 r_{t-1} + \varepsilon_t \quad , \qquad I_t = 1 \; if \; r_{t-1} \geq c \tag{2.20}$$

where c is the threshold value and the Heaviside indicator function I_t is defined as $I_t = 1$ if $r_{t-1} \geq c$, or zero otherwise

Furthermore, ESTAR error-correction model allows for smooth transition between regimes and thus represents a more realistic economic model of the dividends-prices relationship.

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 - exp(-\gamma(s_{t-d} - c)^2/\sigma^2(s_{t-d}))\big) + \varepsilon_t \tag{2.21}$$

where the parameters and the transition variable s_{t-d} change symmetrically with the threshold value c, so that if $y \to \infty$ or $\gamma \to 0$, the equation becomes linear.

The LSTAR error-correction model also allows for smooth transition and captures the asymmetry in the adjustment process followed by the different sign of the determinant:

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 + exp(-\gamma(s_{t-d} - c)/\sigma(s_{t-d}))\big)^{-1} + \varepsilon_t \quad (2.22)$$

The AESTAR error-correction model allows smooth transition between regimes characterised by different speeds of adjustment to the equilibrium:

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\left(1 + exp\big(-\gamma_1^2 s_{t-1}^2 I_t - \gamma_2^2 s_{t-1}^2 (1 - I_t)\big)\right)^{-1} + \varepsilon_t \quad (2.23)$$

where γ_1^2 and γ_2^2 are speeds of mean reversion, and I_t is the Heaviside indicator:

$$I_t = 1 \; if \; s_{t-1} > 0 \quad (2.24)$$

$$I_t = 0 \; if \; s_{t-1} \leq 0$$

Unit root tests

In order to apply the framework and to carry out a forecasting exercise, the relevant data is required to be tested for presence of stationarity. Linear stationarity can be tested using the Dickey-Fuller test or augmented Dickey-Fuller test, whereas non-linear stationarity is tested using tests specially modified for these purposes.

Linear unit root tests

The standard Dickey-Fuller (DF) unit root test is sufficient enough to test linear stationarity and takes on the following form:

$$\Delta y_t = \psi y_{t-1} + u_t \tag{2.25}$$

where Δ is the difference operator, ψ is the test statistic, and u_t is a white noise error term. The null hypothesis of unit root ($H_0: \psi = 0$) is tested against the alternative of stationarity ($H_1: \psi < 0$). Since the statistic ratio does not follow the standard F-distribution under the null hypothesis the test statistic is compared to specially tabulated Dickey-Fuller critical values.

The augmented Dickey-Fuller (ADF) unit root test accounts for autocorrelated error terms, since the standard DF test is only valid if the disturbance term follows a white noise process. In these circumstances the ADF test is preferred as a more general procedure for testing presence of linear non-stationarity. Similarly to the standard DF test, the ADF test statistic follows a non-standard distribution and thus Dickey-Fuller critical values are used. A standard equation for the ADF unit root test as follows:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{2.26}$$

where ψ is the test statistic, p is the number of lags of the dependent variable and u_t is an error term. Test lags of the dependent variable or the augmented test are chosen on the basis of frequency of data combined with a previous knowledge from similar studies. The procedure is testing a null hypothesis of unit root against an alternative of stationarity.

$$H_0: \psi = 0, \qquad series\ contains\ unit\ root$$

$$H_1: \psi < 0, \qquad series\ is\ stationary$$

The DF and ADF tests are the most commonly used unit root tests, however, these are unable to detect non-linear stationarity and can lead to misspecified modelling and hence incorrect results. Consequently, the tests will fail to reject the null hypothesis of the unit root for time-series displaying STAR-type non-linearity which in reality might be globally stationary. As a result, a number of alternative unit root tests were developed in order to account for non-linear stationarity.

Non-linear unit root tests

Kapetanios et al. (2003) developed a relatively easy to apply procedure for testing the presence of non-stationarity in time-series data using exponential smooth transition autoregressive (ESTAR) processes, and the proposed test was found to have better

power comparing to the standard DF test. The test is based on the following form of the ESTAR model:

$$y_t = \beta y_{t-1} + \gamma y_{t-1}[1 - exp(-\theta y_{t-d}^2)] + \varepsilon_t \tag{2.27}$$

which can be further reparameterised as:

$$\Delta y_t = \phi y_{t-1} + \gamma y_{t-1}[1 - exp(-\theta y_{t-d}^2) + \varepsilon_t] \tag{2.28}$$

where $\phi = \beta - 1$.

The procedure developed by Kapetanios et al. (2003) is based on a specific ESTAR model where ϕ equals to zero ($\phi = 0$) and the delay parameter d is set to unity ($d = 1$).

$$\Delta y_t = \gamma y_{t-1}\{1 - exp(-\theta y_{t-1}^2)\} + \varepsilon_t \tag{2.29}$$

Hence, the procedure involves testing the null hypothesis of parameter θ being equal to zero against the alternative of θ being positive. However, since it is not possible to test the null directly due to the fact that the speed of reversion γ is not identified, Kapetanios et al. (2003) propose a *t*-type test statistic following the work of Luukkonen et al. (1988), which is in fact a first-order Taylor series approximated to the ESTAR model.

$$\Delta y_t = \beta y_{t-1}^3 + \varepsilon_t \tag{2.30}$$

Hence, the t-statistic is obtained for $H_0: \beta = 0$ against $H_1: \beta < 0$ as follows:

$$t_{NL} = \hat{\beta}/s.e.(\hat{\beta}) \tag{2.31}$$

where $\hat{\beta}$ is the OLS estimate of β and $s.e.(\hat{\beta})$ is the standard error of $\hat{\beta}$. Asymptotic critical values of the t_{NL} statistic are different for different types of data, such as raw data, de-meaned data and de-trended data (Kapetanios et al., 2003).

Table 2.1. Critical values for ESTAR stationarity test.

Fractile (%)	**Raw data**	**De-meaned data**	**De-trended data**
1	-2.82	-3.48	-3.93
5	-2.22	-2.93	-3.40
10	-1.92	-2.66	-3.13

Sollis et al. (2002) introduced the idea of asymmetry in mean reversion adjustments in the time-series of real exchange rates and the effects of such asymmetry on unit root tests. Sollis (2009) extended the research by further development of the unit root test to allow asymmetry within ESTAR-type non-linear dynamics. The null hypothesis of the unit root is tested against the alternative of globally stationary ESTAR non-linearity which can be then further assessed in terms of exhibiting either symmetric or asymmetric behaviour. The test is based on the ESTAR unit root test by Kapetanios et al. (2003).

$$\Delta y_t = \beta y_{t-1}^3 + \delta y_{t-1}^4 + \varepsilon_t \tag{2.32}$$

The null hypothesis of the unit root is tested as coefficients β and δ, which are equal to zero ($H_0: \beta = \delta = 0$). The critical values for the test for the zero mean, non-zero mean and deterministic trend are in the table below, where T is the sample size.

Table 2.2. Critical values for asymmetric ESTAR stationarity test.

T	**Zero mean**			**Non-zero mean**			**Deterministic trend**		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
50	3.577	4.464	6.781	4.009	4.886	6.891	5.415	6.546	8.799
100	3.527	4.365	6.272	4.157	4.954	6.883	5.460	6.463	8.531
200	3.496	4.297	6.066	4.173	4.971	6.806	5.590	6.597	8.954
Asymptotic	1.837	2.505	4.241	3.725	4.557	6.236	5.372	6.292	8.344

Furthermore, based on the test by Kapetanios et al. (2003), Pascalau (2007) developed a framework for testing general STAR-type stationarity (2.31) and a unit root test which considers a logistic smooth transition (LSTAR) process non-linear stationarity in particular (2.32), where the null hypothesis of unit root is tested against the alternative of ESTAR and LSTAR stationarity for the general STAR test, and against LSTAR stationarity for the LSTAR unit root test.

$H_0: \gamma = \beta = \delta = 0$

$H_1: \gamma + \beta + \delta < 0$

$$\Delta y_t = \gamma y_{t-1}^2 + \beta y_{t-1}^3 + \delta y_{t-1}^4 + \varepsilon_t \tag{2.33}$$

$$\Delta y_t = \gamma y_{t-1}^2 + \delta y_{t-1}^4 + \varepsilon_t \tag{2.34}$$

The hypotheses are tested using the F-tests, and similar to the study by Kapetanios et al. (2003), Pascalau (2007) offers tabulated critical values associated with the test for models with raw data, de-meaned and de-trended data, where F_{NL} is the statistic for the general STAR unit root test and $\bar{F}_{NL}$ is the statistic for the LSTAR test.

Table 2.3. Critical value for the general STAR and LSTAR stationarity tests.

Fractile (%)	**Raw data**	**De-meaned data**	**De-trended data**
$\boldsymbol{F_{NL}}$			
1	4.92	5.16	6.08
5	3.64	3.87	4.72
10	3.05	3.30	6.08
$\boldsymbol{\bar{F}_{NL}}$			
1	4.92	5.16	6.08
5	3.64	3.87	4.72
10	3.05	3.30	6.08

Econometric forecasting

Point forecasts predict a single value of the variable under consideration, whereas interval forecasts attempt to predict a range of values in which the future value of the variable is expected to lie. Interval forecasts are usually given with a specified

confidence level. Harris and Sollis (2003) suggested that since the forecast and forecast error are random values, the interval forecasts would be more appropriate and useful. Furthermore, most financial trading schemes are based on a limit barrier and thus will benefit from knowing the possible range of values which can be given by an interval forecast. Granger et al. (1989) points out the importance of interval forecasting for various economic variables such as GNP growth, prices and unemployment rates. In addition, Value-at-Risk (VaR) and risk management are widely used practical applications of interval forecasting.

A related technique, density forecasts, on the other hand, provides an estimate of the probability distribution of possible future values of the forecasted variable (Wallis, 2003). Thus, allowing for full information about the forecasted density, such as dispersion or tails of the distribution (Mitchell and Hall, 2005).

Interval and density forecasts produce more informative predictions compared to point forecasts (Clements and Taylor, 2003). Both, interval and density forecasts supplement point forecast with a description of uncertainty. While interval forecasting specifies the probability of the forecasted outcome to fall within a specified interval of an upper and lower bound, density forecast offers a complete probability distribution of future outcome (Mitchell and Hall, 2005).

Christoffersen (1998) pointed out that while point forecasts are easy to compute and evaluate, interval forecasts and indeed density forecasts have an advantage over point forecasts in terms of their versatility in practical uses allowing for contingency planning as, by definition, interval forecasts indicate the range of likely outcomes. However, Diebold et al. (1998) point out computing difficulty, lack of demand for this type of

forecasts and lack of evaluation techniques as the reasons for limited research into density forecasts. While the former two points have changed in time with improved availability of computing technology and the fact that financial risk management has increased the requirement of density forecasts, density forecast evaluation methodology on the other hand is the topic for an improved methodology with many studies suggesting various methodology approaches (Diebold et al., 1998; Berkowitz, 2001).

The main body of literature focuses on evaluation of point forecasts awarding a relatively small proportion of research to interval (Chatfield, 1993; Christoffersen, 1998) and density (Diebold et al., 1998; Berkowitz, 2001) forecasts. The basis of interval and density forecasts evaluation involves comparing the forecasted coverage to the true coverage of the data (Baillie and Bolerslev, 1992; McNees and Fine, 1996). A number of researchers have attempted to formulate a single method of evaluation of interval and density forecasts similar to RMSE commonly used to evaluate and compare point forecasts. Christoffersen (1998) proposes a likelihood ratio as means of evaluation of interval forecasts as a model free forecast testing criterion similar to the works of Diebold and Mariano (1995).

Similarly, Mitchell and Hall (2003) proposed the Kullback-Leibler information criterion (KLIC) as a unified statistical tool for evaluation, comparing and combining density forecasts and which offers operational convenience in terms of practical use. The methodology is based on the likelihood ratio proposed by Berkowitz (2001) for evaluation of density forecasts. Mitchell and Hall (2003) suggest KLIC as a statistical method of evaluating density forecast in a similar fashion as root mean squared error (RMSE) is used to statistically evaluate and compare point forecasts.

In addition, Wallis (2005) carries out research into the combining of interval and density forecasts on the suggestion that different forecasts containing different information sets are possible to produce superior combined forecast, which is a similar approach taken for point forecasts in previous literature. However, the researchers suggested further development into the rules of combination and in particular optimal weight methodology of the density and interval forecasts for future research. Hall and Mitchell (2007) continue the research into combining density forecasts with an application of their methodology to UK inflation and find that their methodology of combining weights delivers encouraging results in terms of forecasting performance. Unlike the previous study by Wallis (2005) which used equally weighted forecast combination, Hall and Mitchell (2007) implement the Kullback-Leibler information criterion in order to determine the combination weights by minimizing the distance between the forecasted and the true unknown density. However, the best combined forecast failed to outperform the best individual forecast, thus suggesting density forecast combination as a topic for further research. The possible reason for research findings confirming individual density forecasts to outperform combined forecasts could be lack of research into techniques and rules of combination and optimal weights allocation. The emergent research into the topic of combined density forecasts yet lacks a firm explanation of whether the poor performance of combined density forecasts is due to incorrect combining procedures or due to theoretical underlining.

An interval forecast consists of upper and lower limits, or prediction intervals, between which a future expected value of the forecasted series is expected to lie with certain assigned probability. In relation to computation of the prediction intervals, Chatfield

(1993) distinguishes between conditional and true forecast errors, as well as the importance of forecast error variance in terms of quality of the interval forecast.

An observed time-series $(x_1, x_2, \dots, x_n)$, where n is the number of observations, follows a stochastic process X_t, at time t. The k-step-ahead point forecast conditional on data up to time n is denoted as $\hat{X}_n(k)$ when regarded as a random variable and $\hat{x}_n(k)$ when it is a particular value. Thus the forecast error conditional on data up to time n is the difference between the actual value of the random variable and the point forecast value, which can be expressed as follows:

$$e_n(k) = X_{n+k} - \hat{x}_n(k) \tag{2.35}$$

Since the observed value of $e_n(k)$ becomes available at time $n + k$, the out-of-sample conditional forecast errors are the true forecast errors, while in-sample forecast errors are the residuals from the fitted model. Calculation of interval forecasts involves computing of the expected mean squared prediction error, or PMSE ($E[e_n(k)^2]$), in order to set the prediction intervals. Unbiased forecast where the point forecast is the mean of predictive distribution would thus be characterised with a zero prediction error ($E[e_n(k)] = 0$) and variance ($E[e_n(k)^2] = var[e_n(k)]$). Chatfield (1993) point out that the evaluation in terms of forecast uncertainty relies on the evaluation of the variance of the forecast errors rather than the forecast.

Granger et al. (1989) propose a practical approach of obtaining interval forecast for estimated time-series models, which also allows for possible presence of non-linearity in the series. Further, Chatfield (1993) suggests a general procedure for calculating

prediction intervals whereby a $100(1-\alpha)\%$ prediction interval for X_{n+k} is given in the following form:

$$\hat{x}_n(k) \pm z_{\alpha/2}\sqrt{var[e_n(k)]} \tag{2.36}$$

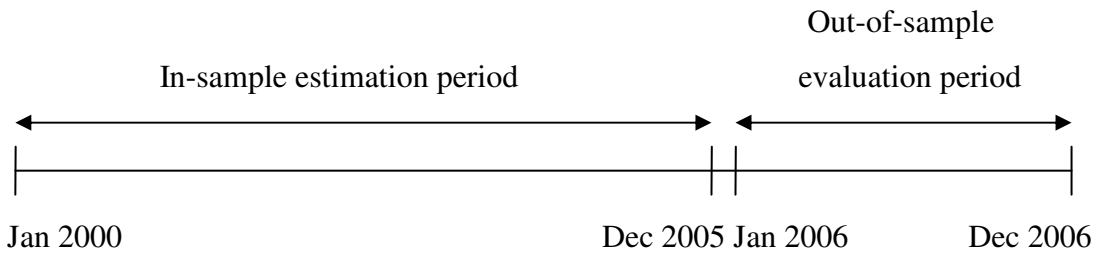
Where $z_{\alpha/2}$ indicates the appropriate percentage point of a standard normal distribution. Eq. 0.02 holds assuming that the forecast is unbiased and that the forecast errors follow normal distribution. However, in practice, $z_{\alpha/2}$ sometimes is assumed to follow a *t*-distribution.

While both interval and density forecasts compared to point forecasts provide likelihood of accuracy and more thorough understanding and comparison opportunities of forecasts, these types of time-series forecasts have been characterised with a number of drawbacks. Thus, for interval and for density forecasts, in particular, problems occur when error distribution is not normal. Estimation techniques for both types of forecasts assume normally distributed error term. This assumption does not consider the common presence of outliers in the time-series data, which result in associated asymmetry and heavy tails of the distribution. Moreover, as with any econometric forecasts, the results depend heavily on an identification of a fitted model and dangers of estimating the wrong model. This remains true for interval and density forecasts, nonetheless, most approaches of computing prediction intervals are based on the assumption that the correct model was fitted. Furthermore, changing structure of the underlying model due to either slow changes in the dynamics of the data or sudden shocks, bears a significant impact on the estimation and hence the performance of interval forecast. In addition, Tay and Wallis (2000) point out the importance of correct presentation of density

forecasts as inappropriate presentation might significantly reduce the practical usefulness of the forecast as a whole by leading to misinterpreted or misleading results. The topic of interval and density forecasts is still a rather sparse subject in academic literature and research, thus resulting in a lack of generally accepted methods of calculating and evaluating the forecasting results.

Furthermore, the notion of a one-step ahead forecast as well as a multi-step ahead forecast should be mentioned. It is evident from the terminology, that the former forecast is generated for the next observation only, while multi-step-ahead forecast is generated for a particular number of forecasts for the future time period. The number of steps of the forecast depends on the type and purpose of the forecast. In addition, when deciding on the forecasting horizon one should appreciate that different models might be superior in producing forecasting for short horizons up, to one or two steps ahead, while producing poor forecasts further ahead, and vice versa.

Moreover, estimation and forecasting periods are different in in-sample and out-of-sample forecasts. According to Brooks (2002), in-sample forecasting involves estimating fitted values using the same data that was used to estimate the model. Naturally, models are expected to produce relatively accurate in-sample forecasts. Hence, for model evaluation purposes and examination of forecast accuracy, the latter sample of observations is withheld from the estimation sample for the purpose of using this so called holdout sample to generate an out-of-sample forecast. Thus, the holdout sample then can be used to assess the accuracy of the forecast by comparing fitted with actual values.



For the purpose of illustration, assume the current data of interest ranges from 1 January 2000 to 31 December 2006. With an intention of carrying out an out-of-sample forecast a researcher may estimate an appropriate time-series model using in-sample estimation from the period of 1 January 2000 to 31 December 2005, withholding the sample from 1 January 2006 to 31 December 2006. The out-of-sample forecast is then estimated for the period from 1 January 2006 to 31 December 2006 and consequently compared with the actual values available as a holdout sample. Note that the number of observations in the out-of-sample forecast and the holdout period is the same since these are using the same data period.

In addition, forecasts can be performed using either recursive or rolling window forecasting techniques. When applying a recursive method, the initial estimation date is fixed and additional observations are added one by one to the whole of the estimation period. While a rolling window technique implies the length of the in-sample period to be fixed, thus the start and the end dates increase with addition of each new observation.

While in-sample forecasts provide a good evaluation of a model in terms of goodness of fit, out-of-sample forecasts provide more accurate assumptions regarding the forecasting accuracy of econometric models. Similarly, recursive forecasts tend to utilise the dynamic patterns of data through constant re-estimation. Furthermore, one-step forecasts are preferred due to the simplicity of estimation and evaluation techniques, as

opposed to a limited evidence of adequate performance of the multi-step forecast. Consequently, empirical chapters in this paper concerned with forecasting methodologies will consider one-step ahead point recursive out-of-sample forecasts.

Tests of forecasting accuracy

As Ericsson (1992) pointed out, the success of any empirical economic model is assessed on the basis of how well it is able to explain significant features of the data thus capturing its true dynamics. In addition, models are tested on their abilities to deliver reliable predictions of the future behaviour of the data, or in other words, forecasting accuracy of the model. Moreover, tests of forecasting accuracy are also used when comparing competing models in order to determine which model generates the superior forecast.

Granger and Newbold (1977) proposed a notion of a cost function as a criterion of optimising of a point forecast. It is based on the assumption that forecast errors have a high probability of occurrence in connection with a random process. Granger and Newbold (1977) suggested the notion of the cost of an error $C(e)$, where the error e is defined as $e_{n,h} = y_{n+h} - f_{n,h}$, where $f_{n,h}$ is the forecast for y_{n+h} based on the information set I_n. Thus the cost of a zero error will equal to zero, $(0) = 0$, while forecast based on not optimal decisions will result in cost of $C(e_{n,h})$. To reduce a cost function the point forecast $f_{n,h}$ is chosen so that the expected cost of forecasting errors

$E_c\{C(e_{n,h})\}$ is minimised. Thus the optimal forecast is the forecast function that minimises the error cost function[4]:

$$f_{n,h} = E_c\{y_{n+h}\} \tag{2.37}$$

There are numerous tests that can be performed to assess accuracy of a time-series forecast. In the case of out-of-sample forecasts the actual values from the holdout sample are compared to the forecasted values and the differences between those values are analysed using various tests of forecasting accuracy appropriate to the specific type of forecast. Brooks (2002) defines a forecast error as the difference between the value of an observation and the value of the forecast made for this observation. Hence, the forecast error can be either positive, when the forecasted value was too low, or negative, when the forecast was too high. Due to this fact, the forecast errors are usually squared or the absolute value is taken to prevent mathematical cancelling out when summed to provide a forecast error value for the whole series. Techniques that assess forecasting errors in such way are usually referred to as statistical loss function tests.

The most commonly known statistical loss function tests include mean squared error (MSE) and mean absolute error (MAE). These tests are used when comparing forecasts from different models performed on the same data and over the same forecasting period. As in most similar loss function tests the model producing the lowest value of MSE or MAE is considered to be more accurate. However, as Harris and Sollis (2003) pointed out, a lower MSE of one forecasting model in comparison to another does not necessarily indicate superiority of the first model for the simple reason that the

[4] For further detailed discussion of theory of forecast optimisation and loss functions refer to Granger and Newbold (1977).

difference between their MSEs may not be significant enough to support that claim. They recommend that the test of equal forecast accuracy developed by Diebold and Mariano (1995) is used to assess whether the difference between MSEs of competing forecasts is statistically significant from zero. Monte Carlo simulations showed the test to be a valuable tool, however, it was found to be over-sized for small forecasts of two or more steps ahead. Consequently, Harvey, Leybourne and Newbold (1997) modified the original Diebold-Mariano test in order to improve its performance. As a result, the new statistic exhibits a much higher performance and is robust for different forecast horizons, as well as autocorrelated and non-normally distributed errors.

Furthermore, Ericsson and Marquez (1993) point out that when presented with few competing forecasting models it should be taken into account that different models may perform well individually in capturing different features of the data's behaviour. Chong and Hendry (1986) proposed the concept of forecast encompassing which relates to the model's informational content. The test allows investigation of whether the forecasts of one model can explain the forecast errors of another, or whether competing models contain no additional information, thus assuring the superiority of the original model.

On the contrary, some researchers argue that regardless of whether the forecasting accuracy tests indicate the superiority of one model over the other, the main indication of a successful forecast is whether it can convey any practical gains in terms of a generated profit when using the forecasting model. Trading rule style tests are very popular and are a relatively easy way to compare performances of different forecasts. These are known as economic loss functions.

While the traditional statistical measures have their drawbacks and are mostly not equipped to deal with non-linear time-series, these are very easy to implement and interpret, and as some researchers would argue, provide a clear overview and enough information to draw conclusive assumptions. Hence, this study will by no means dismiss these techniques, and nevertheless will consider other methods of comparative measurement. This section describes traditional statistical tests of forecasting accuracy as well as some alternative procedures highlighting benefits and drawbacks on each test before applying these to the forecasted series considered in this study.

Statistical loss function tests

The following test procedures are the most commonly used statistical loss function tests that can be applied to evaluating forecasting accuracy of time-series models. These tests are also often used to evaluate econometric models at the estimation stage of model building, where the selection criteria is based on minimising the value of these statistics. Due to the simplicity and relative ease of interpretation a number of these tests are run by researchers as standard practice when attempting a forecasting exercise and are included in most software modelling packages, thus providing readily availability of the tests.

All the tests of forecasting accuracy considered here are applicable to out-of-sample forecasts and will be performed by creating what is known as a dynamic simulation (Pindyck and Rubinfeld, 1998) where forecasted values are compared to the actual values withheld in the holdout sample period. This approach allows researchers to

determine how close the predictive values mimic the corresponding actual data series. Different tests interpret the comparison between forecasted and actual values in different ways although are based around the same principle.

Error magnitude tests

Mean error (ME) is one of the simplest statistical loss function tests measuring the forecasting performance in terms of the magnitude of the forecasting errors, and involves taking the mean value of the sum of differences between actual and forecasted values.

$$ME = \frac{1}{T-1}\sum\left(y_{t+s} - f_{t,s}\right) \tag{2.38}$$

where y_t is the actual value of the variable at time *t*, and *T* is the sample size including the out-of-sample observations.

Thus, mean squared error (MSE) is the sum of residuals, or forecast errors, divided by the number of degrees of freedom, which in essence provides a measurement of residual variance. MSE of forecast error for *s*-step ahead forecast at time *t*, $f_{t,s}$:

$$MSE = \frac{1}{T-1}\sum\left(y_{t+s} - f_{t,s}\right) \tag{2.39}$$

Similarly, mean absolute error (MAE) measures the average absolute of forecast errors for the forecast $f_{t,s}$:

$$MAE = \frac{1}{T-1}\sum_{t=T_i}^{T}\left|y_{t+s} - f_{t,s}\right| \tag{2.40}$$

Mean absolute percentage error (MAPE):

$$MAPE = \frac{100}{T-1}\sum_{t=T_i}^{T}\left|\frac{y_{t+s} - f_{t,s}}{y_{t+s}}\right| \tag{2.41}$$

Adjusted MAPE (AMAPE) is also known as a symmetric MAPE (Brooks, 2002), which corrects for the asymmetry between the actual and forecasted values by dividing the forecast error by the average of actual and forecasted values twice.

$$AMAPE = \frac{100}{T-1}\sum_{t=T_i}^{T}\left|\frac{y_{t+s} - f_{t,s}}{y_{t+s} + f_{t,s}}\right| \tag{2.42}$$

Statistics like ME, MSE and MAE can be used for comparisons between different models as long as these are estimated using the same data and forecasting period. Generally, the model with the lowest value of ME, MSE or MAE statistic is regarded as

the more accurate one. The MSE statistic is also tolerant towards models where there are significantly more much larger forecast errors than smaller errors. However, MSE is scale dependent, which means it requires forecasts to be made using the same data and forecasting period in order to carry out a valid comparison. Both MAPE and AMAPE, on the other hand, can be used to compare a wider range of forecasts as these statistics are interpreted as a percentage. However, MAPE and AMAPE statistics cannot be used when the forecast values and the series can take opposite signs, as in the case of forecast returns, for instance. This is due to a chance that the values can cancel each other out, which in turn will result in extremely large and erratic values of these statistics (Brooks, 2002). Moreover, if absolute values of the series are less than unity, the MAPE statistic becomes unreliable (De Gooijer and Hyndman, 2006).

Another useful comparison measure commonly used by forecasters is root mean square forecast error (RMSE). Pindyck and Rubinfeld (1998) defined RMSE as a measure of deviation of the forecast from the actual variable over time. As ME, MSE and MAE statistics, RMSE is only a comparison measure and can only be used when assessing similar constructed data sets.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(f_{t,s} - y_{t+s}\right)^2} \tag{2.43}$$

To overcome the comparison constraint and measure RMSE in relative terms, there is a similar technique statistic known as Theil's U inequality coefficient, where the

numerator is the RMSE and the denominator scales the whole statistic so that the values of U will fall between 0 and 1.

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(f_{t,s} - y_{t+s}\right)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(f_{t,s}\right)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_{t+s})^2}} \tag{2.44}$$

Evidently, when the statistic's value equals zero it signifies that the forecasted and actual values are equal, hence indicates the best accuracy of the forecast. Value of unity, on the other hand, signifies that the forecasting model is as inaccurate as it can be. However, even though the Theil's inequality coefficient is a very useful statistic, similarly to MSE, it is too influenced by outliers and extreme data points (Brooks, 2002).

Empirical chapters of this dissertation apply different forecasting methodologies to time-series financial data with an intention to determine the preferred superior forecasting model for each data set. Due to the nature of the forecasting exercise employed here, the statistical loss functions will be compared to the same series of data across the forecasting exercise, thus this paper will apply ME, MAE and RMSE statistics for comparative measure.

Diebold-Mariano test of equal forecast accuracy

Diebold and Mariano (1995) introduced a test of equal forecasting accuracy which tests whether the differences in MSEs of competing forecast models are statistically significant. The test is based on the idea that if one of the competing models displays a lower MSE value than the other model it does not necessarily mean that it produces a superior forecast, as the difference between values of MSEs might not be statistically significant. The test is intended for comparing results of competing forecasts of the same quantitative value. The Diebold-Mariano test takes into account two sets of forecasting errors from two forecasting models, e_{1t} and e_{1t}, and runs the hypothesis represented by the expectations operator, E, such that $E[d_t] = 0$, where d_t is the difference between the squared forecast errors, $d_t = e_{1t}^2 - e_{2t}^2$. The mean of this difference can be expressed as $\bar{d} = n^{-1}\sum_{t=1}^{n} d_t$ with the variance of $V(\bar{d}) \approx n^{-1}\left[\gamma_0 + 2\sum_{k=1}^{k-1}\gamma_k\right]$, where γ_k is the k-th autocovariance of d_t, which is estimated as following:

$$\hat{\gamma}_k = n^{-1}\sum_{t=k+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d}) \tag{2.45}$$

The statistic for the Diebold and Mariano (1995) test is testing the null hypothesis of equal forecast accuracy:

$$S_1 = \left[\hat{V}(\bar{d})\right]^{-\frac{1}{2}}\bar{d} \tag{2.46}$$

Diebold and Mariano (1995) have reported good test results using Monte Carlo simulations. The test performed well for small samples and when forecast errors displayed autocorrelation and non-normal distributions. However, the test was found to be over-sized for two or more steps ahead forecasts.

Harvey, Leybourne and Newbold (1997) reviewed the original test proposed by Diebold and Mariano (1995) amongst a few other similar tests, in order to assess and possibly improve performance of the test. Similarly to Diebold and Mariano (1995), Harvey et al. (1997) questioned whether one forecast being more successful than the other by a small amount was significant enough to make a claim of forecast superiority or if it was due to chance. Harvey et al. (1997) modified the original Diebold-Mariano test in order to address the test being over-sized for two or more steps ahead forecasts. By modifying the test statistic and considering the Student's *t* critical values instead of the standard normal distribution, Harvey et al. (1997) have significantly improved the original test making the modified Diebold-Mariano test the best available procedure for comparing forecasts in terms of equal forecasting accuracy. The test demonstrated a very powerful performance and proved to be simple to compute.

The modified Diebold-Mariano statistic by Harvey et al. (1997) is as follows:

$$S_1^* = \left[\frac{t + 1 - 2h + t^{-1}h(h-1)}{t}\right]^{\frac{1}{2}} S_1 \tag{2.47}$$

where S_1 is the original Diebold-Mariano test statistic for h-steps ahead forecast for time t. Critical values for the modified test are taken from the Student's t-distribution with $(t-1)$ degrees of freedom.

Forecast encompassing test

The aim of the forecast encompassing test is to assess whether a forecast from a competing model contains any information that is absent from the original model. If it does not, the forecast from the competing model is said to be encompassed by the forecast from the original model (Harris and Sollis, 2003). Hence, it will be unnecessary to combine these two models in anticipation that it would produce a forecast of a superior quality.

Fang (2003) carried out an extensive research on whether competing individual forecasts can be successfully combined into one which in turn would be much superior. The assessment of forecast superiority was performed using the forecast encompassing tests. As a result, Fang (2003) found that in that particular case each individual forecast contained independent information necessary for forecasting the dependent variable, in other words, neither forecast encompassed the other. However, Fang (2003) did establish that the forecast encompassing test is a complimentary and a necessary tool to such criteria as RMSE and MAE.

The test is carried out by regressing actual levels or change in the dependent variable Y_t on the forecasted values $Y_t s$ (Fang, 2003). The same principle can be applied when

comparing values from two different forecasts. In this case the predicted values of a benchmark forecast are regressed on the predicted values of the alternative forecast.

$$Y_t = \alpha_1 \hat{Y}_{t-s,t}^{(I)} + \alpha_2 \hat{Y}_{t-s,t}^{(II)} + u_t \tag{2.48}$$

This simple version of the forecast encompassing tests H_0 of $\alpha_2 = 0$ against H_1 of $\alpha_1 = 0$. Hence, the first model forecast encompasses the second model when $\alpha_1 \neq 0$, $\alpha_2 = 0$; in the case of the first model forecast being encompassed in the second model forecast $\alpha_2 \neq 0$, $\alpha_1 = 0$. Any other outcome will indicate that neither model encompasses the other. Moreover, if both forecasts contain independent information for forecasting the Y_t, than both α_1 and α_2 should be non-zero ($\alpha_1 \neq 0, \alpha_2 \neq 0$); whereas if neither model contains any information required for forecasting the Y_t both α_1 and α_2 should be zero.

In their original research Chong and Hendry (1986) have used the above regression with a restriction of $\alpha_1 + \alpha_2 = 1$. However, Fair and Shiller (1990) adopted a slightly different approach.

$$Y_t - Y_{t-s} = \alpha + \beta_1\left(\hat{Y}_{t-s}^{(I)} - Y_{t-s}\right) + \beta_2\left(\hat{Y}_{t-s}^{(II)} - Y_{t-s}\right) + u_t \tag{2.49}$$

$H_0: \beta_1 = 0$, $H_1: \beta_2 = 0$

where $\hat{Y}_{t-s}^{(I)}$ is the forecast of Y_t made from the forecasting model *(I)*, and $\hat{Y}_{t-s}^{(II)}$ is the forecast of Y_t from the forecasting model *(II)*. Fair and Shiller (1990) are testing similar hypothesis that the forecasts made by model *(I)* contain no relevant information for forecasting the $Y_t(H_0)$ against the hypothesis that model *(II)* contains no relevant information (H_1).

Fair and Shiller (1990) do not put a constraint on β and γ to sum to unity on the argument that if the forecasts from both models are just noise, they expect both estimates to be zero. Furthermore, in the case of Y_t being a result of two independently distributed processes each of the competing forecast models could specify each of those processes individually, thus having both coefficient estimates equal to unity, which would sum up to two. Similarly, Fair and Shiller (1990) do not restrict the constant term α to be equal to zero, since in the case of both models being a noise and the estimates of β_1 and β_2 equal zero, the constant is required to account for the non-zero mean of the dependent variable. It is also suggested that the u_t is likely to be heteroscedastic and can be treated as a general forecast error term.

Similarly, Ericsson and Marquez (1993) also pointed out that the original forecast encompassing test by Chong and Hendry (1986) was designed for static linear models and assumed *i.i.d.* forecast errors. Ericsson and Merquez (1993) thus generalised the test to accommodate these points as well as to include a constant term. This allowed for the test to be performed for multi-step ahead forecasts on several competing models at the same time and allowed the uncertainty from estimating model coefficients, in the cases when these are unknown. As a result, their test consists of GLS estimation of the following equations using $\left[\Phi^{(h)}\right]^{-\frac{1}{2}}$, where $\Phi^{(h)}$ is an approximately diagonal matrix.

$$v_{T+s}^{(h)} = k_l \mu_{T+s}^{(l)} + e_{T+s}^{(h)} \tag{2.50}$$

$$v_{T+s}^{(h)} = k_0 + \sum_{l \neq h} k_l \mu_{T+s}^{l} + e_{T+s}^{(h)} \qquad s = 1, \dots, s \tag{2.51}$$

Where $v_{T+s}^{(h)}\left(\equiv A \cdot y_{T+s} - \mu_{T+s}^{(h)}\right)$ is the actual forecast error of model h, $\mu_{T+s}^{(h)}$ is the forecast of model l ($l \neq h$), $e_{T+s}^{(h)}$ is the error term of the regression. Equation 2.51 contains a non-zero constant term k_0. The procedure is to test $k_0 = 0$ and $k_l = 0$ as a joint hypothesis; and $k_0 = 0$ given $k_l = 0$. GLS estimation is used to account for any autocorrelation in the forecast errors that will most likely be present in non-linear as well as linear models due to coefficient uncertainty. Ericsson and Marquez (1993) also cautioned that non-linearity of a model might produce non-normality in the forecast errors, though according to the researchers, it should not affect the forecast encompassing test statistic.

On the contrary, Harvey and Newbold (2000) consider the forecast encompassing test to lack robustness due to forecast error non-normality and recommend their modification of Diebold-Mariano-type test for forecast encompassing used on multiple models. They found the modified test to be a preferred option, especially in large samples, however, to have limitations when applied to small samples. In their argument, Harvey and Newbold (2000) suggest that the test's drawback is not significant when considering its reliability for the large size samples. Harvey and Newbold (2000) considered the forecast encompassing process in terms of a weight average linear forecast combination where in the case of the inferior forecast to be encompassed in the other model, the optimal weight of the inferior forecast is zero. In addition, the forecast encompassing

test besides its direct use in forecast comparisons may also be used as an indicator of misspecification of a model and hence suggest further improvement of that particular model (Ericsson and Marquez, 1993).

This study will apply the standard Diebold-Mariano (Diebold and Mariano, 1995) test, as well as the modified Diebold-Mariano (Harvey and Newbold, 2000) test to the forecasts generated using linear and non-linear models, as well as forecasting errors of those models (Fair and Shiller, 1990).

Combined forecast test

Winkler and Clemen (1992) suggest that the basis for combining forecasts is very intuitive as such approach intends to reduce the risk of a particularly poor forecast. However, this risk might be counteracted if the approach results in weights allocated to each individual forecast that are too sensitive or extreme. Hence, they advise on methods that reduce weights variabilities, including a simple average and outperformance measurement (Gupta and Wilton, 1987). However, the simplest and thus most commonly used method of selecting the combining weights is the simple arithmetic average. This method has proven to be robust and relatively accurate. It is usually used as a benchmark and was often found to perform better than alternative more complex methods (Clemen and Winkler, 1986).

Assuming that $f_1, f_2, \dots, f_k$ are the forecasts for the variable in question, y, the combined forecast $f_c(y)$ using the equal weighting method can be expressed as follows:

$$f_c(y) = \frac{1}{k}\sum_{i=1}^{k} f_i \tag{2.52}$$

This approach does not take into account differences between combined forecasts in terms of their accuracy, as the logical choice would be to allocate a stronger weight to a more accurate forecast. In other words, this approach implies that the forecasts are exchangeable (Clemen and Winkler, 1986; Gupta and Wilton, 1987). Hence, it is evident that regardless of its practical usefulness this method lacks theoretical justification, as it does not utilise information contained in the past data patterns, and, as pointed out by De Gooijer and Hyndman (2006), does not take into account the dependence among the forecasts' errors. The simple equal weighting approach was further criticised by Gupta and Wilton (1987) for not accommodating for any additional information available to a researcher or a decision maker, including correlation between forecast errors and different functional structures of each model. Despite the downfalls of this approach, it performs well in empirical studies. Gupta and Wilton (1987) explain this due to the models used in the combination forecast having similar variances. Approaches that, on the other hand, have a solid theoretical base, such as minimum variance approach (see Gupta and Wilton, 1987), perform poorly in empirical studies, not robust enough and generally too sensitive to data non-stationarity and are outperformed by the equal weighting method.

Gupta and Wilton (1987) also dismissed the possibility of a judgemental method on allocating weights on the basis that such approach will be too complex to implement and thus might not use all the available information efficiently. Instead Gupta and Wilton (1987) proposed the Odds-Matrix method for weight allocation of combined

forecast and found it to address all the necessary properties that are desired for this purpose. They found the method is flexible enough to adopt weights should new information about the data become available. Gupta and Wilton (1987) attempted to provide a procedure for allocating weights for combined forecasts such that the weights will be intuitively meaningful and not dependent upon large amounts of data. They have tested the proposed approach against previous procedures and found that the new methods perform equally well when used on large data and significantly better when used on sparse data. More so, as one of the methods to deal with the problem of non-stationarity when combining forecasts Clemen and Winkler (1986) suggested allocation of heavier weights to most recent observations.

In addition, Fang (2003) warned against simple combining of multiple linear forecasts in order to achieve lower values of RMSE than that of an individual forecast. Combined forecasts will inevitably have lower RMSE values due to greater sample variability from the combined forecasts. Fang (2003) doubted whether smaller RMSE does indeed signify superiority of a forecast. Due to these factors forecast combination might appear challenging and difficult to interpret. Moreover, combined forecast weights can also be determined by OLS, however, since there is a possibility of serial correlation in the combined forecast errors, the weights are inefficient. The forecast encompassing test, on the other hand, can be a valuable tool in model specification and forecasting accuracy assessment. Nonetheless, despite the criticism and clear downfalls of the simple average, equal approach seems to perform consistently well in empirical studies and investigations. Hence, one might question whether it is worth an effort to determine a more theoretically robust approach which will still offer the same results and the same ease of application. Therefore, this paper will employ forecast combination in

conjunction with the results of forecast encompassing tests using a simple equal weighing method.

Economic loss function tests

Some researchers argue that since the main objective of time-series models is practical forecasting it makes logical sense to assess these models on the basis of their potential profitability. However, Leitch and Tanner (1991) draw the attention to the fact that most economic forecasts completely overlook their profitability. Some researchers point out that forecasting models that might perform poorly as indicated by the statistical base criteria may yet prove to be very useful in yielding a profit when used for trading. Hence, real life practitioners will value models that accurately predict the sign of returns or turning points in a series, rather the ones that have the lowest statistics. Furthermore, there are tests that assess the ability of a forecasting model to predict the direction of changes of future values, and correct magnitude or percentage change in values of those predictions.

Leitch and Tanner (1991) support a similar argument that the conventional statistical methods of forecast evaluation have little to do with the forecasts' profitability. In their study, the researchers compared the standard statistics, such as average absolute error (AAE), the root mean squared error (RMSE) and the Theil's *U* statistic, which all assess the magnitude of the forecast error, and found that none of these criteria relate to profitability of the forecast in question. Hence, considering their argument, Leitch and Tanner (1991) question whether standard conventional error measuring criteria justifies

payment to professional forecasting bodies allocated by companies, which presumably are profit maximising. Further investigation by Leitch and Tanner (1991) of the quality of interest rate forecasts and relation of criteria of forecasting accuracy to profitability have established that while error measuring criteria is not related to the profitability of the forecast, relation between profitability and directional accuracy forecasts seem to be more reasonable. The results suggest that in practice the preference will be given to the directional accuracy forecast as it demonstrates strong statistical association while error measuring tests relate to profitability merely marginally.

A number of studies were undertaken to investigate the results behind research into the presence of predictability, based on the paper by Brock et al. (1992). However, while most of these are enquiring whether there is any predictability in the data patterns, Ready (2002) is apprehensive as to whether any predictability present in the data is sufficient enough to generate profit after transaction costs. Furthering the argument, Ready (2002) points out that it is not always essential to consider the profitability of trading rule net transaction costs, as an investigation into profitability generating abilities of any financial modelling will be beneficial to practitioners in terms of deeper understanding of market dynamics.

The aim of this exercise is to test whether it is possible to create a technical analysis on the basis of modelling patterns uncovered in past data and to exploit these in order to generate profitability in terms of excess returns, and whether certain types of models have beneficial advantage in doing so. However, since the purpose of this exercise is to use the trading rule approach as merely a test of accuracy of forecasting models considered earlier rather than a realistic trading strategy, there is no need to address the issue of transaction costs here. This study will use method of profit calculation for the

trading rule approach loosely based on procedure suggested by Leitch and Tanner (1991) where the profit is calculated on the basis of whether the forecasting error, or in essence, return, is positive or negative. Hence, it will be assumed that if the forecast error is positive, i.e. the forecasted value is above the actual value, the long position (buy) will be taken on the contract; and the short position (sell) if the forecast error is negative, in other words, the forecasted value is below the actual value.

The trading rule methodology considered in empirical chapters of this paper should not be mistaken with the practical approach of creating a successful trading procedure aimed at generating profit thus implying its use by traders. This investigation implements the trading rule technique to assess the accuracy of the forecasts drawn earlier in this chapter as an additional variation of forecasting accuracy tests. Hence, the so called profit calculated here using the trading rule method will be an indicator of comparative success of each forecasting model, and by no means is an implication of profitability of such forecast, as this is a completely different concept to what is considered in this study. While some researchers are concerned with investigating a presence of predictability in daily returns patterns, intention of other researchers, for instance Ready (2002), is to realise predictability sufficiently strong to generate profit substantial enough to account for transaction costs. The latter approach may seem very practical and clearly a logical choice, however, on closer inspection such tactic requires much detailed and complex consideration of particular needs and requirements of different types of practitioners it is aimed at. For instance, certain types of investors might only be interested in seasonal directional changes of the market over a long-term period, rather than short term profitability based on the magnitude of market changes. Similarly, daily market activities might present little interest to policy makers concerned

with general market behaviour whether these are profitable or not. In addition, to find suitable levels of profitability in a particular data set one will have to consider transaction costs very specific to certain contracts presented in that data, which in turn implies an individual approach to a forecasting exercise, rather than a general investigation intended in this study.

2.4. Conclusion

Time-series forecasting models may grant a deeper understanding of a series as it examines behaviour patterns which in turn may spark a new found interest for certain extraneous factors that might offer an explanation of the series dynamics (Newbold and Granger, 1974). It is evident that econometric modelling and forecasting is important across a wide range of disciplines (Holden et al., 1990; Diebold and Mariano, 1995; Granato and Suzuki, 1996; Montgomery et al., 1998). A wide range of models provides extended flexibility and a variety of approaches for modelling different characteristics of the data, however, at the same time such a broad array creates additional challenges in terms of correct model specification, danger of overfitting the data and basing the model on spurious assumptions. Thus, Chatfield (1997) recommends comparing of out-of-sample forecasting performances of fitted models as opposed to only an in-sample comparison. Nonetheless, the univariate time-series models seem to be the predominant choice for forecasting due to ease of computation and interpretation of the results.

Furthermore, issues of econometric forecasting also include a component of real life practitioners and market participants, and thus challenges associated with practical modelling, forecasting and forecast accuracy assessment. Chatfield (1997) brought attention to the fact that forecasts are used in different ways by practitioners, thus sales forecasts are used as a target setting technique and that judgemental forecasts are still used extensively despite the lack of theoretical support of accurate forecasting performance. Chatfield (1997) also pointed out that while companies rate accuracy as the most important rating criterion of a forecast, there is no clear definition of how exactly it is measured in practice. Moreover, while the extensive range of forecasting computer software allows flexibility of the modelling and forecasting process, satisfying an array of users and providing easy-to-use packages, it also presents an increased possibility of estimating an incorrect model and misinterpretation of the result due to misuse of the forecasting package (Chatfield, 1997). Chatfield (1997) points out that the availability of computational advantages could also result in overfitting, whereby an econometric model could be fitted to data to produce relatively satisfying forecasting results, however, there is a danger of fitting certain models when they are not appropriate by ignoring the theoretical and logical reasoning of why a particular model should be applied to particular data. Chatfield (1997) also reminds that any econometric forecast is based on an assumption and comparing of competing forecasts should be performed on the results of out-of-sample forecasts as opposed to in-sample estimations. Summarising the review of forecasting methods in the 1990s Chatfield (1997) concludes that forecasting is very much the same as in previous decades in terms of difficulties and challenges faced by forecasters, only characterised by a wider range of models and extensive availability of software.

Chapter 3
Daily stock returns forecasting

3.1. Introduction

This chapter intends to apply time-series non-linear framework to daily returns of four leading stock indices, namely FTSE 100, S&P, DAX and Nikkei, with the purpose of recursive out-of-sample forecasting. In view of the fact that the main purpose of econometric modelling appears to be application of these models to forecasting, this paper is concentrating on fulfilling this objective and extending research into non-linear model forecasting.

Given that a number of studies highlight the importance of forecasts in general (Brooks, 2002), in planning and operations of companies (Holden et al., 1990), political science (Granato and Suzuki, 1996) and for economic policy-makers (Montgomery et al., 1998), there is no doubt that forecasts are required in a wide range of disciplines. The main focus of this chapter is daily stock return forecasts which are also required by a broad spectrum of market practitioners. Moreover, the degree of sensitivity of non-linearity to the frequency of data is still not entirely clear. For instance, according to Abhyankar et al. (1995) who suggest the use of high-frequency data, microstructural dynamics in the financial time-series are more apparent at higher frequencies. In addition, high-frequency data provides a large sample size for empirical investigation.

Daily stock returns predictability has been much debated over the years. The reason for this is inconsistency of returns stock predictability with efficient market hypothesis (EMH), which states that the stock price incorporates all publicly available information. Cuthbertson and Nitzsche (2004) describe efficient markets being driven by simple supply and demand mechanism of a competitive market where rational traders react and consequently adjust the stock prices according to the available information relevant to the determination of fundamental asset prices. According to the theory, due to any relevant information being costless and publicly available while new information, such as news, being unpredictable by definition, there is no opportunity to accumulate excess profit in a perfectly efficient market. Thus, Abhyankar et al. (1995; 1997) point out that returns stock predictability is inconsistent with the theory of efficient markets, however, find evidence of predictability and non-linear dependence in high-frequency FTSE returns and daily returns of S&P 500, DAX, Nikkei 225 and FTSE 100. Attempts to explain the stock market predictability suggested market inefficiency or time-varying expected returns (Brock et al., 1992; Pesaran and Timmermann, 1995). Furthermore, the presence of cyclical behaviour and asymmetric adjustments in economic and financial series implied the presence of non-linear predictability (Tong 1990; De Gooijer et al., 1992; Abhyankar et al., 1997; McMillan, 2001; Sarantis, 2001; McMillan, 2002; Bali et al., 2008; Hartmann et al., 2008; Guidolin et al., 2008). The presence of these non-linearities could be attributed to the presence of market frictions, including transaction costs, borrowing and short selling constraints, limit to arbitrage (He and Modest, 1995; Kilian and Taylor, 2003; McMillan, 2005), as well as the presence of speculative bubbles (Evans, 1991; Froot and Obstfeld, 1991; Bohl, 2003; Psaradakis et al., 2004) and interaction between noise traders and informed arbitrageurs

(Kirman, 1991, 1993; Shleifer, 2000; McMillan, 2002, 2005; McMillan and Speight, 2006).

Section 3.2 of this chapter contains a brief reminder of the methodology discussed in Section 2.3 of Chapter 2, with further discussion of the STAR-type model estimation procedure in more technical detail. Empirical results in Section 3.3 contain plots and diagrams with descriptive statistics for each time-series considered in this chapter, as well as the results of the non-linearity tests. The estimated models then are tested for goodness of fit and the results are presented in the view of the forecasting exercise. Linear and non-linear forecasts are compared in terms of forecasting performance using a number of tests of forecasting accuracy, including the tests of forecasting error magnitude, the Diebold and Mariano test of equal forecasting accuracy, forecast encompassing and trade rule tests. Moreover, the same tests are then applied to the combinations of linear and non-linear forecasts. Section 3.4 summarises the results and concludes.

3.2. Methodology

A simple random walk model and linear ARIMA models will be estimated as benchmarks for the STAR-type models. Forecasting abilities of all linear and non-linear models will then be compared using forecasting accuracy tests. A random walk model with a drift (δ) is applied in this chapter:

$$y_t = \delta + y_{t-1} + \varepsilon_t \tag{3.1}$$

where y_t is the price returns level at time *t*, y_{t-1} is the price returns level at time *t-1* and ε_t is an error term.

The autoregressive integrated moving average process, ARIMA (*p*, *d, q*), is a combination of an autoregressive process of order *p*, AR (*p*), and a moving average of order *q*, MA (*q*), where *d* is the order of integration, or in other words, the number of times the series has to be differenced in order to achieve stationarity. For stationary series *d* equals zero, thus ARIMA (*p, d, q*) becomes ARMA (*p, q*). The general form for the ARIMA (*p*, *d, q*) process is as follows:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ \dots \tag{3.2}$$
$$- \theta_q \varepsilon_{t-q} + u_t$$

where ϕ_p are the coefficients of the AR process component and θ_q are coefficients of the MA process component, and u_t is an error term. ARIMA models are estimated using the Box and Jenkins approach introduced by Box and Jenkins (1976) and involves three stages of model building: identification, estimation and diagnostic checking. The first stage of model identification involves determining the order of the model, i.e. the values of *p* and *q*. The value of integrating order, *d*, is determined following the results of the stationarity test. After parameters estimation, the adequacy of the estimates is

tested with diagnostic checking of the model using an information criteria approach. Akaike's information criteria (AIC) and Schwarz's Bayesian information criteria (SBIC) are the most commonly used procedures in ARIMA modelling. In addition, residual tests such as tests for remaining autocorrelation and ARCH-LM test are performed as model misspecification tests.

Further to linear alternatives this paper will estimate smooth transition-type models for price returns series for the data considered. The formulae for a standard smooth transition (STR) model is as follows:

$$y_t = \phi' z_t + \theta' z_t G(\gamma, c, s_t) + u_t \quad , \quad u_t \sim iid(0, \delta^2) \qquad (3.3)$$

$$G(\gamma, c, s_t) = \left(1 + exp\left\{-\gamma \prod_{K=1}^{K} (s_t - c_K)\right\}\right)^{-1} \quad , \quad u_t \sim iid(0, \delta^2) \qquad (3.4)$$

where ϕ is a parameter of the linear part of the equation and θ' is a parameter of the non-linear part. $G\ (\gamma, c, s_t)$ is the transition function which depends on the transition variable, s_t, the slope parameter, γ, and the vector of location parameters, c. The transition variable, s_t, can be either part of z_t, which in the case of SETAR (self-exciting threshold autoregressive) will be the dependent variable itself, y_t, or the transition variable can be represented by another variable, such a trend, for instance. The term K can be set either to unity ($K = 1$) to attain an LSTAR (logistic STAR)

model, or it can be set to be equal to two ($K = 2$) for an ESTAR (exponential STAR) model.

There are three stages in smooth transition modelling, which include specification, estimation and evaluation. The initial stage of specification involves testing the time-series for the presence of STAR-type non-linearity and choosing the transition variable. The results will suggest whether LSTAR, ESTAR or a linear model should best fit the data. Furthermore, the estimation phase involves finding the starting values for non-linear estimation through a grid search and estimating the model based on those starting values. Results are then evaluated using a number of tests, such as misspecification tests, autocorrelation of the disturbance term, test for remaining non-linearity, ARCH test and test of non-normality. There are also graphical tests that might give an indication of whether the model was estimated correctly.

Once the significant non-linearity is reported and either ESTAR or LSTAR models are chosen, a non-linear optimisation routine known as a grid search is applied in order to estimate the starting values of STAR model parameters. The grid search requires the transition variable, s_t, to be known, which is accomplished in the first stage of the specification. The procedure involves creating a linear grid within a vector of location parameters, c, and a long-linear grid in the slope parameter, γ, and calculating the residual sum of squares for each of those values. The values that offer the minimum residual sum of squares are chosen as starting values for model estimation.

After the starting values have been established the Newton-Raphson algorithm is applied to maximise the likelihood function which estimates the remaining parameters of the model. Further misspecification tests are carried out on the estimated model

including a test for remaining residual autocorrelation, a test of parameter constancy, the ARCH-LM test and the Jarque-Bera normality test. In addition, the model can be tested for any remaining additive STAR-type non-linearity. The parameter constancy test, in its turn, tests whether parameters are constant or continuously change. In addition, graphical analysis may serve as a good indication tool. Thus, the tests that allow to determine validity and the goodness of fit of the estimated models used in this chapter include a test of no error autocorrelation, a test of no remaining non-linearity, and the ARCH-LM test.

The test of no error autocorrelation used in this study is based on the test commonly known as the Breusch-Godfrey test. In the case of STAR modelling this particular test is preferred over the more popular Durbin-Watson autocorrelation test. The reason for this is that the Durbin-Watson test is constructed in a way that tests relationship only between an error and its immediate previous value. In other words, it is only valid if autocorrelation is present in the first lag. The Breusch-Godfrey test, on the other hand, examines the relationship between an error and several lagged error values at the same time. Another reason for not choosing the Durbin-Watson test is that for the test to be valid there are certain conditions that have to be fulfilled, including a constant term in the regression and non-stochastic regressors. In addition, the regression must not contain lags of dependent variable. In other words, the regression should be static in nature, as opposed to dynamic. These conditions defy the very essence of the STAR-type modelling and thus a different approach is required.

However, Brooks (2002) points out that the Breusch-Godfrey test presents some difficulty in its conduct in terms of determining the appropriate value of the number of lags of residuals, r. As there is no particular rule or procedure for choosing the correct

value, it is usually down to a researcher to employ a trial-and-error approach. The frequency of data might give an initial idea about the number of lags. The test is a joint hypothesis test with a critical value following a Chi-squared distribution. The Breusch-Godfrey test for autocorrelation of rth order involves regressing residuals u_t estimated using OLS:

$$u_t = p_1 u_{t-1} + p_2 u_{t-2} + p_3 u_{t-3} + \cdots + p_r u_{t-r} + v_t \qquad , \; v_t \sim N(0, \sigma_v^2) \quad (3.5)$$

where the error term follows normal distribution, $v_t \sim N(0, \sigma_v^2)$. The test statistic following Chi-square distribution is: $(T - r)R^2 \sim \chi_r^2$, where T is the number of observations and R^2 is obtained from the above regression (3.5). The null hypothesis of no serial correlation to the order of r is tested against the alternative of autocorrelation.

$H_0: p_1 = 0 \; and \; p_2 = 0 \; and \; \ldots \; p_r = 0$

$H_1: p_1 \neq 0 \; or \; p_r \neq 0 \; or \; \ldots \; p_r \neq 0$

H_0 of no serial correlation is rejected if the test statistic is greater than the value of the critical value from the Chi-squared statistical tables.

Another test considered here is a test of no remaining non-linearity, which is based on the account that in the case of a correctly fitted model the residuals should contain no

remaining non-linear structure. The test naturally assumes that the remaining non-linearity is a STAR-type non-linearity.

$$y_t = \phi' z_t + \theta' z_t G(\gamma_1, c_1, s_{1t}) + \psi' z_t H(\gamma_2, c_2, s_{2t}) + u_t \tag{3.6}$$

where $u_t \sim iid(0, \sigma^2)$ and H is a transition function for that regression, i.e. different from the one used in the main model. The alternative hypothesis is defined as:

$$y_t = \beta_0' z_t + \theta' z_t G(\gamma_1, c_1, s_{1t}) + \sum_{j=1}^{3} \beta_j' \tilde{z}_t s_{2t}^j + u_t^* \tag{3.7}$$

The following auxiliary model is used to test the above model, where $\hat{u}_t$ is regressed on $\left(\hat{z}_t' s_{2t}, \hat{z}_t' s_{2t}^2, \hat{z}_t' s_{2t}^3\right)'$ and the partial derivatives of the log-likelihood function with respect to the parameters of the alternative model. The null hypothesis for this test of no remaining non-linearity is that $\beta_1 = \beta_2 = \beta_3 = 0$. The test statistic follows F-distribution and is treated in the same fashion as a standard non-linearity test.

The ARCH-LM test is used to test for presence of ARCH in the residuals (Engle, 1982). The residuals, $\hat{u}_t$, of a regression in question are squared and regressed on their own lags. The number of lags signifies the order of ARCH the test is run for. Hence, the regression for the ARCH test of order q will be as follows:

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \gamma_2 \hat{u}_{t-2}^2 + \cdots + \gamma_q \hat{u}_{t-q}^2 + \upsilon_t \quad (3.8)$$

where q is number of lags and ν_t is an error term. The value of R^2 obtained from this regression forms the test statistic, TR^2, where T is the number of observations. TR^2 is compared to a critical value obtained from the Chi-squared distribution table $\chi^2(q)$ to test the following hypotheses:

$$H_0: \gamma_1 = 0 \; and \; \gamma_2 - 0 \; and \; \gamma_3 = 0 \; and \; \ldots and \; \gamma_q = 0$$

$$H_1: \gamma_1 \neq 0 \; or \; \gamma_2 \neq 0 \; or \; \gamma_3 \neq 0 \; or \; \ldots or \; \gamma_q \neq 0$$

If test statistic is greater than the critical value, the null hypothesis of no ARCH is rejected.

3.3. Empirical results

This study will analyse daily time-series data over a twenty year period from 1st January 1988 to 31st December 2007, which consists of 5217 observations. The data consists of four price indices of major world economies. These include FTSE 100 for UK; S&P

500 Composite for US; DAX 30 Performance for Germany; and Nikkei 225 Stock Average for Japan.[5]

Descriptive statistics

Descriptive statistics are carried out to give an initial indication of the nature of the data and include plots of the data against time, histograms, measures of central tendency and dispersion, and normality tests. A histogram provides a good insight into the shape of the distribution of the data, while skewness and kurtosis indicate the symmetry and thickness of the tails of a distribution respectively. The Jarque-Bera statistic is generally regarded as a good measure of normality of the distribution. It follows a chi-square distribution with two degrees of freedom.

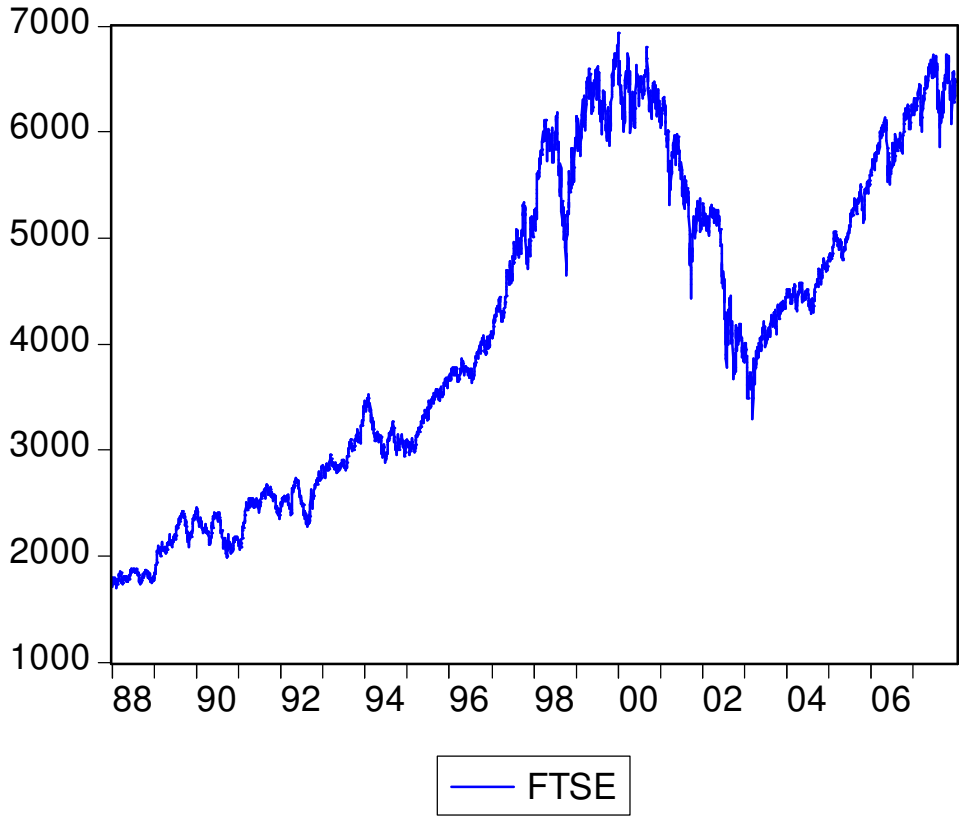
The diagram below (Figure 3.1) illustrates the FTSE 100 index plotted against time. The observation shows that the values for the index have increased dramatically in the late 1990s with a decline over early 2000s, following less dramatic increase toward the end of the sample. Moreover, up to late the 1990s it is seems to be less volatile compared to the early 2000s, and again displaying less volatile behaviour between 2003 and the beginning of 2007.

The period of 1995 – 2001 is of a particular interest, as can be seen on the diagram below. The period is known as the dot-com bubble or the IT bubble. During this period stock markets of Western economies showed a rapid growth in the Internet sector and

[5] All the data was obtained from the Datastream database. All estimations and tests were performed using EViews 3.1, PCGive 10 and JMulTi software.

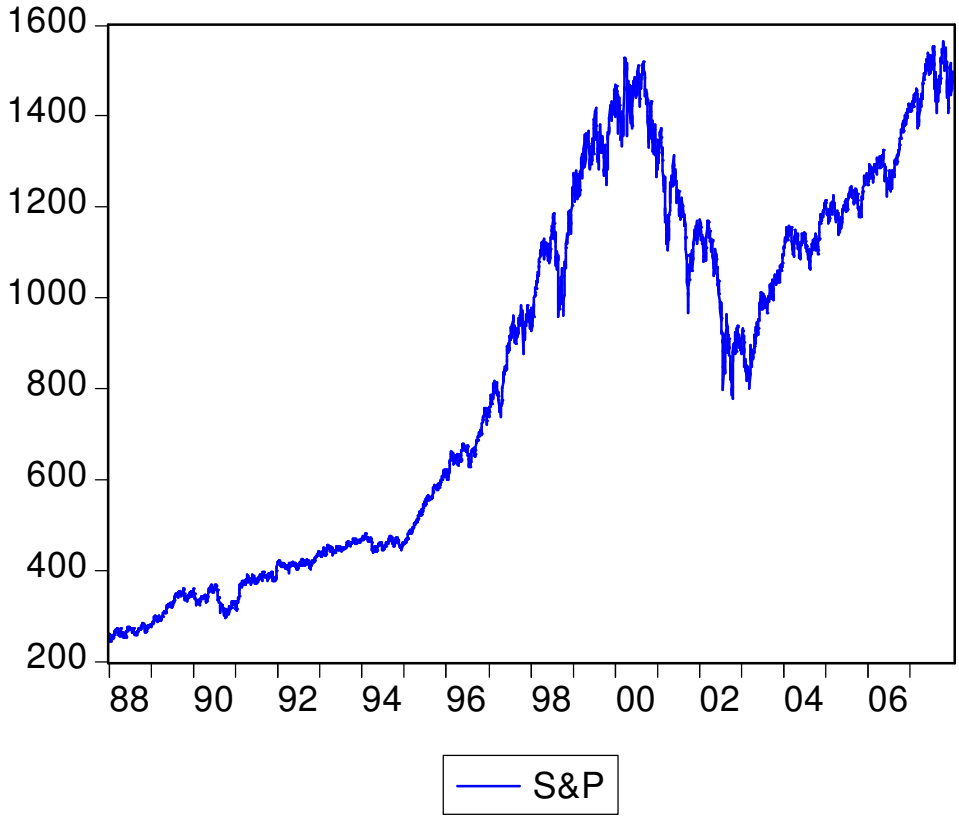
similar or related industries. The IT bubble was created when speculators started to buy stocks which showed a fast increase in value in the expectation that the stock price will increase even further. However, for most of these shares, the price did not reflect their true value, and as a result large number of companies' stock prices became overvalued. Subsequently, the bubble burst, causing the share prices to fall dramatically and many businesses thus endured bankruptcy. The effects of the dot-com bubble can also be seen on the S&P and DAX indices. In addition, for all time-series the period of the dot-com bubble is accompanied by a number of outliers or extreme data points.

Figure 3.1. FTSE index time-series data.



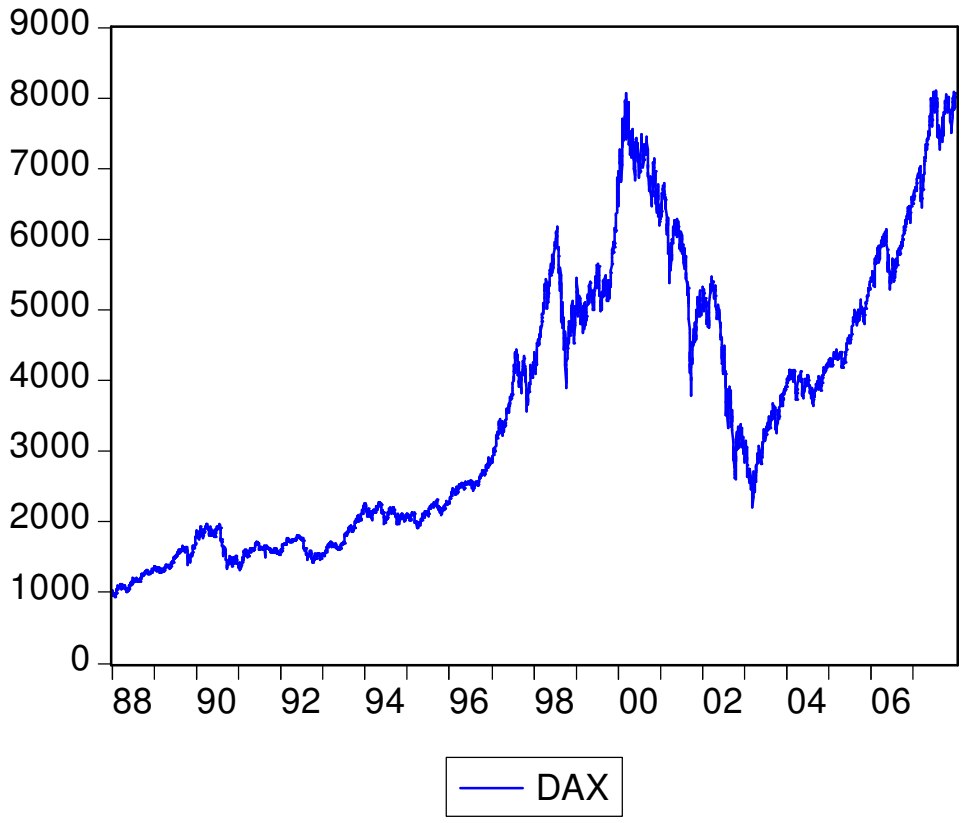
According to the plot of data in Figure 3.2, the US price index has a similar pattern as seen in the FTSE 100 index.

Figure 3.2. S&P index time-series data.



The German DAX price index displays once again a similar pattern seen in UK and US indices. However, even though the index seems to be affected analogously with the UK and US, the rapid falls in values appear to be sharper than those of previous time-series.

Figure 3.3. DAX index time-series data.



The Japanese price index somewhat appears to be very different to the UK, US and German indices. It corresponds with other series in some dramatic movements, however it seems to react to those outliers differently. In addition, the overall pattern diverges from the common outline of the other three time-series. The index has a sequence of fairly high values in the beginning of the sample, which eventually descends in very rapid fragments maintaining the rate throughout, with the exception of the early 2000s.

Historically, Japan is an industrial-based economy focusing on manufacturing and processing industries due to the deficiency of natural resources, which in turn explains the lack of agricultural industry. The economy is characterised by being very efficient and competitive, however is limited to international trade in some sectors. The

remarkable economic performance in the post-war period accelerated Japan into becoming one of the most successful developed economies, continuing into the 1980s with high growth of high-technology industries. However, the economic bubble of the 1980s resulted in over-investment coupled with banks allowing risky loans, consequently culminating to the Tokyo Stock Exchange crash in 1989. This event is clearly visible on the time-series plot as a sharp drop in the index after reaching its all-time high. Subsequent to the lowest value of the Nikkei index in 2003, the Japanese economy seems to undergo a sustained recovery up to the end of the sample. For more detailed discussion of Japanese economy cycles refer to Chakraborty (2009).

Figure 3.4. Nikkei index time-series data.

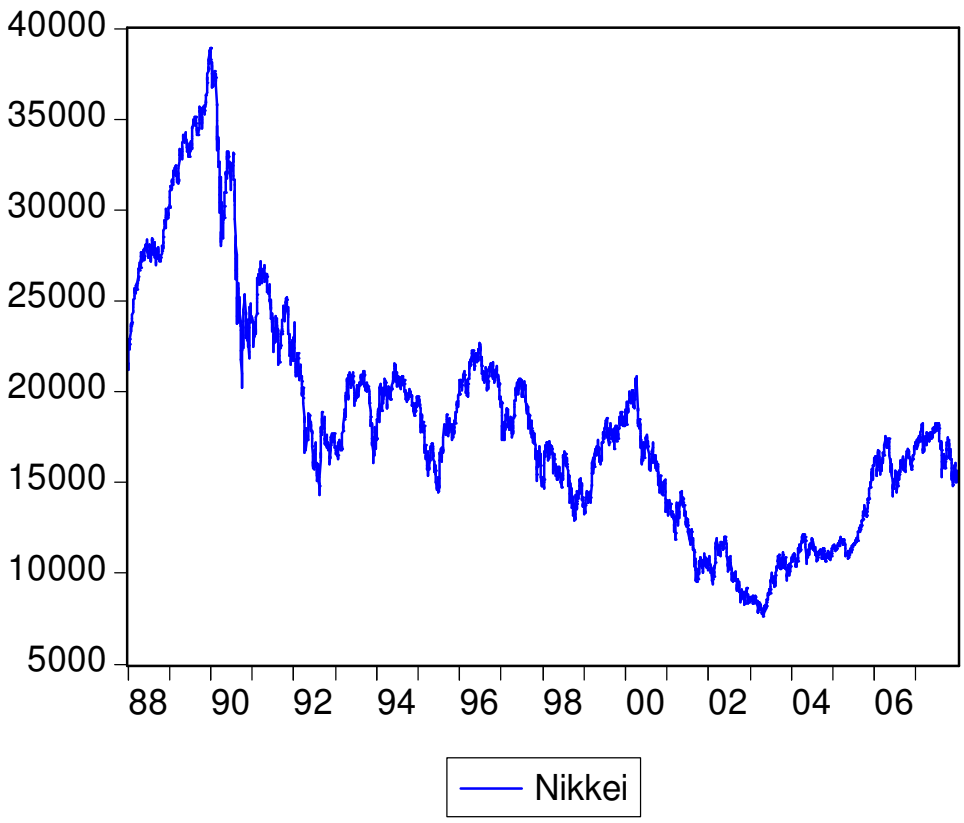
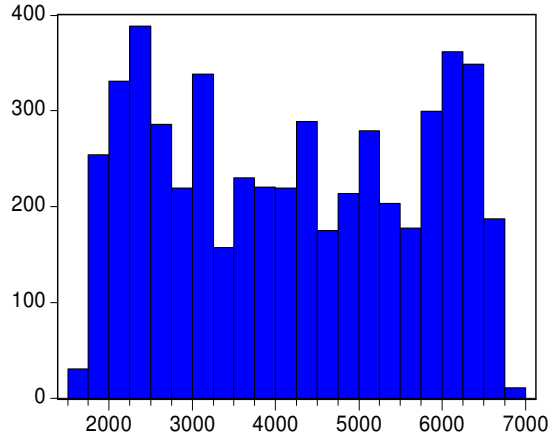


Figure 3.5 represents histograms and main descriptive statistics for real data for all four indices. The values of skewness for FTSE, S&P, DAX and Nikkei indices are all positive and close to zero, which indicates the thickness of the upper tail of distribution, meaning that the distribution seems to be skewed to the right. This suggests that all distributions are characterised by comparatively few high values. In addition, the kurtosis value for FTSE, S&P and DAX indicates a thin tail, which can easily be seen on the diagram. The Nikkei index, on the other hand, has a distribution with thicker than normal distribution tails. For all four series considered here the Jarque-Bera statistic was greater than the critical value hence, the hypothesis of normality was rejected for all four time-series.

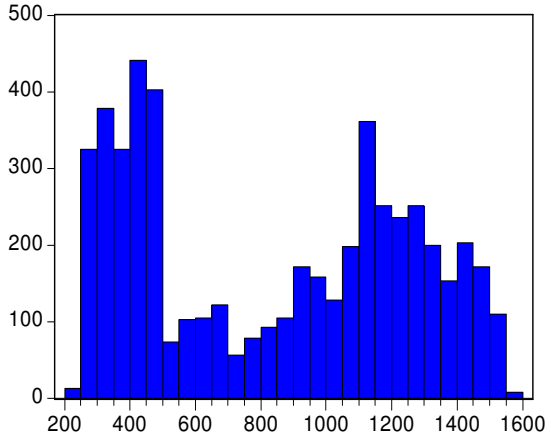
Nevertheless, the above descriptive statistics were performed on the actual levels of the time-series, i.e. prices, as opposed to the returns series. Hence, while the analysis of the real data provides an overview of the series, it is normally not expected to draw any strong conclusions from such results due to high volatility and strong probability of non-stationarity of the data. Moreover, this investigation is concerned with predictability of the stock price returns.

Figure 3.5. Real time-series data histograms: FTSE, S&P, DAX, Nikkei.



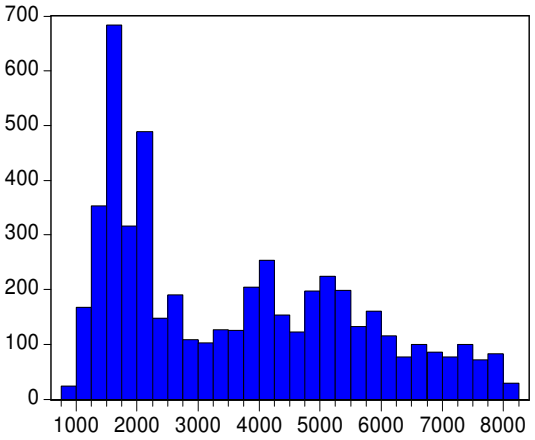
Series: FTSE
Sample 1/01/1988 12/31/2007
Observations 5217

Mean	4185.424
Median	4161.060
Maximum	6930.200
Minimum	1694.500
Std. Dev.	1522.962
Skewness	0.037111
Kurtosis	1.655411
Jarque-Bera	394.1937
Probability	0.000000



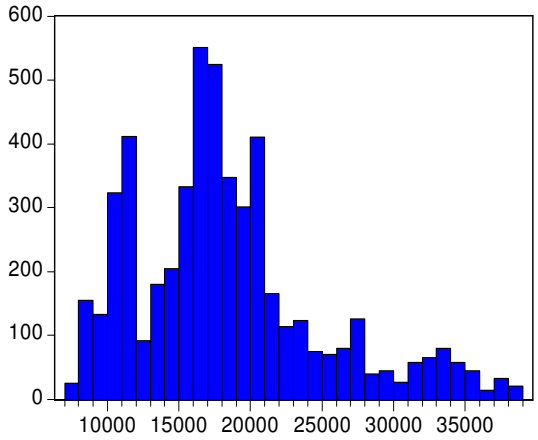
Series: S&P
Sample 1/01/1988 12/31/2007
Observations 5217

Mean	841.4577
Median	896.3400
Maximum	1565.150
Minimum	242.6300
Std. Dev.	410.3756
Skewness	0.044028
Kurtosis	1.488340
Jarque-Bera	498.4129
Probability	0.000000



Series: DAX
Sample 1/01/1988 12/31/2007
Observations 5217

Mean	3619.340
Median	3315.930
Maximum	8105.690
Minimum	931.1800
Std. Dev.	1963.179
Skewness	0.511617
Kurtosis	2.061022
Jarque-Bera	419.2481
Probability	0.000000



Series: NIKKEI
Sample 1/01/1988 12/31/2007
Observations 5217

Mean	18319.23
Median	17358.19
Maximum	38915.87
Minimum	7607.880
Std. Dev.	6505.680
Skewness	0.907362
Kurtosis	3.623321
Jarque-Bera	800.3213
Probability	0.000000

There are a number of ways the returns can be calculated, however, for the purpose of a forecasting exercise, which is the main intent of this chapter, the returns are calculated as a first difference of the logarithm of the original data. The returns data tends to be more stable and stationary compared to the price series. Diagrams and histograms are provided in figures below (Figure 3.6 – 3.7). It is evident from the returns diagrams that high volatility in prices corresponds with high volatility in returns with economic bubbles characterised by extreme outliers in returns series. For most indices the widest spread of stock returns is during the bubble of the late 1990s and early 2000s. For the Japanese index, on the other hand, high volatility is observed throughout the sample in a fairly consistent pattern with a distinctive outlier in 1989 which equates to the Tokyo Stock Exchange crash. For all returns series, with the exception of Japan, negative value for skewness and excess kurtosis indicate distributions to have thick tails and to be skewed to the left, suggesting that all series have relatively few low values. This effect can easily be observed on the diagrams. The distribution of the Nikkei index seems to be closer to the shape of the normal distribution with it being slightly skewed to the right. However, similarly to the result of the price series analysis, as expected, the hypothesis of normality is still rejected for all of the time-series returns.

Figure 3.6. Returns series: FTSE, S&P, DAX, Nikkei.

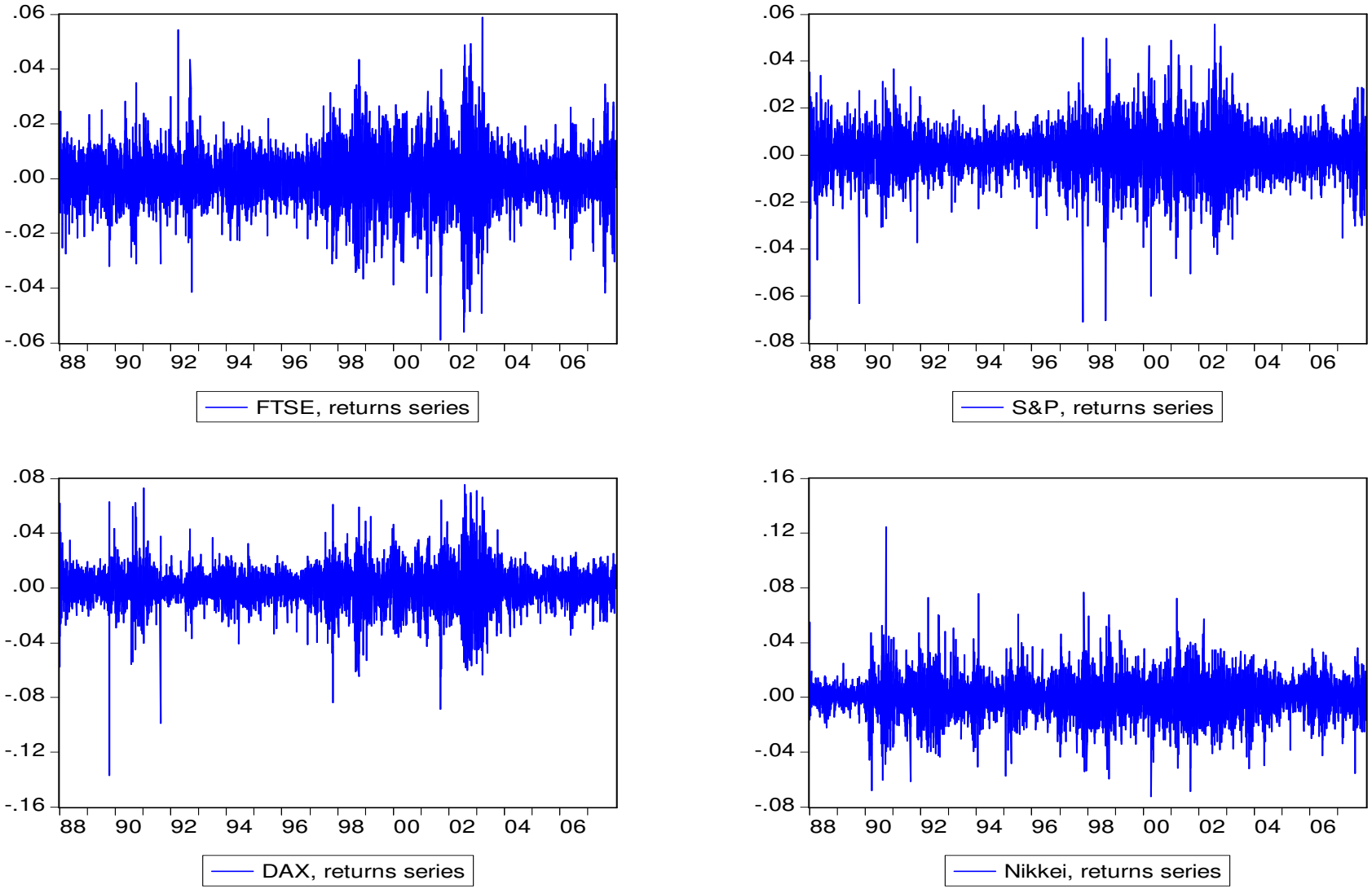
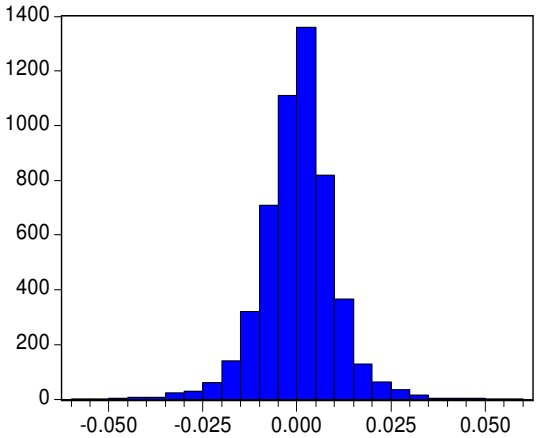
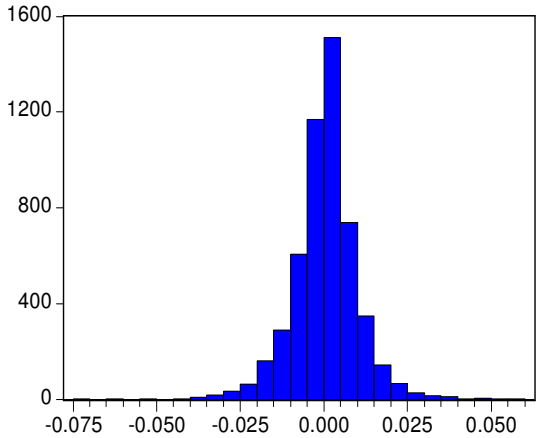


Figure 3.7. Returns time-series data histograms: FTSE, S&P, DAX, Nikkei.



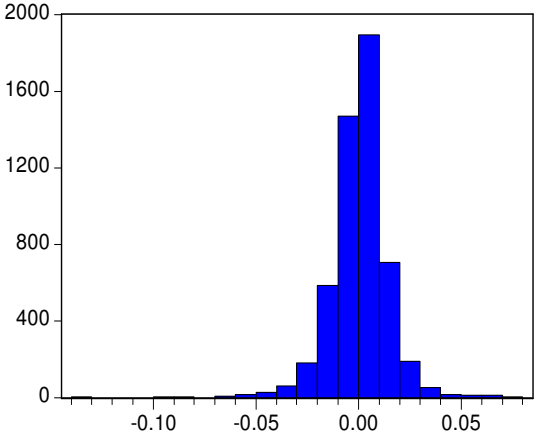
Series: FTSE, returns series
Sample 1/04/1988 12/31/2007
Observations 5216

Mean	0.000254
Median	0.000131
Maximum	0.059026
Minimum	-0.058853
Std. Dev.	0.009849
Skewness	-0.146365
Kurtosis	6.217022
Jarque-Bera	2267.857
Probability	0.000000



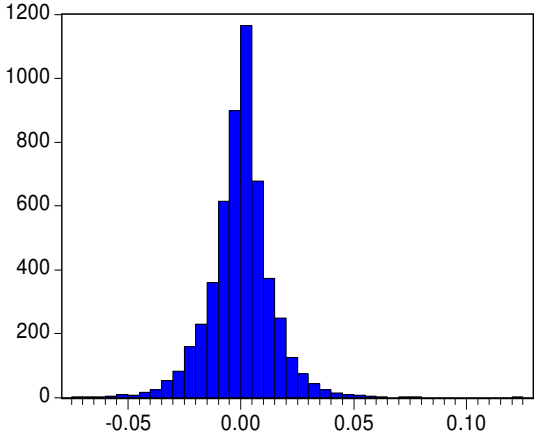
Series: S&P, returns series
Sample 1/04/1988 12/31/2007
Observations 5216

Mean	0.000342
Median	0.000258
Maximum	0.055732
Minimum	-0.071127
Std. Dev.	0.009758
Skewness	-0.234589
Kurtosis	7.572150
Jarque-Bera	4591.098
Probability	0.000000



Series: DAX, returns series
Sample 1/04/1988 12/31/2007
Observations 5216

Mean	0.000400
Median	0.000479
Maximum	0.075527
Minimum	-0.137099
Std. Dev.	0.013626
Skewness	-0.421770
Kurtosis	9.189051
Jarque-Bera	8479.459
Probability	0.000000



Series:Nikkei, returns series
Sample 1/04/1988 12/31/2007
Observations 5216

Mean	-6.57E-05
Median	0.000000
Maximum	0.124303
Minimum	-0.072340
Std. Dev.	0.013616
Skewness	0.137958
Kurtosis	6.934756
Jarque-Bera	3381.367
Probability	0.000000

Unit root tests

The augmented Dickey-Fuller (ADF) test is the most commonly used unit root test due to its simplicity and ability to account for autocorrelated error term. The ADF test statistic follows a non-standard distribution, hence, a set of special critical values are used. Brooks (2002) offers a standard equation for ADF unit root test as follows:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{3.9}$$

where Δ is the difference operator which indicates how many times the series has to be differenced in order to achieve stationarity, ψ is the test statistic, p is the number of lags of the dependent variable and u_t is an error term. There is no strict rule on choosing the number of lags of the dependent variable, hence augmented test lags are chosen on the basis of frequency of data combined with a previous knowledge from similar studies, whilst ensuring white noise residuals. The procedure tests a null hypothesis of unit root against an alternative of stationarity.

$H_0: \psi = 0, \qquad series\ contains\ unit\ root$

$H_1: \psi < 0, \qquad series\ is\ stationary$

For all price indices the null hypothesis of unit root could not be rejected (Table 3.1). As expected, since all the test statistics were more negative than the critical values, the null hypothesis of unit root was rejected for all returns series, hence implying stationarity for all returns. The result was the same for all three variations of the ADF test where the dependent variable is a random walk, a random walk with a drift (intercept), or a random walk with a drift around a stochastic trend (intercept and trend) (Gujarati, 2003).

Testing the presence of non-linearity

The initial stage of STAR model specification involves testing the time-series for the presence of STAR-type non-linearity and choosing the appropriate transition variable. The transition variable, s_t, can be either part of z_t, the dependent variable itself (y_t), or the transition variable can be represented by a trend. The results of the non-linearity test will also suggest whether the LSTAR, ESTAR or a linear model should best fit the data. The following auxiliary regression is applied if the transition variable s_t is an element of z_t:

$$y_t = \beta_0' z_t + \sum_{j=1}^{3} \beta_j' \tilde{z}_t s_t^j + u_t^* \tag{3.10}$$

Table 3.1. ADF test results for price series and returns data series.

PRICES								
ADF test statistic	FTSE	S&P	DAX	Nikkei	Critical value at 1%	Critical value at 5%	Critical value at 10%	Conclusion
Intercept	-1.1701	-0.7453	-0.2110	-1.3593	-3.4348	-2.8626	-2.5674	non-stationary
Intercept and trend	-1.9945	-1.9467	-1.4375	-2.1952	-3.9654	-3.4134	-3.1284	non-stationary
None	0.9298	1.2645	1.4673	-0.7688	-2.5662	-1.9394	-1.6156	non-stationary
RETURNS								
ADF test statistic	FTSE	S&P	DAX	Nikkei	Critical value at 1%	Critical value at 5%	Critical value at 10%	Conclusion
Intercept	-52.4203	-52.6816	-52.1287	-53.9532	-3.4348	-2.8626	-2.5674	stationary
Intercept and trend	-52.4217	-52.6873	-52.1250	-53.9486	-3.9654	-3.4134	-3.1284	stationary
None	-52.3757	-52.5911	-52.0641	-53.9564	-2.5652	-1.9394	-1.6156	stationary

where $z_t = (1, \tilde{z}_t)'$, β_0' is a linear parameter, and β_j' is a non-linear parameter.

The following regression is applied when s_t is not an element of z_t:

$$y_t = \beta_0' z_t + \sum_{j=1}^{3} \beta_j' z_t s_t^j + u_t^* \tag{3.11}$$

Thus, the null hypothesis of non-linearity ($H_0: \beta_1 = \beta_2 = \beta_3 = 0$) is tested by applying an *F*-test. The *p*-values of the *F*-test results for the transition variables are represented in the table below (Table 3.2).

Table 3.2. STAR non-linearity tests results.

	Transition variable	***p*-values of *F*-test**				**Suggested model**
		$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	
FTSE	Returns	0.0000	0.0004	0.1522	0.0053	LSTAR
	Trend	0.0002	0.5185	0.0930	0.0000	LSTAR
S&P	Returns	0.0000	0.9989	0.0000	0.0000	ESTAR
	Trend	0.1219	0.6641	0.1352	0.0726	Linear
DAX	Returns	0.0000	0.4678	0.0000	0.0001	ESTAR
	Trend	0.2936	0.2912	0.1094	0.8139	Linear
Nikkei	Returns	0.0032	0.0868	0.0083	0.0482	ESTAR
	Trend	0.0360	0.5169	0.0078	0.2906	ESTAR

Linear and non-linear model estimation

Following the methodology in Section 3.2, a random walk model with a drift (δ) was estimated for all four price returns series.

$$y_t = \delta + y_{t-1} + \varepsilon_t \tag{3.12}$$

where stock returns, y_t, are regressed on their own previous values, y_{t-1}, and a random disturbance term, ε_t.

Subsequent to that, ARIMA (p, d, q) models were estimated using the Box-Jenkins approach:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} + u_t \tag{3.13}$$

where the current stock return, y_t, is depended on the weighted average of the variable's past values (AR component) and past random disturbance terms (MA component) going back p and q periods, respectively, and an *i.i.d.* error term, u_t. Thus, ϕ_p are the autoregressive coefficients, while θ_q are the coefficients of the moving average process.

The integration order, *d*, was set to zero due to stationarity of the returns, thus suggesting an ARMA (p, q) model for price returns. The optimal order of each model

was determined at the diagnostic checking stage of the procedure using the Akaike's information criteria (AIC). The results of the Box-Jenkins procedure confirmed the following models: ARMA (1,3) for FTSE returns series, ARMA (2,1) for S&P, ARMA (3,2) for DAX, and ARMA (0,3) for Nikkei index.

The results of STAR model estimation for all four time-series are represented in tables below (Table 3.3 – 3.8). For each time-series both variable itself, i.e. the return series, and the trend were considered as a transition variable. At the specification stage either LSTAR, ESTAR or a linear model were chosen for each series on the basis of a non-linearity test. Whenever a linear model was suggested, the particular model was disregarded in this particular instance as linear models were already estimated for all series as benchmarks regardless of the non-linearity test. As a result, the LSTAR model was only suggested for the FTSE series with the variable itself and the trend as transition variables, while all other series were estimated using the ESTAR model.

In addition, a test of no error autocorrelation, a test of no remaining non-linearity, and the ARCH-LM test were performed as tests of goodness of fit for each of the estimated non-linear models (Table 3.9). In the no error autocorrelation test for all time-series, the null hypothesis of no autocorrelation could not be rejected suggesting presence of autocorrelation in the errors of the estimated models. The test for no remaining non-linearity demonstrated that the models were estimated correctly as there was no evidence of STAR-type non-linearity present in the residuals. The null hypothesis of no ARCH effect was rejected for all of the time-series, thus assuming the presence of ARCH, which is an expected result for the daily returns series.

Table 3.3. FTSE; Transition variable FTSE returns $(t-1)$; Suggested model LSTAR.

Variable	**Estimate**	**SD**	**t-statistic**
Linear part			
Constant	0.0088	0.0000	0.0000
FTSE returns (t-1)	-4.1186	0.0000	-0.0000
Non-linear part			
Constant	-0.0087	0.0000	-0.0000
FTSE returns (t-1)	4.1513	0.0000	0.0000
Gamma (γ)	0.7369	0.0000	0.0000
C1	-0.0894	0.0000	0.0000

Table 3.4. FTSE; Transition variable TREND; Suggested model LSTAR.

Variable	**Estimate**	**SD**	**t-statistic**
Linear part			
Constant	0.0003	0.0002	1.7431
FTSE returns (t-1)	0.0805	0.0219	3.6718
Non-linear part			
Constant	-0.0002	0.0003	-0.6196
FTSE returns (t-1)	-0.1869	0.0439	-4.2553
Gamma (γ)	6.9113	3.9482	1.7505
C1	3499.7574	218.3356	16.0292

Table 3.5. S&P; Transition variable S&P returns $(t-1)$; Suggested model ESTAR.

Variable	**Estimate**	**SD**	**t-statistic**
Linear part			
Constant	-0.0007	0.0014	-0.5400
S&P returns (t-1)	-0.0311	0.0329	-0.9469
Non-linear part			
Constant	0.0150	0.0058	2.6237
S&P returns (t-1)	-0.3434	0.1027	-3.3438
Gamma (γ)	0.1917	0.1128	1.6999
C1	-0.0582	0.0075	7.7640
C2	0.0237	0.0050	4.7520

Table 3.6. DAX; Transition variable DAX returns $(t-1)$; Suggested model ESTAR.

Variable	**Estimate**	**SD**	**t-statistic**
Linear part			
Constant	0.0003	0.0002	1.8172
DAX returns (t-1)	0.0220	0.0169	1.3053
Non-linear part			
Constant	0.0088	0.0016	5.5283
DAX returns (t-1)	-0.2562	0.0381	-6.7278
Gamma (γ)	2.4211	1.6129	1.4862
C1	-0.0561	0.0014	40.1357
C2	0.0265	0.0037	7.1675

Table 3.7. Nikkei; Transition variable Nikkei returns $(t-1)$; Suggested model ESTAR.

Variable	**Estimate**	**SD**	**t-statistic**
Linear part			
Constant	-0.0094	0.0056	-1.6880
Nikkei returns (t-1)	-0.3853	0.1885	-2.0442
Non-linear part			
Constant	0.0095	0.0057	1.6670
Nikkei returns (t-1)	0.3734	0.1832	2.0384
Gamma (γ)	0.8411	0.6750	1.2461
C1	-0.0677	0.0103	6.5737
C2	-0.0162	0.0074	2.1972

Table 3.8. Nikkei; Transition variable TREND; Suggested model ESTAR.

Variable	**Estimate**	**SD**	**t-statistic**
Linear part			
Constant	-0.00128	0.0021	-0.5997
Nikkei returns (t-1)	-0.29646	0.4505	-0.6581
Non-linear part			
Constant	0.0040	0.0000	0.0000
Nikkei returns (t-1)	0.9613	0.0000	0.0000
Gamma (γ)	0.2491	0.3771	0.6605
C1	-273.3744	0.0000	0.0000
C2	6090.2297	0.0000	0.0000

Table 3.9. Goodness of fit tests results.

	FTSE (LSTAR)	**FTSE Trend (LSTAR)**	**S&P (ESTAR)**	**DAX (ESTAR)**	**Nikkei (ESTAR)**	**Nikkei Trend (ESTAR)**
Test of no error autocorrelation						
F-value	5.7465	2.6071	0.0258	1.1503	35.1559	0.0724
Critical value	254	254	254	254	254	254
Test of no remaining non-linearity						
Transition variable	FTSE returns	FTSE returns	S&P returns	DAX returns	Nikkei returns	Nikkei returns
F-value	0.0074	0.0003	0.5766	0.2849	0.9689	0.0024
Critical value	254	254	254	254	254	254
ARCH-LM test						
Test statistic	850.9093	864.8856	322.0965	449.7600	305.5739	289.8454
Critical values	146.57	146.57	146.57	146.57	146.57	146.57
F- statistic	203.4019	207.4078	68.6643	98.4509	64.9226	61.3840
Critical values	4.37	4.37	4.37	4.37	4.37	4.37

Forecasting

Subsequent to the estimation and testing procedures, the following models are used to carry out an out-of-sample one-step ahead recursive forecast, where *STAR-trend* models refer to STAR-type models with the trend being estimated as the transition variable, as opposed to the returns series itself:

Table 3.10. Linear and non-linear models of daily stock returns.

	Linear forecast	*Non-linear forecast*
FTSE	Random walk model	LSTAR
	ARMA (1,3)	LSTAR-trend
S&P	Random walk model	ESTAR
	ARMA (2,1)	
DAX	Random walk model	ESTAR
	ARMA (3,2)	
Nikkei	Random walk model	ESTAR
	ARMA (0,3)	ESTAR-trend

The twenty year period data from 1st January 1988 to 31st December 2007, which consists of 5217 observations, is divided into in- and out-of-sample, where the seven year holdout, or evaluation sample, is set to be in the range from 28th December 2000 to 31st December 2007, thus including 1825 observations. This study anticipates that the sample of nearly thirteen years of daily data combined with a recursive approach and allowing for non-linear applications will be sufficient to elevate the behaviour patterns

in the data and hence produce a relatively accurate forecast. The results of the forecasts are then further assessed by comparing the forecasting ability of each model using various methods of forecast accuracy measures including statistical loss functions and the technical trading rule approach.

Forecasting accuracy tests

All the forecasts generated in this chapter are assessed and compared in terms of their forecasting accuracy in expectation of determining which of the models produces a superior forecast. Each forecast will be assessed using a range of measures of forecasting accuracy including conventional statistical measures, such as ME, MAE and RMSE; Diebold and Mariano tests of equal forecast accuracy; forecast encompassing tests; statistical measures of combined forecasts; and trading rule style forecasting accuracy tests. Since the aim of this exercise is to determine whether any specific model demonstrates considerable superiority over other models, all the forecasts will be assessed and compared within each separate time-series, as opposed to inter-comparison across all the data sets.

ME, MAE, RMSE

Table 3.11 includes the results of standard statistical measurements for each single forecast. It is evident from the table that the random walk model for all series seems to have the best value for almost all statistics. However, the differences between the values

are extremely diminutive and, thus, it is difficult to determine any definite conclusions at this point whether these differences are statistically significant.

Table 3.11. ME, MAE and RMSE statistics for daily returns data forecasts.

FTSE	**Random Walk**	**Linear**	**LSTAR**	**LSTAR (Trend)**
ME	0.00002*	-0.0002	-0.0002	-0.0002
MAE	0.0077*	0.0077*	0.0077*	0.0077*
RMSE	0.0111*	0.0111*	0.0112	0.0111*
S&P	**Random Walk**	**Linear**	**ESTAR**	
ME	0.00005*	-0.0003	-0.0002	
MAE	0.0073*	0.0073*	0.0074	
RMSE	0.0104*	0.0104*	0.0105	
DAX	**Random Walk**	**Linear**	**ESTAR**	
ME	0.0001*	-0.0002	-0.0002	
MAE	0.0106*	0.0106*	0.0106*	
RMSE	0.0154*	0.0154*	0.0154*	
NIKKEI	**Random Walk**	**Linear**	**ESTAR**	**ESTAR (Trend)**
ME	0.00005*	0.0002	0.0002	0.0001
MAE	0.0096*	0.0096*	0.0096*	0.0096*
RMSE	0.0132*	0.0133	0.0132*	0.0133
Note: * indicates the best statistic				

Diebold – Mariano test

The Diebold-Mariano test of equal forecasting accuracy (Diebold and Mariano, 1995) assesses whether the differences in MSEs of competing forecasts are statistically significant. According to the theory of the test, lower values of MSEs of one forecast in

comparison to the alternative do not necessarily translate into the superiority of this forecast. The test statistic follows standard normal distribution and tests the null hypothesis of equal forecast accuracy against the alternative:

$$S_1 = \left[\hat{V}(\bar{d})\right]^{-\frac{1}{2}}\bar{d} \tag{3.14}$$

where $\bar{d}$ is the mean of the coefficient d_t, which is the difference between the sets of squared forecast errors from two competing models, $d_t = e_{1t}^2 - e_{2t}^2$; and $\hat{V}(\bar{d})$ is an estimate of the variance of $\bar{d}$.

The modified Diebold-Mariano (Harvey et al., 1997) test statistic follows the *t*-distribution with $(t-1)$ degrees of freedom.

$$S_1^* = \left[\frac{t + 1 - 2h + t^{-1}h(h-1)}{t}\right]^{\frac{1}{2}} S_1 \tag{3.15}$$

where S_1 is the original Diebold-Mariano test statistic for h-steps ahead forecast for time *t*. Critical values for the modified test are taken from the Student's *t*-distribution

Table 3.12 includes results of the standard Diebold-Mariano and modified Diebold-Mariano tests. All test statistics following the modified Diebold-Mariano test are insignificant at 1%, 5% and 10% levels of significance, which according to the test,

implies that none of the differences between the MSEs of forecasting models considered here are statistically significant.

Table 3.12. Diebold-Mariano test results for daily returns data forecasts.

	DM test statistic	**Modified DM test statistic**
FTSE		
Random walk – ARMA (1,3)	- 0.01379	- 0.01376
Random walk – LSTAR	- 0.05079	- 0.05067
ARMA (1,3) – LSTAR	- 0.03064	- 0.03057
Random walk – LSTAR trend	- 0.04415	- 0.04405
ARMA (1, 3) – LSTAR trend	- 0.00728	- 0.00726
S&P		
Random walk – ARMA (2, 1)	- 0.01739	- 0.01735
Random walk – ESTAR	- 0.05240	- 0.05228
ARMA (2, 1) – ESTAR	- 0.02578	- 0.02572
DAX		
Random walk – ARMA(3, 2)	- 0.04116	- 0.04106
Random walk – ESTAR	- 0.01629	- 0.01625
ARMA (3, 2) - ESTAR	0.02169	0.02164
Nikkei		
Random walk – ARMA (0, 3)	- 0.00496	- 0.00494
Random walk – ESTAR	0.01000	0.00997
ARMA (0,3) – ESTAR	0.00984	0.00982
Random walk – ESTAR trend	- 0.02311	- 0.02306
ARMA (0,3) – ESTAR trend	0.0000535	0.0000534

Forecast encompassing test

This study will use two types of forecast encompassing tests. The first considers whether one forecast encompasses the other, whereas the second test considers whether the forecast errors of one model can explain the forecast errors of the other model. Hence, the first forecasting encompassing test implemented here tests whether the forecasts from a simple linear random walk model encompasses STAR-type model forecasts for each series. The STAR-type models will also be tested against linear ARMA models. The following equation is a variation of model used by Fang (2003) and is considered for the former type of encompassing test:

$$y_{t+s} = \alpha + \beta_1 f_{t,s}^{RW} + \beta_2 f_{t,s}^{STAR} + u_t \tag{3.16}$$

where $f_{t,s}^{RW}$ is the forecast obtained from a random walk model and $f_{t,s}^{STAR}$ is the forecast generated by the STAR model. The forecast encompassing test used in this study also allows for a constant and an error term. Moreover, the hypothesis testing procedure is based on the encompassing test methodology applied by Clements and Harvey (2007), where the null hypothesis of $\beta_2 = 0$ is tested against a one-sided alternative of $\beta_2 > 0$. Thus, the statistical significance of the β_1 coefficient will signify the first forecasting model encompassing the alternative forecast, and the positive statistical significance of the β_2 coefficient will indicate the first model being encompassed by the alternative forecast. Since this paper does not impose the restriction of unity of the sum of the coefficients, the statistical significance of both coefficients will imply that both

forecasting models contain independent information required for the forecasting of the dependent variable.

Results in Table 3.13 present the forecast encompassing test results for all the daily data returns forecasts. As a result, β_2 coefficients for all the data sets and models were found to be insignificant at 5% level of significance. β_1 coefficients were found to be significant at 5% level of significance for the random walk model for FTSE and S&P series, and for the linear model for FTSE series. Thus, the results suggest that the random walk model encompasses LSTAR and LSTAR-*trend* models for FTSE, and ESTAR for S&P; while the linear ARMA model encompasses LSTAR-*trend* for FTSE series and ESTAR model for DAX. Moreover, for the Nikkei series it seems that neither of the forecasts, including random walk, ARMA model and STAR-type models, are able to contribute significant independent information for returns series forecasting, as insignificance of the coefficients suggest that all of the forecasts are very noisy.

Table 3.13. Forecasting encompassing test.

	t-statistic for β_1	**t-statistic for β_2**
FTSE		
Random walk - LSTAR	- 3.0138*	0.1244
Random walk – LSTAR *trend*	- 2.1510*	0.8706
ARMA – LSTAR	0.7117	- 0.3401
ARMA – LSTAR *trend*	2.3302*	- 3.2003
S&P		
Random walk - ESTAR	2.1118*	- 0.4816
ARMA – ESTAR	0.8332	- 0.9730
DAX		
Random walk - ESTAR	1.8837	0.7870
ARMA - ESTAR	- 2.1991*	0.7226
Nikkei		
Random walk - ESTAR	0.5114	1.4106
Random walk – ESTAR *trend*	1.6505	- 1.3909
ARMA – ESTAR	0.9957	1.5028
ARMA – ESTAR *trend*	1.1227	0.1015
Note: * indicates statistical significance at 5%.		

The second forecast encompassing test used in this study is based on the approach suggested by Fair and Shiller (1990), whereby the regression coefficients are not restricted to equal unity, there is no constraint on the constant term α and the error term is not assumed to be independent and identically distributed (*i.d.d.*). The general equation used in this exercise to implement the forecast encompassing test is as follows:

$$y_{t+s} = \alpha + \beta_1\left(f_{t,s}^{RW} - y_{t,s}\right) + \beta_2\left(f_{t,s}^{STAR} - y_{t,s}\right) + u_t \qquad (3.17)$$

where y_{t+s} is the actual returns series at the time *t*, $f_{t,s}^{RW}$ is the random walk forecast for *s* steps ahead for y_{t+s}, and $f_{t,s}^{STAR}$, consequently, is the STAR model forecast for the same variable and time period. The same method is applied when carrying out forecast encompassing test for linear and STAR models, where $f_{t,s}^{RW}$ is replaced with the forecast generated by the ARMA model, $f_{t,s}^{ARMA}$. The regression is testing the null hypothesis ($H_0: \beta_1 = 0$) of forecasts made by the random walk model to be encompassed by the forecast made with a STAR, and hence containing no relevant information for forecasting the returns series y_t; against an alternative hypothesis ($H_1: \beta_2 > 0$) of a STAR model forecast being encompassed by the random walk model forecast. In order to test the hypotheses, both coefficients, β_1 and β_2, are tested for significance using a standard *t*-test.

The test results (Table 3.14) show that all the β_1 coefficients for forecast models are statistically significant at 5% level of significance, with the exception of linear models in the combination with ESTAR-*trend* for the Nikkei index. Thus, the linear ARMA model encompasses LSTAR and LSTAR-trend models for FTSE index; LSTAR for S&P; ESTAR for DAX; and ESTAR for the Nikkei series. While the random walk model encompasses ESTAR model for S&P; ESTAR for DAX series; and the ESTAR model forecast for Nikkei. However, β_2 coefficients were found to be significant for a few non-linear models together with significant coefficients for linear forecasts, suggesting that both linear and non-linear models contain independent information required for forecasting the price returns series. These combinations include the random walk model and LSTAR model, as well as random walk and LSTAR-*trend* models for FTSE; and a combination of a random walk and ESTAR-*trend* models for the Nikkei series.

Table. 3.14. Forecasting errors encompassing test.

	t-statistic for β_1	**t-statistic for β_2**
FTSE		
Random walk - LSTAR	- 695.0298*	2.7810*
Random walk – LSTAR trend	- 421.6723*	66.0126*
ARMA – LSTAR	- 60.2214*	- 22.7316
ARMA – LSTAR trend	2.5180*	- 73.0253
S&P		
Random walk - ESTAR	- 209.4365*	- 18.5475
ARMA – ESTAR	- 48.4028*	- 34.6721
DAX		
Random walk - ESTAR	- 115.8646*	- 15.6294
ARMA - ESTAR	- 71.9033*	- 27.0736
Nikkei		
Random walk - ESTAR	- 98.5322*	- 3.7355
Random walk – ESTAR trend	- 21.8602*	5.5050*
ARMA – ESTAR	- 25.1779*	- 53.5528
ARMA – ESTAR trend	- 0.3136	- 125.6118
Note: * indicates statistical significance at 5%.		

ME, MAE, RMSE of a combined forecast

Furthermore, following the results of the forecast encompassing test, a combination of linear and non-linear forecasts was performed using a simple weighted average approach. Combination of the forecasts involved running a regression for each data set combining the appropriate linear and non-linear models specified earlier. Thus, each combined forecast involved regressing actual returns (y_{t+s}) on a combination consisting of forecasted series for s steps ahead at time t obtained from a random walk model ($f_{t,s}^{RW}$), linear ARMA model ($f_{t,s}^{ARMA}$) and the STAR model ($f_{t,s}^{STAR}$). The following is a

general equation for the combined forecasts procedure (3.18), while Table 3.15 offers individual equations for each series. In addition, Table 3.16 represents the standard statistics drawn for each of the combined forecasts as an indication of forecasting success.

$$y_{t+s} = \alpha + \beta_1 f_{t,s}^{RW} + \beta_2 f_{t,s}^{ARMA} + (1 - \beta_1 - \beta_2) f_{t,s}^{STAR} + \varepsilon_t \tag{3.18}$$

Table 3.15. Individual combined forecast equations for daily returns data series.

Time series	**Individual combined forecast equation**
FTSE	$y_{t+s}^{FTSE} = \alpha + \beta_1 f_{t,s}^{RW} + \beta_2 f_{t,s}^{ARMA(1,3)} + (1 - \beta_1 - \beta_2) f_{t,s}^{LSTAR} + \varepsilon_t$ $y_{t+s}^{FTSE} = \alpha + \beta_1 f_{t,s}^{RW} + \beta_2 f_{t,s}^{ARMA(1,3)} + (1 - \beta_1 - \beta_2) f_{t,s}^{LSTAR(trend)} + \varepsilon_t$
Nikkei	$y_{t+s}^{Nikkei} = \alpha + \beta_1 f_{t,s}^{RW} + \beta_2 f_{t,s}^{ARMA(0,3)} + (1 - \beta_1 - \beta_2) f_{t,s}^{ESTAR(trend)} + \varepsilon_t$

Table 3.16. Statistics results for combination forecasts of daily returns data series.

FTSE	**Combination**	**Combination trend**
ME	0.00003*	0.0002
MAE	0.0077*	0.0077*
RMSE	0.0111*	0.0111*
NIKKEI		**Combination trend**
ME		-0.0001*
MAE		0.0096*
RMSE		0.0132*
Note: * indicates the best statistic		

When the above statistical results are compared with the results from Table 3.11, it is apparent that statistics for random walk model forecasts and for combined forecasts mostly have the smallest statistical values and hence indicating a possible preference for these models.

Trade rule

In addition to the previous comparative measure, the forecasts were also assessed using the trade rule approach. The trading trigger in this case is whether the forecast level for each data point is above or below zero. Hence, a positive forecast will be a signal to buy (long), and a negative forecast will be a signal to sell (short). The trade rule is run for all the forecasts, including random walk models and linear model forecasts. The results in the table below demonstrate an average return per day for each individual forecasting model using the trading rule. Essentially, the negative return indicates an overall loss, and consequently, a positive value is a result of profit gain.

Table 3.17. Trade rule test results for daily returns forecasting series.

	FTSE	**S&P**	**DAX**	**Nikkei**
Random walk	- 0.0004	0.0007*	0.0006*	0.00007
Linear ARMA	0.0002	0.0002	0.0001	0.00007
ESTAR / LSTAR	- 0.0002	- 0.0004	- 0.0002	0.0002
ESTAR / LSTAR trend	- 0.0001			0.00002
Combination RW and LSTAR	0.0044			
Combination RW and LSTAR-trend	0.0045*			
Combination RW and ESTAR-trend				0.0064*
Note: * indicated the best statistic RW – random walk				

Linear ARMA model forecasts produce positive results for all data sets, followed closely by the random walk model forecast, which generates a loss only for one of the data sets, namely FTSE. The Nikkei index generates positive trade values in all forecasts, with the highest value belonging to the combination forecast of random walk and ESTAR-*trend*, thus suggesting the most stable result across the four series. Moreover, combination forecasts for FTSE and Nikkei generate the best result in terms of the highest average profit per trading day. However, as it was mentioned before, the trading rule approach considered here should not be treated as a realistic profit generating procedure, as it is merely an extensive test of forecasting accuracy. Moreover, the total magnitude of these hypothetical profit gains and losses is somewhat to be desired better, as the comparative difference between those does not seem significant enough to draw strong conclusions. These results are to be expected for daily stock market data, as it is characteristically very noisy. Therefore, it is also expected that longer term data series, for instance, monthly data or long-horizon data, will produce much more reasonable and conclusive results based on the trade rule methodology.

Drawing from the results of the forecasting accuracy tests and taking into account specific behaviour and characteristics of daily data, it can be concluded that the best forecasting model in terms of combination of forecasting accuracy and ease of implementation, the random walk model seems to be the best choice for the purpose of a forecasting exercise. However, there is no clear evidence of the random walk model significantly outperforming the linear and STAR model in terms of forecasting accuracy. The random walk model is preferred in this instance due to the ease of implementation and interpretation.

Interval forecasts for daily data[6]

Interval forecasts provide a prediction of a range of values in which the future value of the variable is expected to lie. This study will apply a technique based on a study by Christoffersen (1998) in order to carry out interval forecasts on the linear and non-linear models considered in this chapter. The methodology involves setting interval prediction barriers in the form of upper and lower limits each with assigned certain probability, with further evaluation of goodness of fit of the forecast using a success ratio approach.

The upper and lower limits are set as a time-series of forecasted values plus or minus respectively the standard error term at the 95% level of confidence assuming normal distribution (Figures 3.8 – 3.11). The goodness of fit test will determine the success rate of the forecast value falling inside the set limits.

[6] The main objective of this thesis is an investigation of point forecasting with non-linear models and does not include a thorough examination into interval forecasts. The subject of interval forecasts is an important area of time-series research that lacks extensive empirical examination in the literature. I would like to thank my examiners for their valuable comments and recommendations for further research within the field of forecasting.

Figure 3.8. FTSE interval random walk.

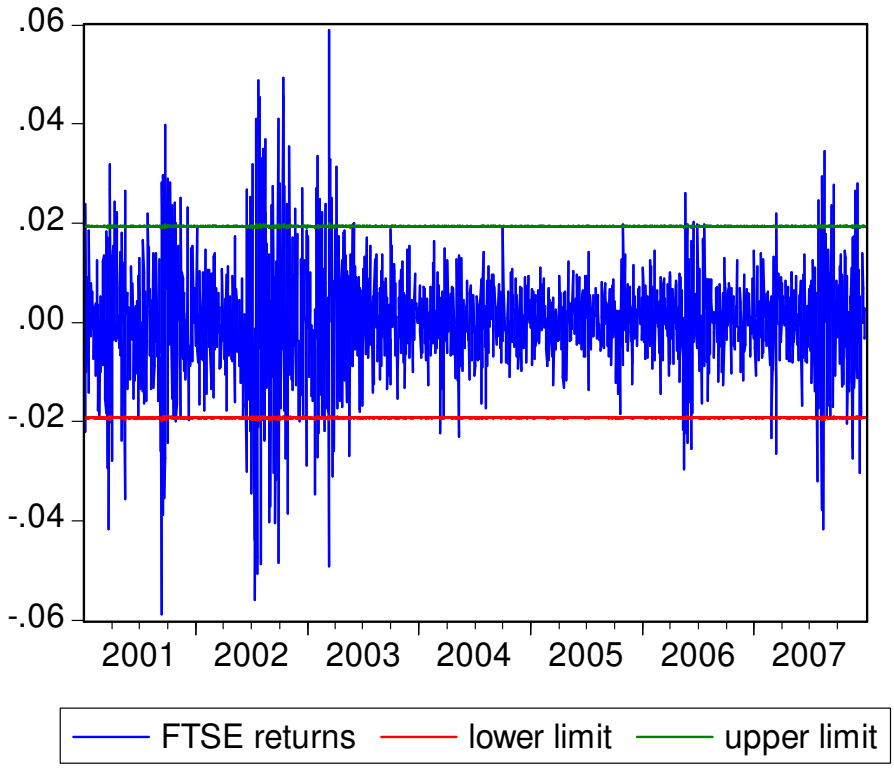


Figure 3.9. FTSE interval linear forecast.

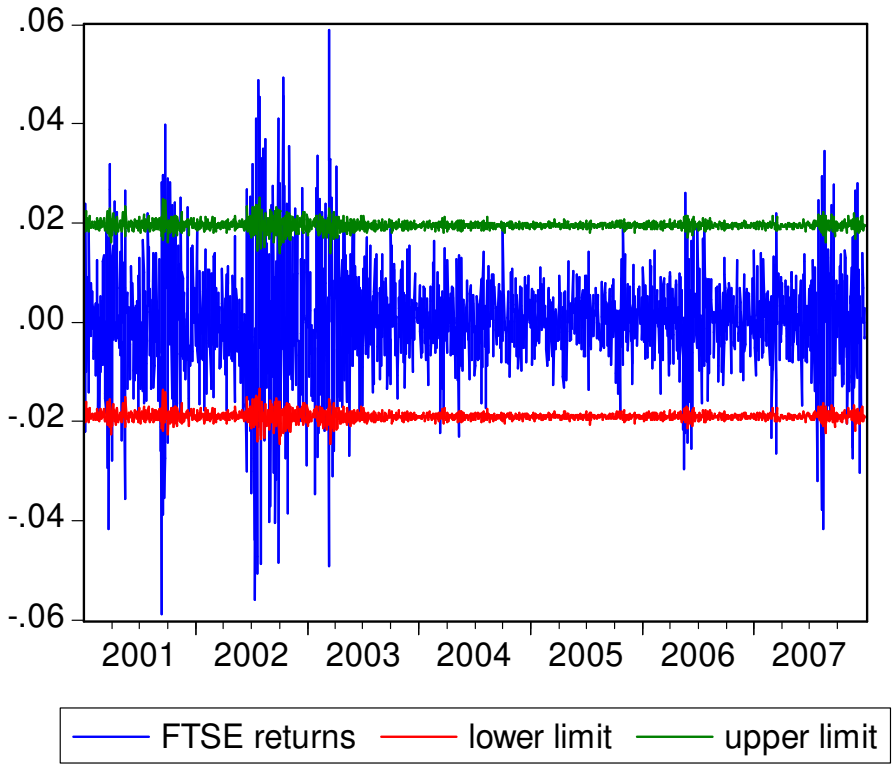


Figure 3.10. FTSE interval LSTAR forecast.

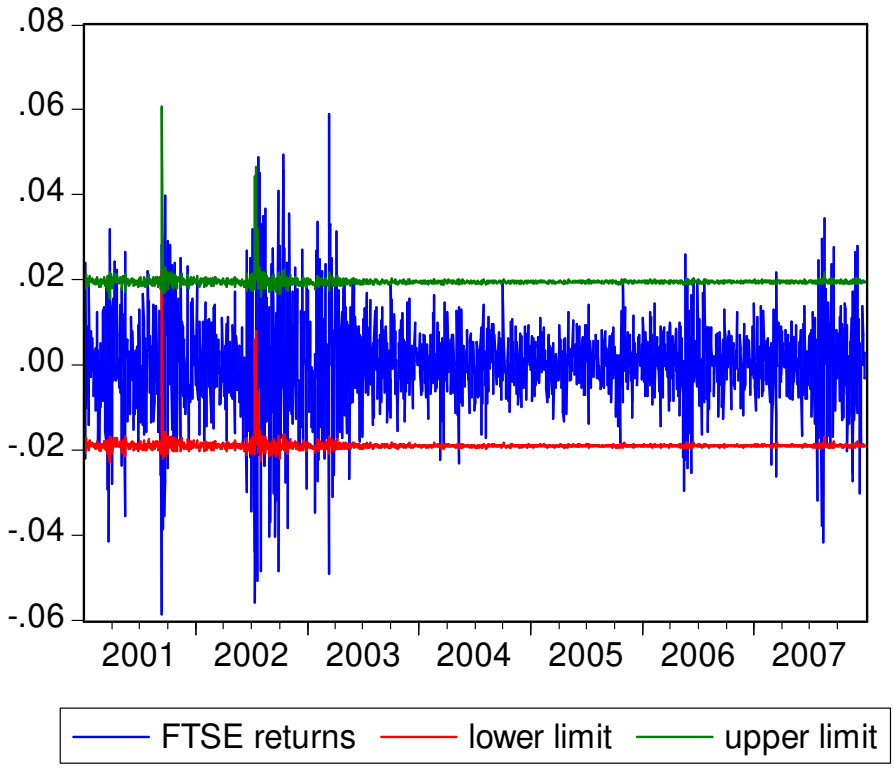
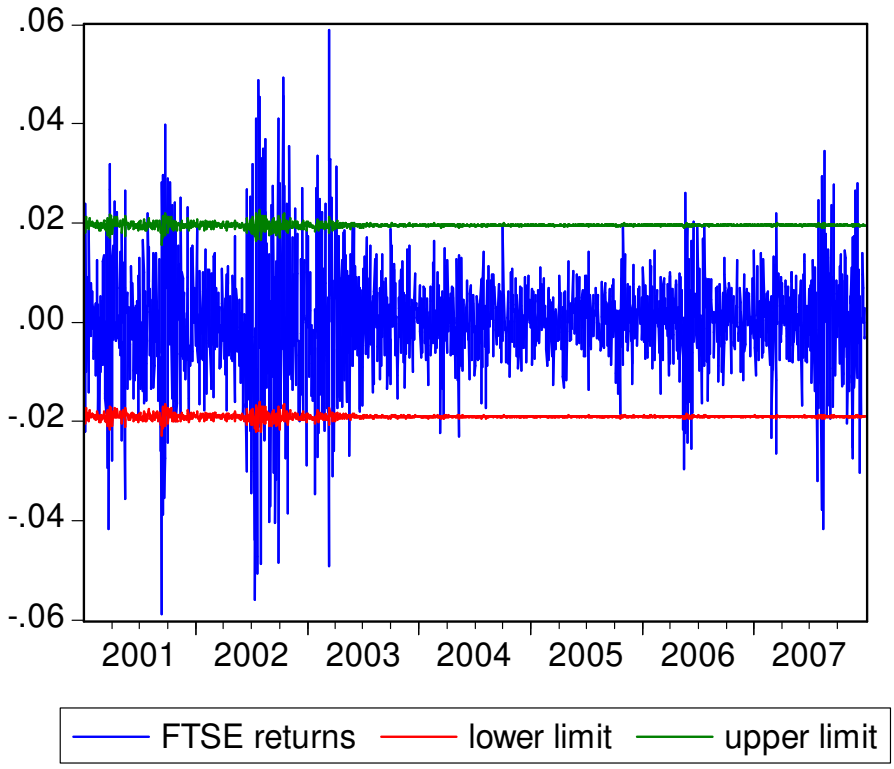


Figure 3.11. FTSE interval LSTAR trend forecast.



The success rate of the interval forecast can be easily seen on the graphical representation, where the actual returns will be either within or outside the set limit barriers, thus indicating success or failure of the forecast respectively. Naturally, the upper and lower limits for the random walk interval forecast are characterised by a somewhat less volatile line as opposed to limit barriers of the linear and non-linear interval forecasts which mimic the movements of the actual returns series. Moreover, all the interval forecasts share a characteristic of a common trend level. As expected, while the most of actual returns series values lie within the interval forecast, the outliers and extreme points rest outside the prediction barriers. The most successful forecast based on graphical representation in terms of following outliers is the LSTAR forecast (Figure 3.10), where the model attempts to correct for extreme value in the beginning of the sample characterised with high volatility.

The out-of-sample goodness of fit evaluation of interval forecast applied in this chapter is based on assessing the success ratio of the indicator variable, I_t, for a given interval forecast, $\left(L_{t|t-1}(p), U_{t|t-1}(p)\right)$ for time *t*, made at time *t-1*, with the coverage probability, *p*, for a time-series of a random variable, y_t, which is defined as follows:

$$I_t = \begin{cases} 1, & if \;\; y_t \in \left[L_{t|t-1}(p), U_{t|t-1}(p)\right] \\ 0, & if \;\; y_t \notin \left[L_{t|t-1}(p), U_{t|t-1}(p)\right] \end{cases} \tag{3.19}$$

Where, $L_{t|t-1}(p)$ and $U_{t|t-1}(p)$ are lower and upper limits respectively. In other words, zero value is assigned to every forecasted value outside the prediction barriers, while forecasts within the range are assigned a value of unity. The mean of the indicator

variable is the success ratio of the interval forecast. The results (Table 3.18) suggest that none of the interval forecasts performed in this section surpassed the limit required by the 95% confidence interval.

Table 3.18. Interval forecast success ratio results.

	Success Ratio
FTSE	
Random walk	0.9243
ARMA (1, 3)	0.9183
LSTAR	0.9178
LSTAR-trend	0.9210
S&P	
Random walk	0.9287
ARMA (2, 1)	0.9276
ESTAR	0.9265
DAX	
Random walk	0.9216
ARMA (3, 2)	0.9205
ESTAR	0.9227
Nikkei	
Random walk	0.9468
ARMA (0, 3)	0.9484
ESTAR	0.9490
ESTAR-trend	0.9473

These results could be explained by the fact that the goodness of fit evaluation procedure was based on the assumption of normal *t* distribution. Generally the distribution of financial daily data is characterised with fat tails due to daily data being very noisy and containing extreme values. Similarly to the results of the point forecast

in this chapter it seems that the argument of the daily data lacking defined patterns still holds when applying interval forecasting techniques. Therefore, the suggestion that the less noisy monthly data will demonstrate more clearly defined forecasting performance of non-linear as well as linear models is also applicable to interval forecast.

3.4. Conclusion

This chapter intended to assess the forecasting abilities of non-linear STAR-type models using daily stock price data over the period of twenty years between 1988 and 2007 using four price indices of four major world economies, including FTSE 100, S&P 500 Composite, DAX 30 Performance and Nikkei 225 Stock Average.

Results of the empirical investigation suggest the presence of stock returns predictability and presence of STAR-type non-linearity. These results are consistent with extensive literature on the issue of forecastability of stock returns and successful use of STAR-type models in forecasting these dynamics (Abhyankar et al., 1995; Clements and Smith, 1999; Clements and Smith, 2001; McMillan, 2001; Lekkos and Milas, 2004; McMillan, 2004; Teräsvirta et al., 2005). Moreover, in parallel with notion of traders interaction in financial markets suggested by McMillan (2001) and the presence of market frictions including transaction costs, limits to arbitrage, short selling and borrowing constraints (Martens et al., 1998; Kapetanious et al., 2003; McMillan, 2005b), it can be argued that small changes in pricing equilibria can be foregone and not

corrected immediately, thus, displaying non-linear dynamics within the series. STAR models produce reasonably accurate results in comparison with linear alternatives, however, any additional gains achieved by non-linear framework are only marginal to the results of a random walk and ARMA models. Hence, drawing from the results of the forecasting accuracy tests and taking into account specific behaviour and characteristics of daily data, and combining aspects of forecasting accuracy and ease of implementation, it can be concluded that the random walk model seems to be the most superior model for the purpose of forecasting daily stock returns. It has to be noted, however, that there is no clear evidence of exceeding superiority of the random walk model compared to other linear and non-linear approaches. Nevertheless, it is assumed that for forecasting high-frequency data on a daily level it is vital that the model is fast and easy to apply in addition to clear interpretation of results, which the random walk models appears to provide. The conclusion of these results is similar to those of an empirical study of high-frequency stock returns by Abhyankar et al. (1995), where the researchers confirmed the presence of non-linearity, however found the time-series to be adequately explained by a simple alternative, namely the GARCH (generalised autoregressive conditional heteroskedasticity) model process. Thus, while Abhyankar et al. (1995) encourage the use of high-frequency data due to the fact that it allows for a larger sample and thus increases the likelihood of better understanding the underlying process, there is also a possibility that small changes in high-frequency time-series returns might be too noisy and would not fully reflect the long-run dynamics.

Moreover, Fair and Shiller (1990) pointed out the fact that a specific model displays either good or poor forecasting abilities for one sample period might not necessarily mean it will have the same results for a different forecasting period. One of the reasons

for this could be a change in economic structure or other events that will change the behavioural dynamics of the data. Furthermore, Montgomery et al. (1998) found that in their study of US unemployment rate with the aid of non-linear forecasting models, the quarterly series is much smoother comparing to a more frequent monthly series. Both series shared similar cyclical and trend characteristics, however it is evident that there is a strong possibility that the long-horizon data might utilise the benefits of the non-linear forecasting much more efficiently than data sets with much higher frequency. Hence, this chapter will be concluded on the notion that the results obtained here suggest the use of the random walk model as the best forecasting model for daily stock returns in terms of the ease of implementation and relative forecasting accuracy it provides. However, it is not to suggest that the non-linearity should be disregarded and that researchers should consider its presence. This study further anticipates that an investigation of non-linear forecasting models should be extended to long-horizon data, as a non-linear approach seems to be more appropriate in this case.

Chapter 4
Long-horizon forecasting

4.1. Introduction

This chapter will concentrate on the topic of long-horizon stock predictability, in particular, the possibility of predicting stock market returns using price-dividend and price-earnings ratios. Based on the present value model introduced by Campbell and Shiller (1987), there has been a debate amongst researchers whether it is possible to use the current dividend-price ratio, or dividend yield, as a reliable enough measure of the expected stock returns in order to predict future stock returns. Literature on the out-of-sample forecasting ability of the dividend yield for stock returns is somewhat limited, with previous studies concentrating on stock returns in-sample predictability in order to examine the validity of the present value model. Thus, this chapter will apply non-linear STAR-type modelling to a present value framework with the intention to extend research into out-of-sample stock returns predictability by examining whether a forecasting exercise can be improved with a non-linear error-correction approach.

The introduction of the present value model raised research interest with numerous studies extending Campbell and Shiller's (1987) original work, including an introduction of a time-varying discount rate. Furthermore, unexpected significant rise in stock prices and subsequent fall in late 1990s and early 2000s as a result of the dot-com bubble have raised new interest in the present value model and forced academics to

focus on re-examining its validity and relationships between stock prices and dividends. In addition, an extensive research into returns stock predictability is to some extent fairly confusing with numerous studies offering various testing procedures and eventually different conclusions (Campbell and Shiller, 1987; Goetzmann and Jorion, 1993; Torous et al., 2004; Campbell and Yogo, 2006; Cochrane, 2008).

A number of studies that consider the dynamics in the dividend-price relationship assess the validity of the present value model in terms of testing for the presence of linear cointegration between stock returns and determinants. However, this approach assumes a constant discount rate, whereas non-linear modelling allows for a time-varying discount rate. The results are somewhat mixed with more recent studies suggesting non-linear dynamics in the relationship between stock market returns and dividend yield (McMillan, 2004; Kanas, 2005; Rapach et al., 2005; Bali et al., 2008). Researchers proposing a non-linear approach to the validity of the present value model suggest that market fundamentals still support stock return predictability with a long-term equilibrium but with non-linear adjustments. These non-linearities in the stock market returns-dividends relationship are suggested to be explained by the presence of transaction costs in a trading market and an interaction between informed and uninformed or noise traders. Moreover, McMillan (2001) pointed out that these non-linear adjustments to the fundamental equilibrium characterised by the presence of transaction costs and traders' interaction are persistent and exhibit slow mean reversion, which in turn implies the presence of market inefficiencies. In addition, Kanas (2005) suggests that when applying a non-linear approach to the dividend-price relationship empirical tests and methods of assessment of such modelling should be tailored

specifically for non-linear purposes as the results obtained with conventional linear techniques may be spurious.

The present value model is supported by strong theoretical analysis, however, the model has been challenging to validate using real life stock market returns. Some researchers attribute this to the presence of transaction costs in financial markets thus creating non-linear dynamics in the stock prices time series. Due to arbitrage opportunities arising from large deviations from long-run equilibrium it will be ensured that these deviations will be corrected, however small deviations that are below the transaction costs trading barrier will remain uncorrected. Hence, the different speed of adjustment depending on the size of price deviations is better to be described by a non-linear model, and in particular STAR-type models.

This empirical chapter will examine price returns of four stock market indices including FTSE All Share, S&P, DAX and Nikkei. However, developing on the results and conclusions of the previous chapter (Chapter 3), the emphasis is on long-horizon data, namely monthly stock returns. Furthermore, developing the investigation of long-horizon time-series data, the methodology will be applied to the monthly data in periods of three, six and twelve months in a form of a buy-and-hold strategy. The stock price returns are modelled using error-correction methodology with the dividend yield and price-earnings ratio as determinant variables. Recursive out-of-sample forecasting is then applied to all time periods of the four time-series in non-linear as well as linear framework, followed by the assessment using tests of forecasting accuracy.

This chapter will discuss the progression of the theory of stock returns predictability in a form of a literature review, highlighting various explanations of dynamics between stock prices and dividends, as well as empirical attempts to validate the present value

model (Section 4.2). Section 4.3 outlines the methodology applied in this chapter. Section 4.4 on monthly returns empirical results discusses statistical characteristics of the data including non-linear unit root tests, completing with the forecasting exercise and implementation of forecasting accuracy tests. Long-horizon buy-and-hold strategy is applied in Section 4.5. Section 4.6 concludes the chapter.

4.2. Literature review

Introduction to the present value model

The efficient market hypothesis (EMH) implies a relationship between stock market prices and dividends in terms that according to the theory, current prices reflect all available information, including dividends. Consistent with the concept of the efficient markets, the long-horizon equity stock returns were believed to be unforecastable. However, Campbell and Shiller (1987) introduced the present value model which relates the stock price to discounted future dividends and, thus, represents the fundamental values for the stock prices. Numerous research studies have been carried out into the validity of the present value model which in turn will support the possible predictability of stock returns. The present value model is in essence a simple stochastic model. Campbell and Shiller (1988b) point out that the dividend-price ratio can be interpreted as expectations about future dividends, or in other words, as a reflection of expectations for future dividends in current stock prices. The dividend-price ratio will

be high when, in the former case, dividends are predicted to decrease, or, in the latter case, when discount rates are high. Campbell and Shiller (1988b) attempted to investigate whether these interpretations could explain time variation in price-dividend ratio assuming rational market expectations. They also put large emphasis on the importance of log dividends and discount rates in reflecting the state of economy.

Campbell and Shiller (1988b) proposed that the present value model suggests the variations in expected stock returns to be captured by the dividend-price ratio. Assuming constant dividend growth, the present value model can be used to price stocks, hence, dividend yields, by definition, have been used to evaluate expected future returns. Subsequently Campbell and Shiller (1988b) extended their previous research to include time-varying discount rate in the dividend-price ratio, as opposed to constant discount rate considered previously, and proposed a model of a linear approximation of a relationship between stock prices, stock returns and dividends which allowed for the discount rate to vary over time. Campbell and Shiller's (1987, 1988a) research was closely related to cointegration and error-correction concepts introduced by Engle and Granger (1987), on the basis of which Campbell and Shiller (1987, 1988a) proposed a test to confirm the validity of the present value model on a condition of stationarity of the variables in the first difference. The tests involved examining bonds and stocks in the context of the present value model using a single-equation regression based on the cointegration procedure by Engle and Granger (1987). Thus, the validity of the present value model can be tested with an error-correction model, which relates the changes in a time-series variable to the changes in the variable's own lags multiplied by a cointegration vector. With the assumption of a constant discount rate the stock prices and dividend levels are theoretically cointegrated, or in other words, follow an

integrated process of order of one, i.e. *I(1)* (Campbell and Shiller, 1987). Similarly, if the present value model is valid, assuming a time-varying discount rate instead of a constant one, the log difference between dividends and prices follows a stationary process (Campbell and Shiller, 1988a, 1988b).

While error-correction models are usually used to model adjustment of cointegrated variables, Campbell and Shiller (1988a) suggested its use when one variable forecasts another and thus applied both concepts to stock price dividend relationship. Campbell and Shiller (1988a) pointed out that even though market participants, such as managers, responsible for setting dividend levels, they do not directly influence dividends, but do behave in the manner of a structural error-correction model. Campbell and Shiller (1988a) called this phenomenon a reduced-form error-correction behaviour.

Part of the debate provoked by the present value model was the fact that the presence of return stock predictability can be interpreted as evidence of market inefficiency. The alternative interpretation would be evidence of time-variation in expected returns (Torous et al., 2004). Consequently, McMillan (2001) points out that consistent with Campbell and Shiller's (1988a) extended version of the present value model it has been assumed that the linear stock predictability occurs from time-varying returns. Indeed, numerous studies have supported the presence of stock predictability when accounting for time-varying discount rate. However, despite structural simplicity, the present value model raised much controversy through empirical evidence of its validity being mixed with some researchers providing supporting evidence (Fama and French, 1988; Campbell and Shiller, 2001; Lewellen, 2004; Torous, Valkanov and Yan, 2004; Campbell and Yogo, 2006) while others do not find return stock predictability (Wolf,

2000; Lanne, 2002; Ang and Bekaert, 2007, Valkanov, 2003, Campbell and Yogo, 2006).

Stock predictability

The earlier studies that found evidence of market returns to be predictable were criticised for presence of biases, which further research tried to correct for. Stambaugh (1999) reports a bias in the OLS slope coefficients in a standard predictive regression when investigating dividend yield as a stochastic regressor for stock returns. Also, Campbell and Yogo (2006) drew attention to the fact that predictor variables such as the dividend-price and price-earnings ratios are highly persistent and might contain a unit root thus leading to over-rejection of the null hypothesis of no predictability when employing standard conventional statistics. After modifying testing procedure to account for this fact, researchers found evidence of presence of predictability in US stock returns. Campbell and Yogo (2006) based their approach on methodology used by Lewellen (2004) who found strong evidence of returns stock predictability. Lewellen (2004) provides evidence of stock predictability using dividend yield in the post-war period of 1946-2000. Similarly, after finding reliable evidence of predictability for returns over horizons less than one year, Torous et al. (2004) suggest that previous studies have not accounted for persistent behaviour of the explanatory variables and thus suffered from over-rejection of the null hypothesis of no predictability when using standard statistics. In addition, the type of regression used by the present value model is known as a predictive regression, which assumes stationarity of the explanatory variables. However, following most recent research into financial market variables, it is

evident that such an assumption seems unrealistic for most explanatory variables used in predictive regressions (Torous et al., 2004).

Wolf (2000) suggested that some studies on return stock predictability suffer from statistical pitfalls and structure dependency in model building. Hence, Wolf (2000) employed a new statistical method of subsampling, which allows avoiding the need to fit a structural model to fit the data. As Wolf (2000) pointed out, while generalised method of moments (GMM) and vector autoregression (VAR) are common methods used in approaching stock predictability, a bootstrap approach, on the other hand, does not rely on model estimation hence avoiding any possible model misspecifications. However, assumptions drawn by the bootstrap method in the context of stock predictability seem to lack asymptotic consistency. As a solution Wolf (2000) proposed the use of a subsampling approach which is completely model free and more asymptotically consistent comparing to bootstrap methodology. Wolf (2000) found the subsampling method to be superior to bootstrap, VAR and GMM methods, however found evidence of stock return predictability only for the long horizon data. Moreover, due to the fact that there was no support for predictability for short- and medium-horizon data, and presence of strong dependency in long-horizon residuals, the study concluded that there was no convincing evidence of stock return predictability when using dividend yields. Similarly Lanne (2002) also pointed out that there are clearly problems in testing stock return predictability employing simple regression model framework as it requires the use of standard *t*- and *F*-tests leading to spurious results. Lanne (2002) argues that most previous studies on stock returns predictability using strongly autocorrelated variables have ignored near unit root problem and hence their findings will be spurious. As a solution, Lanne (2002) developed a substitute to a

standard *t*-test, but found no predictability in US stock return data between 1928 and 1996.

Goetzmann and Jorion (1993) re-examined long-horizon stock returns predictability when using dividend yields. Researchers criticized previous studies for applying bias methods and, similarly to Wolf (2000), proposed the use of a non-parametric technique known as the bootstrap approach which implements the observed distribution of the data in order to model the distribution of a test statistic. Goetzmann and Jorion (1993) agreed that the bootstrap approach has certain limitations, mainly poor standard error distribution approximation for small samples, which can lead to underestimation of confidence intervals. However, in the case of large samples, on the other hand, the bootstrap methodology allows for the control of potentially bias factors such as using overlapping return intervals, the lagged correlation between independent and dependent regression variables, and their idiosyncrasies in the returns distribution or in the error structure. The outcome of the study, however, produced misleading regression coefficients, *t*-statistic and R^2, confirming the null hypothesis of no returns predictability.

Valkanov (2003) pointed out that using standard statistics in long-horizon regressions leads to spurious results due to non-standard asymptotic properties of the *t*-statistic of the least square estimator of the slope coefficient and the R^2. While Valkanov (2003) found weak predictability of returns using dividend yield for the data of the pre-war period, the post-war period was characterised by evidence of predictability which, disappointingly, was described as somewhat unimpressive. Moreover, Ang and Bekaert (2007) find stock returns predictable over short-horizons, however results also demonstrated that predictability by the dividend yield was not statistically significant. In

addition, Ang and Bekaert (2007) did not find excess return predictability when using the earnings yield as a determining variable. As suggested by the researchers, weak evidence of returns predictability could be a consequence of using univariate linear models which lack the ability to capture complex dynamics of stock returns. Hence, it was suggested that a predictability model incorporating structural break and regime shifts might produce different results.

Moreover, Cochrane (2008) suggested an entirely different approach for testing the presence of predictability in stock returns by examining whether it is possible that returns are *not* predictable. Cochrane (2008) argues that the logic behind the present value model suggests that if returns are not forecastable for cointegration relationship to hold, either the dividend growth has to be forecastable or for the dividend-price ratio to be constant. However, while observation of the dividend-price ratio variation in the financial market confirms that the later scenario is not plausible, Cochrane (2008) does not find any evidence to support dividend growth predictability to confirm the former argument. Dividend growth and expected returns are the most promising determining variables for stock prices. Hence, according to Cochrane (2008), the observed variation of dividend-price ratio and the absence of dividend growth predictability are strong evidence of predictability of stock returns.

On the contrary, Chen (2009) felt that while there has been a number of extensive research studies carried out into aggregate stock returns predictability using a dividend yield, there is still limited research done into dividend growth predictability. Chen (2009) suggests that since, by definition, the dividend yield is a sum of future expected dividend growth, it is implied that dividend yield variations will be reflected in similar variations in expected returns and the expected dividend growth. Since movements of

these variables are of a great importance and have major economic implications, Chen (2009) carries out an empirical study into stock return and dividend growth predictability. Similar to Lewellen (2004), Chen (2009) finds structural differences in different historical time periods, so that the null of no returns predictability cannot be rejected for the period before 1926, whereas it is strongly rejected for the post-war period after 1945. Results for the period between 1926 – 1945 generate mixed outputs. Hence, Chen (2009) suggests that evidence of returns predictability is mainly a post-war phenomenon. Possible explanations of such dramatic changes in returns predictability in the post-war period include increased number of firms, and thus the market index containing greater diversity of firms, implementation of different dividend policies by different firms, and a general decrease of dividend volatility. However, Chen (2009) struggles to find sufficient evidence in support of any of the above explanations of the reversal of predictability. Consequently, the topic raises many questions, and Chen (2009) suggests further investigation into the issue.

Non-linear tests of stock predictability

Conversely, growing evidence of the presence of non-linear dynamics in financial time-series, together with the failure of the linear present value model to explain stock prices dynamics suggests a non-linear approach to the price-dividend relationship. Furthermore, an increasing discrepancy between stock and fundamental prices evident in the late 1990s casts a final doubt that the stock prices follow linear stationary perfectly cointegrated behaviour implied by the present value model. Campbell and

Shiller (2001) report an unusually bearish behaviour within the US stock market in 1998 which resulted in a shift in stock prices from the fundamental values and historical averages. However, Lewellen (2004) reports finding strong evidence of stock returns predictability even during the period of the unusual price dynamics. Bohl and Siklos (2004) propose a more plausible approach by taking an assumption that the present value model is valid as a long-run framework for the US stock prices, and recognising the presence of asymmetries in the short-run. As pointed by Bohl and Siklos (2004), there are a number of possible reasons for mixed empirical evidence of the long-term validity of the present value model including the presence of non-linearities, structural breaks and outliers. It is indeed possible to integrate crashes and non-fundamental stock price behaviour that occurred during the 1990s by not including the transversality condition of the standard present value model. Numerous studies were carried out in order to explain stock price behaviour as a function of dividends. An increasing number of researchers conclude that the prices and dividends are in fact cointegrated, however the mean reversion processes is characterised in a non-linear fashion. McMillan (2007), for instance, observed that some researchers argued that the deviation from the fundamentals in the 1990s was a result of an extended bubble that eventually burst, and were concentrating on determining a technique which would allow to capture this type of stock price behaviour. Non-linearities in the present value model are usually explained by the presence of non-fundamental components. In addition, Psaradakis et al. (2004) identified the presence of a time-varying discount factor and the presence of bubbles as possible explanations for short-term deviations of prices from the fundamental values and long-term price-dividend relationship. The most promising theoretical justifications of such dynamics include the presence of speculative bubbles

(Blanchard and Watson, 1982; West, 1988; Evans, 1991); noise traders' models (Kirman, 1991, 1993; Shleifer, 2000) and the theory on booms and slumps in economic activity (Phelps, 1994; Phelps and Zoega, 2001).

Besides the debate on the presence of stock returns predictability, there is an ongoing discussion concerning the predictability in short- and long-horizons. Rapach et al. (2005) investigated the presence of long-horizon predictability in real stock prices using a predictive regression model with price-dividend and price-earnings ratios as fundamental valuation ratios following previous research that seemed to detect predictability in the long-horizons, but not in the short-horizons (e.g. Campbell and Shiller, 1998). Possible explanation of such pattern of stock price predictability could be attributed to presence of non-linearity. The argument was based on the work by Berkowitz and Giorgianni (1996) who addressed long-horizon predictability of nominal exchange rates using monetary fundamentals as valuation ratios. The researchers argued that a linear framework does not provide sufficient justification for the stock predictability as it implies that the stock predictability is for *all* horizons or for *no* horizons. This contradicts with numerous findings of long-horizon predictability and the lack of such in the short-horizons. Using their approach Rapach et al. (2005) adopted the methodology in order to implement the Monte Carlo simulations to the long-horizon stock price predictability. While results from a linear predictive regression demonstrated the ability of both price-dividend and price-earnings ratios to predict stock price in the long- but not short-horizons, the parsimonious exponential smooth transition autoregressive (ESTAR) model proved not only to fit the data sufficiently well, but also to allow for non-linear mean reversion, thus providing plausible explanation for the long-horizon predictability pattern. As a result, Rapach et al. (2005) agree that a non-

linear framework provides a sufficient explanation for the pattern of stock price predictability for at least the dividend-price ratio. Moreover, Kilian (1999) argues that the observed pattern of long-horizon predictability together with the absence of predictability in the short-horizon can be interpreted as indirect evidence of the presence of non-linearity in the data generating process.

Kanas (2005) employed a non-linear cointegration approach to confirm the presence of non-linearities in the stock price and dividend relationship and thus validated the present value model in non-linear fashion. Bali et al. (2008) also found evidence of stock returns predictability by employing a non-linear test of mean reversion. Hartmann et al. (2008) find evidence of predictability of stock returns using macroeconomic variables and incorporating structural breaks by assessing publicly available and easily accessible information on economic and financial crises.

Reasons for non-linear behaviour

The idea behind non-linear dynamics within the stock market time-series caused by the interaction between informed and uninformed traders is that the deviation from the fundamental values has to be sufficiently large for the arbitrage traders to participate in the market trading, thus correcting the values to the long-term equilibrium. Similarly, the presence of transaction costs will ensure that the arbitragers will only engage in trading if the return exceeds the required limit. Hence, the stock returns time-series will be characterised by bounds of inactivity around the equilibrium which, in turn, causes small and large returns to exhibit different dynamics (McMillan, 2001).

However, according to the theories of behavioural finance, investors will exhibit different assets trading behaviour following different states of the market. Thus, great importance is placed on the phenomenon of the market sentiment where noise traders will demonstrate a positive response in rising markets by overreacting to positive news and hence overvaluing stocks bringing the prices up in the excess of what is required by the news, otherwise known as trend-chasing (McMillan, 2005; McMillan and Speight, 2006). The arrival of bad news will, on the other hand, result in noise traders showing bearish, more conservative behaviour that is close to the characteristics of fundamental traders. As a result, the fundamental traders are trying to take advantage of these known noise traders' strategies by recognising the market triggers that set off the noise traders and engaging in the trade by taking long positions in order to drive the asset prices even further before short selling. In other words, fundamental traders purposefully de-stabilise the market in the process of a profit gain. In the light of these practises, McMillan and Speight (2006) suggest that the predictability implied by the present value model will be weaker in the rising market due to the market sentiment, while observing a stronger connection of stock returns to the fundamentals in the falling markets. Thus, when the prices are undervalued market forces readjust these back to the equilibrium more quickly. Whereas adjustments to the fundamental values of overvalued prices will happen at a slower rate due to trend-chasing and interaction between noise and arbitrage traders. Furthermore, this demonstrates the asymmetry in returns predictability following positive and negative dividend yield values. In other words, prices being below or above the fundamental values, where the predictability

will be stronger in the periods when the prices are close to the fundamentals while the overvalued market will be characterised with a weaker predictability.

In addition, McMillan (2004) suggests that since the arbitrage traders will not engage in trading unless the deviation from the fundamental values exceeded arbitragers' trading barrier, the behaviour of the stock returns will differ depending on the size of the disequilibrium. Thus small deviations may be foregone uncorrected due to the presence of transaction costs and fundamental traders not engaging in the trade. Deviations exceeding transaction costs, but still considered to be of small return deviation, will be corrected at a faster rate as a result of an increased number of market arbitrage participants due to possible profit opportunities. Fundamental traders will only engage in trading activities when returns disequilibrium is sufficient enough to produce a profit (McMillan, 2005). However, large return deviations will then be characterised by a slow mean reversion as noise traders engage in trend-chasing and the arbitrage trades are reluctant to act upon the mispricing due to greater risk from adverse market movements. Consequently, difference in sizes of price deviations will result in different rates of adjustments to the equilibrium, hence suggesting non-linear dynamics in the process of reversion to the equilibrium. Furthermore along the lines of the behavioural finance argument, Bali et al. (2008) points out a time-varying or state-dependent nature of the investors' relative risk aversion, such that it increases in falling markets due to short sale, liquidity, or financing constraints. Bali et al. (2008) provides evidence of significantly increased speed of mean reversion during falling markets as a result of an increased investors' risk aversion.

McMillan (2004) demonstrates the preference of a non-linear error-correction model over a linear alternative, as the regimes presented by the ESTAR model capture

different dynamics resulting from small and large price divergences as well as allowing for the smooth transition between these regimes. McMillan (2003) also supports the view that the interaction of noise traders and informed arbitrage traders is one of the reasons for different dynamics in the stock prices associated with small and large returns, thus partially being the reason for linear predictability to be rejected for stock market returns in the light of the present value model.

Market bubbles

Similarly, numerous studies have concentrated on investigating periodically collapsing and speculative bubbles in the context of the stock returns predictability. The presence of bubbles would explain the deviation of stock prices from fundamentals since the absence of bubbles would be indicated by the stock prices and dividends to be cointegrated in a linear fashion as suggested by the present value model. Evans (1991) found no evidence of periodically collapsing bubbles. However, Bohl (2003) suggests that the results obtained by Evans (1991) were based on using inappropriate testing techniques that are not suitable for non-linear processes in cointegration systems. One of the practical difficulties when faced with a periodically collapsing bubble is that the bubble component follows a non-linear process and thus naturally falls outside the alternative hypothesis of the standard unit root test (Bohl, 2003). Hence, Bohl (2003) applied the momentum threshold autoregressive (MTAR) model in order to capture distinctive asymmetric non-linear long-run relationships between real stock prices and dividends and thus examine the presence of periodically collapsing bubbles in stock prices by implementing a non-linear cointegration framework. As a result, Bohl (2003) finds no evidence of the presence of periodically collapsing bubbles in the US stock

market during the sample of 1871 – 1995. However, after extending the sample period to 1871 – 2001 and thus including the period of rapid increases in share prices in the 1990s, Bohl (2003) is able to report the presence of periodically collapsing bubbles, hence implying that the phenomena are period specific. In addition, besides theoretical doubts and obvious difficulties of identifying and modelling the bubbles in the stock market prices, Bohl (2003) reminds that it is practically impossible to prove the existence of bubbles.

Bohl (2003) used a momentum threshold autoregressive (MTAR) model in order to utilise a cointegration framework with asymmetric adjustment while investigating the presence of periodically collapsing bubbles in the stock market. The MTAR model was developed by Enders and Siklos (2001) in order to empirically capture and investigate periodically collapsing bubbles within a cointegration framework. The evidence of asymmetry in deviations from the equilibrium would indicate the presence of periodically collapsing bubbles. Results obtained from the study by Bohl (2003) showed that there was no asymmetry revealed in the US stock market residuals in the subsample prior to the rapid share price increase from 1995. However, the presence of periodically collapsing bubbles in the late 1990s in the US stock prices is confirmed by the MTAR model once the sample is extended to 1871 – 2001. Further on, Bohl and Siklos (2004) also consider the MTAR model as standard linear cointegration methods evidently would be inappropriate in this instance. Whereas the MTAR model is able to accommodate asymmetric price adjustments to the equilibrium in short-run, or in other words, a non-linear error correction mechanism, while preserving the linear long-run relationship between stock prices and dividends. In addition, the researchers claim ease of implementation of the model, which is to be suggested for use by market

practitioners. Bohl and Siklos (2004) applied their approach to US stock market data between 1871 and 2001, and found that for most of the data the log dividend-price ratio followed a stationary process with asymmetric short-run adjustments. Bohl and Siklos (2004) argue that by allowing time-varying expected returns, which seems to be a more realistic assumption in the first place, it also results in a much simpler model. Since this methodology allows for non-linear short-run adjustments, it seems that in comparison to conventional unit root and cointegration techniques Bohl and Siklos' (2004) approach provides a better method of capturing properties of the log dividend-price ratio.

Brooks and Katsaris (2003) draw parallels between the market behaviour in the late 1990s and financial crashes known as 'Black Thursday' in 1929 and 'Black Monday' in 1987. Researchers suggest that the market behaviour observed during the two market crashes share similar development characteristics between the peak and the market collapse. While the significance of fundamental values can be questioned as a result of these occurrences, some researchers explain it with irrational investor behaviour or presence of speculative bubbles, which in turn can be created by informational asymmetry and inaccurate estimation of market fundamentals (Brooks and Katsaris, 2003). Hence, Brooks and Katsaris (2003) proposed an idea of a speculative bubble as a determinant variable for stock prices during the 1990s when the long-run relation between dividends and stock prices did not hold. However, due to a bubble component being a random variable it is extremely difficult to detect. Brooks and Katsaris (2003) employed three different techniques of identifying presence of bubbles with cointegration methodology producing most promising results. The technique involves testing dividends and price series for stationarity and long-run cointegration relationships, where the absence of such relationship could be attributed to the presence

of a speculative bubble. While Brooks and Katsaris (2003) agree that the methodology has its flaws such as not being able to identify all types of bubbles and relying on theory of market prices being determined by these fundamental values, the cointegration system testing is described as the best available tool for identifying presence of speculative bubbles as it is responsive to small samples and model misspecification. Results obtained in the study showed cointegrating relationship between dividends and prices to exist until 1993 for the FTSE All Share index suggesting presence of a speculative bubble after 1993. However, as mentioned by Johansen (1991), the absence of cointegration does not necessarily mean the presence of bubbles. Moreover, as Brooks and Katsaris (2003) pointed out, these results could also be interpreted as a result of a structural change in the long-run relationship between the prices and dividends, thus suggesting models are able to account for structural or regime changes for further investigations of the issue. In addition, the imperfection of the identification technique could also imply that the FTSE index series is characterised with speculative bubbles prior 1993, however the presence of which have not been detected by the methodology. Despite the fact that Brooks and Katsaris' (2003) investigation provides evidence that comply with the presence of bubbles in financial price series, the null hypothesis of absence of bubbles cannot be truly rejected as there is a possibility of non-observable variables, such as investors' expectations and market sentiment, that might cause the dramatic changes in the fundamental price-dividend relationship. In addition, Brooks and Katsaris (2003) suggested investor irrationality or shifts in investors' preferences as possible reasons for divergence from the fundamentals. As a result, the standard present value model in its linear form is insufficient to explain complex dynamics of price movements. Similarly, Kilian and Taylor (2003) suggest that the

market dynamics and market price movements due to the presence of noisy, or uninformed, and arbitrage, or informed, traders are better explained by the means of a non-linear mean reversion approach.

Psaradakis, Sola and Spagnolo (2004) proposed a two-state Markov error-correction model in order to accommodate different rates of adjustments to the long-run equilibrium of US stock prices and dividends and found evidence in favour of these types of models. The researchers argued that dynamic adjustment of Markov error-correction models allows to explain the evolution of stock prices in periods when long-run cointegration processes between stock prices and fundamental values seems to fail or deviations from the equilibrium are corrected at a different speed. Psaradakis et al. (2004) chose Markov-type models as these are adequate for modelling an abrupt change in regime caused by a sudden shock rather than smooth adjustment to a new regime. In addition, Markov error-correction framework was chosen on the basis that it is able to identify periods of unusually high dividend-price ratios which correspond to periods of occurrence of an intrinsic bubble which, according to Psaradakis et al. (2004), indicates the prices divergence from fundamentals. However, while the results of the investigation demonstrated the ability of the Markov switching process to successfully identify periods of disequilibrium, it is unclear whether the identified deviations are caused by an intrinsic bubble or time-varying discount factor. Moreover, research has found no consistent evidence of the presence of periodically collapsing bubbles.

Market frictions and traders interaction

Non-linear dynamics, as suggested to arise from the presence of transaction costs and the interaction between arbitrage and noise traders, require a versatile type of model

able to capture these non-linear adjustments. STAR-type models are very adaptable in terms that these allow for gradual adjustment between the regimes, which are consistent with slow mean reversion and are able to capture two types of asymmetry. Sollis et al. (2002) also argue in support of the smooth transaction between regimes as opposed to abrupt change of TAR models in the context of exchange rates. STAR models are capable of capturing market behaviour dynamics which vary when returns differ in sign, i.e. positive and negative returns. In other words, the market will behave differently depending on its state. McMillan (2001) points out that investors' psychology and, thus, their behaviour will depend on whether the market is falling or rising, and hence if investors' are bullish or bearish. Therefore, the logistic STAR (LSTAR) model can capture the direction of disequilibrium where the model parameters change depending on whether returns are above or below the threshold value, which in the case of returns predictability would be negative or positive returns. Whereas, the exponential STAR (ESTAR) model captures different dynamics when the returns are large or small, in other words, it describes the size of disequilibrium (McMillan, 2001). McMillan (2003) carried out research into non-linear stock predictability using ESTAR modelling on the example of US stock market data. The study found that the ESTAR model performed well thus confirming the view that the market participants' behaviour differs between large and small returns.

McMillan (2005) applied the quadratic-logistic smooth transition autoregressive model (QLSTAR), first suggested by Jansen and Teräsvirta (1996), to international daily market index data. This smooth regime transition model is able to capture non-linear dynamics consistently with the noise trader behaviour where the mean reversion rate differs resulting from large and small returns, while accounting for different variations

following positive and negative returns. Results confirmed the noise traders' interaction theory, whereby the speed of transition differs between rising and falling markets due to cognitive biases resulting in slow mean reversion. McMillan (2005) found the non-linear QLSTAR model to outperform the linear model in in-sample as well as out-of-sample forecast, and to provide evidence of the presence of return stock predictability and non-linearity in the price dividends relationship. In addition, McMillan (2005) found that the returns predictability occurs only in the outer regimes characterised by the large returns, while the inner regime of small returns exhibits random walk patterns with a drift.

Due to the presence of transaction costs and cognitive biases of noise traders, the equilibrium of fundamental price dividends mean reversion relationship is characterised by the presence of a barrier band that the prices deviations have to surpass before arbitrage traders engage in active trade. Moreover, according to McMillan (2005), the band displays a non-symmetric quality for positive and negative returns deviations due to short-selling constraints that restrict arbitrage trades to sell overpriced assets, and due to the noise traders tendency to become over-confident and exhibit bullish behaviour in rising markets while acting more conservative during falling markets. Thus this will result in negative deviations being corrected sooner than positive deviation, which, in its turn, will follow slow reversion to the equilibrium, thus leading to a greater mispricing during up markets (positive returns) than in down markets (negative returns). Kilian and Taylor (2003) point out that while the limits to arbitrage and other market frictions prevent fundamental traders from correcting market mispricing immediately, progressively over- or undervalued assets reduce the risk to arbitrage, thus encouraging fundamental traders to engage in the trade. Hence, as asset prices deviate further from

the fundamentals, the speed of reversion will become unevenly faster producing an asymmetry in the adjustments to the equilibrium. Faster speed of mean reversion during the falling markets is also confirmed by Bali et al. (2008). Considering that the noise traders are more likely to overreact to good news in the rising market and consequently overvalue the stock via trend chasing, while they tend to behave bearish in the falling markets, it is evident that these different reactions will result in different stock price dynamics depending of the sign of disequilibrium. While the logistic STAR (LSTAR) models are able to accommodate for sign asymmetry between rising and falling markets, McMillan (2007) suggests asymmetric ESTAR model in order to allow for the asymmetry in the sign of the deviation as well as the differences in behaviour following large and small deviations.

Conclusion

More researchers turned to investigate possible asymmetries within the mean reversion relationship between stock prices and dividends, as well as other financial time-series. Thus, in the light of validity of the stock market predictability, Bali et al. (2008) proposed a test for non-linear mean reversion. In addition, Sollis et al. (2002) carried out a study into purchasing power parity of the exchange rates where they have proposed a test for time-series mean reversion based on smooth transition autoregressive models. In the study, one of the tests used forced mean reversion to be symmetric while the other test allowed asymmetry in the adjustment. As a result, the proposed test displayed stronger evidence against the unit root hypothesis compared to the standard Dickey-Fuller test.

In addition, a vast majority of studies into the price-dividend relationship assessment consider US market data, thus creating a vulnerable point for criticism of data specific results. Hence, while a considerable majority of empirical studies into return stock predictability concentrate on the US market data, a number of researchers have attempted to extend their examinations to different economic markets in order to perceive whether the predictability phenomena is market specific, including Brooks and Katsaris (2003) adopting a bubble hypothesis to UK data; Kanas (2005) accessing non-linear dynamics of US, UK, Germany and Japan; Kapetanios et al. (2006) examining international indices; and McMillan (2007) applying asymmetric ESTAR models to thirteen countries including South East Asia markets. McMillan and Speight (2006) found evidence of long-horizon predictability for all six South East Asia markets considered. Their findings are supportive of the noise traders' behaviour theory in the sense that there is evidence of different signs of a de-meaned dividend yield having different effects on the levels of predictability of stock returns. They also found that the forecasting power increases with the horizon based on R^2 values, thus the most efficient horizon being between twelve and 48 months of monthly data for six South-East Asia financial markets. McMillan (2005) found that different limits to arbitrage and differences in fundamental traders' knowledge in different countries may be the reason for differences in forecasting advantages of various non-linear models since the Asia-Pacific economies, which are still considered to be evolving, demonstrated greater forecasting power when compared to the more established European economies. In addition, McMillan and Speight (2006) found no common pattern of non-linearity in their international data sets. Hartmann et al. (2008) observed that the structural changes as a result of economic and financial crises are more frequent in emerging market

economies compared to economically developed and mature industrialised countries, thus suggesting different implications for stock market dynamics.

The current investigation aims to find the best non-linear framework of the STAR-type models to suit monthly stock market data. For these reasons and in order to avoid any ambiguity caused by unanticipated irregularities of a developing market, this study will be applying non-linear methodology to well established developed financial markets, including the UK, US, Germany and Japan. While other studies have attempted to provide a wide sample of different economies, this study is inclined to focus on examining the forecasting properties of different non-linear STAR-type models on the cross section of four established and developed markets of the US, UK, Germany and Japan.

4.3. Methodology

The methodology for this chapter relies on methods described in Section 2.3 and includes additional discussion of the STAR-type error-correction model in the context of the present value approach.

The fundamental value for prices as suggested by the original present value model can be described in the following form:

$$P_t = \sum_{i=1}^{\infty} \delta^i E_t D_{t+i} \tag{4.1}$$

Thus, the stock price, P_t, is a function of expected dividends, D_{t+i}, and the time-varying discount rate, δ^i, where E_t is the expectations factor at the time *t*.

The standard present value model can be further re-written in terms of the dividend-price ratio. The real one-period return, r_{t+1}, which is determined by the capital gain of (P_{t+1}/P_t) and the dividend yield (D_{t+1}/P_t), can be defined as:

$$r_{t+1} \equiv ln(1 + R_{t+1}) = ln[(P_{t+1} + D_{t+1})/P_t] \tag{4.2}$$

Where P_t is the stock price at time t, D_{t+1} is the dividends paid during the period $t + 1$, and R_{t+1} is one-period holding period return. The equation 4.2 can be linearised:

$$r_{t+1} \approx \rho p_{t+1} - p_t + (1 - \rho) d_{t+1} + k \tag{4.3}$$

Where p, d, r are the logarithms of prices, dividends and the discount rate respectively, ρ is the linearisation parameter and k is the linearisation constant.

By defining the log dividend-price ratio as:

$$\delta_t = d_t - p_t \tag{4.4}$$

the linearised equation 4.3 is expressed as follows:

$$r_{t+1} = \delta_t - \rho\delta_{t+1} + \Delta d_{t+1} + k \tag{4.5}$$

The above equation (4.5) implies that the one-period returns can be forecasted by forecasting the dividend-price ratio (δ_{t+1}) and the change or growth in dividends. This equation can be solved to generate the expression for the price level of the stock, which is the original log-linear approximation allowing for a time-varying rate proposed by Campbell and Shiller (1988a; 1988b):

$$p_t = (k/1-\rho) + E_t\left[(1-\rho)\sum_{i=0}^{\infty}\rho^i d_{t+i+1} - \sum_{i+0}^{\infty}\rho^i r_{t+i+1}\right] \tag{4.6}$$

Further, imposing the transversality condition ($lim(\delta^n E_t D_{t+n}) = 0, as\ n \to \infty$), equation (4.6) can be re-written in terms of the log dividend yield or dividend-price ratio so that the ratio depends on expectations of future changes in dividends and the discount rate:

$$d_t - p_t = (k/1-\rho) + E_t \sum_{i=0}^{\infty} \rho^i (-\Delta d_{t+i+1} + r_{t+i+1}) \tag{4.7}$$

If the present value model holds, log prices and log dividends would be cointegrated with a cointegration vector (1, -1). Thus, the testing of the present value model involves testing the dividend-price ratio for stationarity and for the presence of cointegration relationship between log prices and log dividends.

For testing predictability of stock returns using the dividend yield, which is the main intention of this chapter, the above equation 4.7 can be re-written to express the relationship between stock returns and dividends as following:

$$r_t = \alpha + \beta(d_{t-1} - p_{t-1}) + \varepsilon_t \tag{4.8}$$

Where α and β are equation coefficients.

Furthermore, the STAR-type models, namely exponential STAR (ESTAR) (4.9), logistic STAR (LSTAR) (4.10) and asymmetric ESTAR (AESTAR) (4.11), will be applied in this chapter to the forecasting of price returns series using dividend-price or price-earnings ratios in the form of an error-correction system:

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 - exp(-\gamma(s_{t-d} - c)^2/\sigma^2(s_{t-d}))\big) + \varepsilon_t \tag{4.9}$$

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 + exp(-\gamma(s_{t-d} - c)/\sigma(s_{t-d}))\big)^{-1} + \varepsilon_t \quad (4.10)$$

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\left(1 + exp\big(-\gamma_1^2 s_{t-1}^2 I_t - \gamma_2^2 s_{t-1}^2 (1 - I_t)\big)\right)^{-1} + \varepsilon_t \quad (4.11)$$

where the returns, r_t, are regressed using the transition variable s_{t-d} with the threshold value c, so that if $y \to \infty$ or $\gamma \to 0$, the equation becomes linear, π and θ are autoregression coefficients, and ε_t is the error term. The AESTAR function (4.11) becomes a symmetric model when speeds of adjustments are identical ($\gamma_1^2 = \gamma_2^2$), and I_t is the indicator function for the AESTAR model which depend on whether the transition variable above or below zero:

$$I_t = 1 \; if \; s_{t-1} > 0 \quad (4.12)$$

$$I_t = 0 \; if \; s_{t-1} \leq 0$$

The dividend yield or the dividend-price ratio and price-earnings ratio are used as a currency free comparative measure of financial assets. Thus, dividend yield (Y_t), or dividend-price ratio, measures annual dividend payout (D_t) as a percentage of the assets stock price (P_t) in the following form: $Y_t = D_t/P_t$. The terms dividend yield and dividend-price ratio will be used interchangeably in this paper. For more details see

Hull (2003), and Reilly and Brown (2003). The dividend yield and price-earnings ratio will be used as a transition variable (s_{t-d}) in the error-correction framework in order to access stock returns predictability and forecasting performance of STAR models.

4.4. Empirical results for monthly returns

This chapter will analyse time-series monthly data over a thirty six year period from January 1973 to February 2009. The data consists of four price indices including FTSE All Share, S&P 500, DAX 30 Performance, and Nikkei 225 Stock Average; dividend yield series and price-earnings ratio for each index over the same time period.

Descriptive statistics

A graph below (Figure 4.1) illustrates diagrams for all four monthly price indices considered here plotted against time. The observation shows that the values for the indices are slightly less volatile in comparison to the daily data plots of the same series in Chapter 3. The data follows the same pattern characterised by a dramatic increase during the late 1990s and subsequent decline in early 2000s, with the exception of the Nikkei series which appears to react somewhat differently to common global market influences. The lack of responsiveness of the Japanese index to the global market could be explained by prolonged recession. Moreover, the effects of the financial crisis in

2007 are clearly seen in all fours series in a form of a significant decline up to the end of the sample in 2009.

Histograms of the four price series exhibit similar statistics to the daily data discussed in Chapter 3. The null hypothesis of normality was rejected for all four series on the basis of the Jarque-Bera statistic. Positive skewness and kurtosis values in the non-symmetric distribution indicate that the upper tail is thicker than the lower tail and that the tails in general are thinner than those of a normal distribution, suggesting that the main mass of the distribution is concentrated on the right of the distribution having fewer high values.

An augmented Dickey-Fuller (ADF) unit root test performed on the monthly price indexes reveals expected non-stationarity of the price time series for all four data sets (Table 4.1).

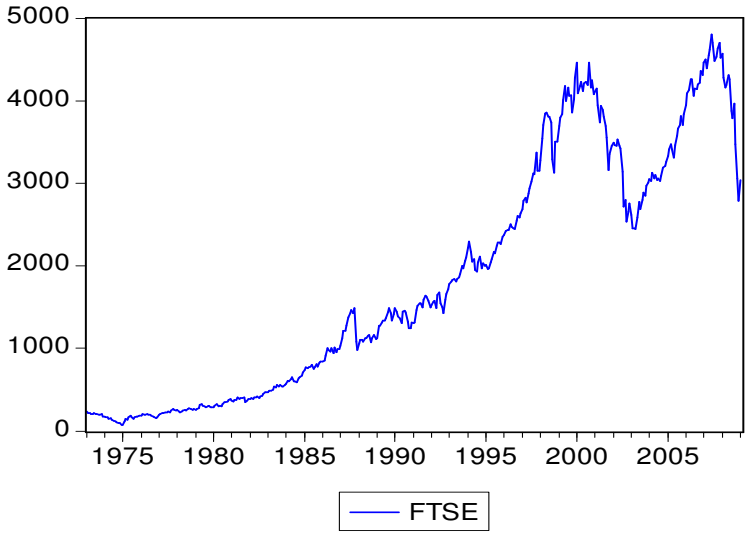
Table 4.1. Price time-series ADF test results.

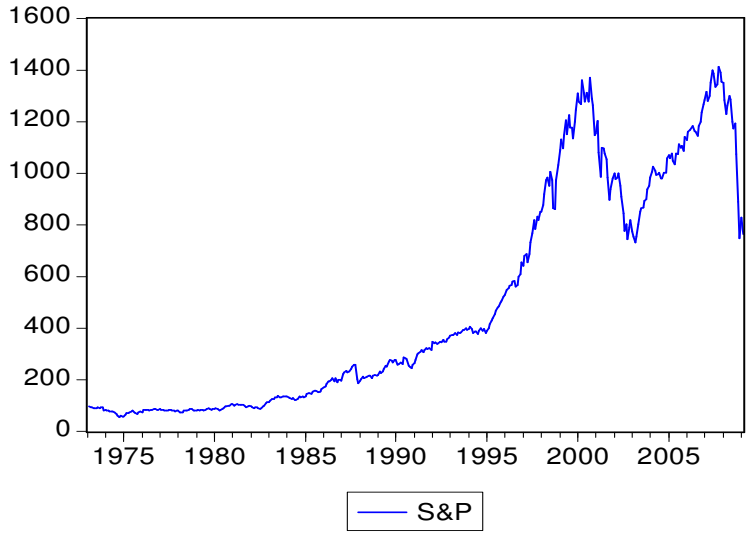
ADF critical values		**Test statistics**	
1 % critical value	-3.4617	FTSE	-1.9011
5 % critical value	-2.8748	S&P	-1.7024
10 % critical value	-2.5738	DAX	-1.6253
		Nikkei	-2.0375

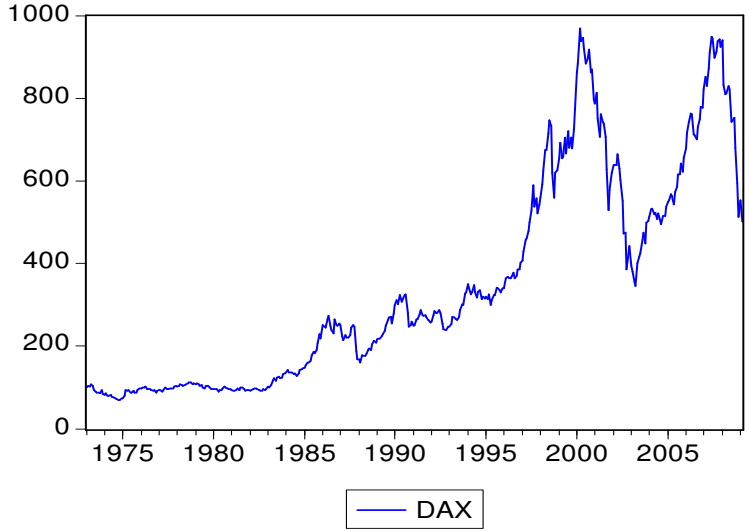
Returns are calculated as first difference logarithms for all data series (Figure 4.2 – 4.5). Returns exhibit different distribution characteristics to the prices series. The null hypothesis of normality is still rejected for all four returns series using the Jarque-Bera statistic. Kurtosis values are all slightly larger indicating excess peakedness and, hence, suggesting low number of fairly extreme deviations rather than more size distributed

moderate values. Skewness is positive only for the FTSE returns suggesting that the upper tail of distribution is thicker, whereas negative skewness for S&P, DAX and Nikkei returns implies thicker lower tail and thus a larger number of higher value returns (Figure 4.6). Figures 4.7 and 4.8 represent plots of the price-dividend ratio, or dividend yield, and price-earnings ratio.

Figure 4.1. Monthly price indices: FTSE, S&P, DAX, Nikkei.







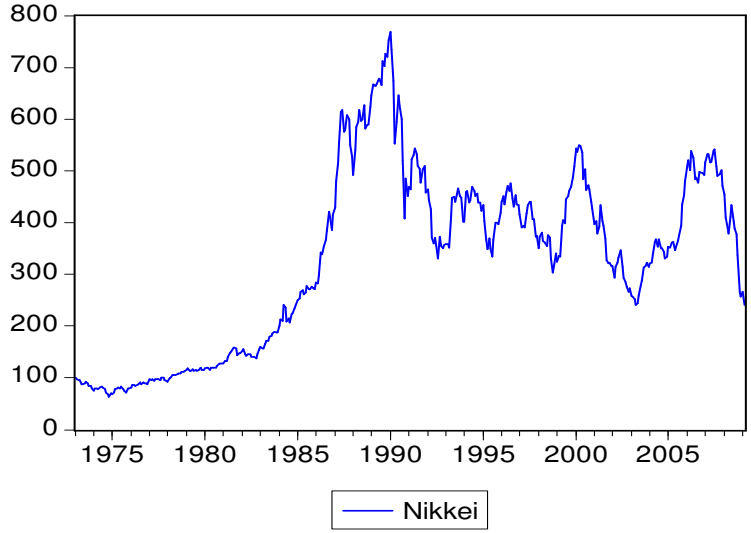


Figure 4.2. Price returns, FTSE.

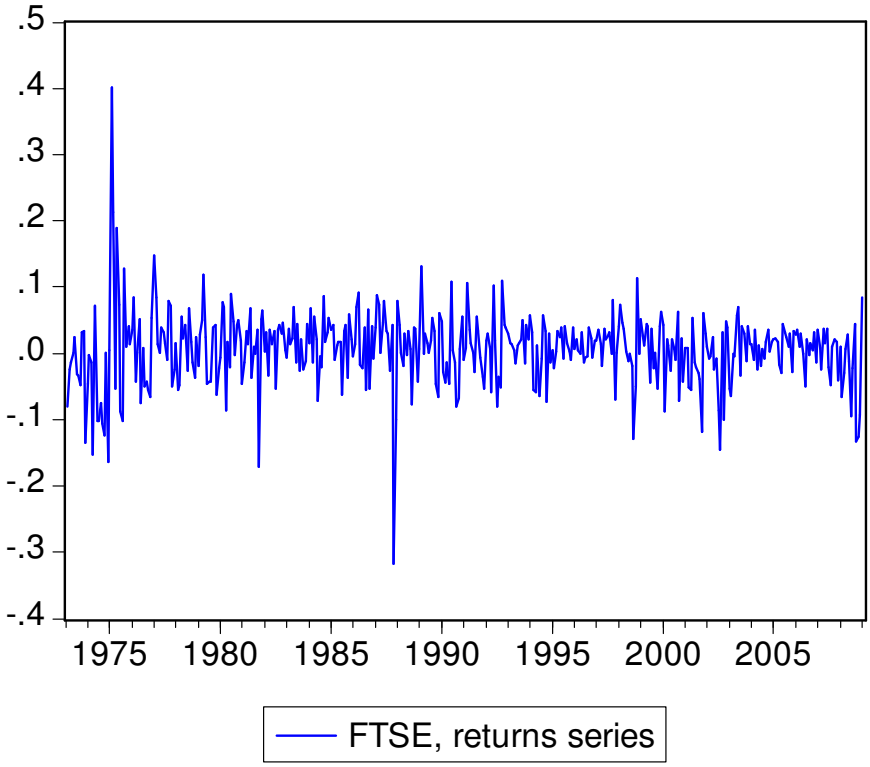


Figure 4.3. Price returns, S&P.

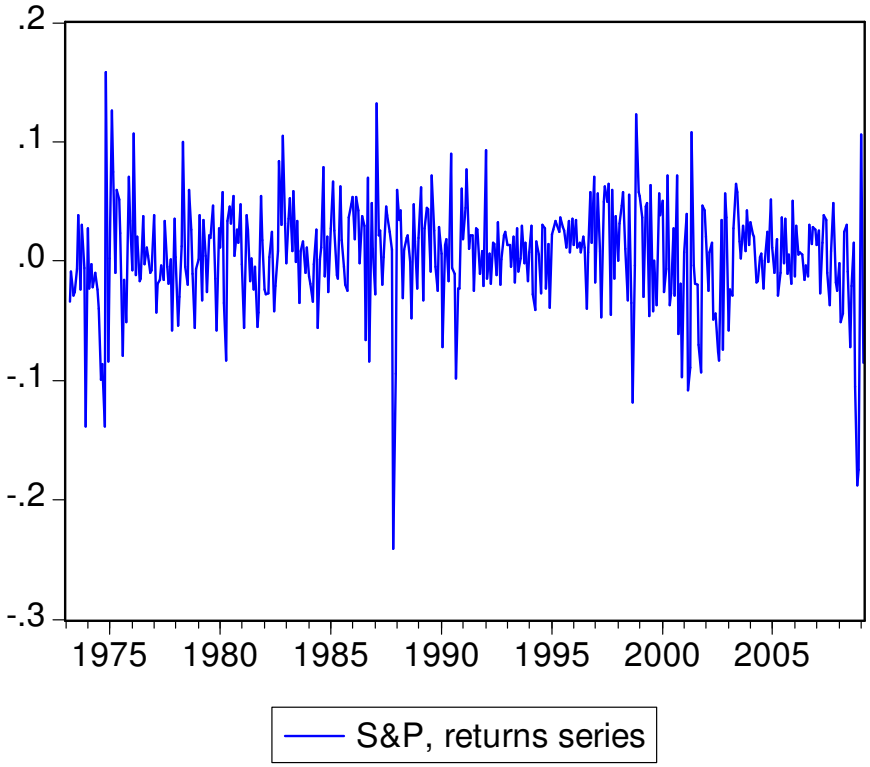


Figure 4.4. Price returns, DAX.

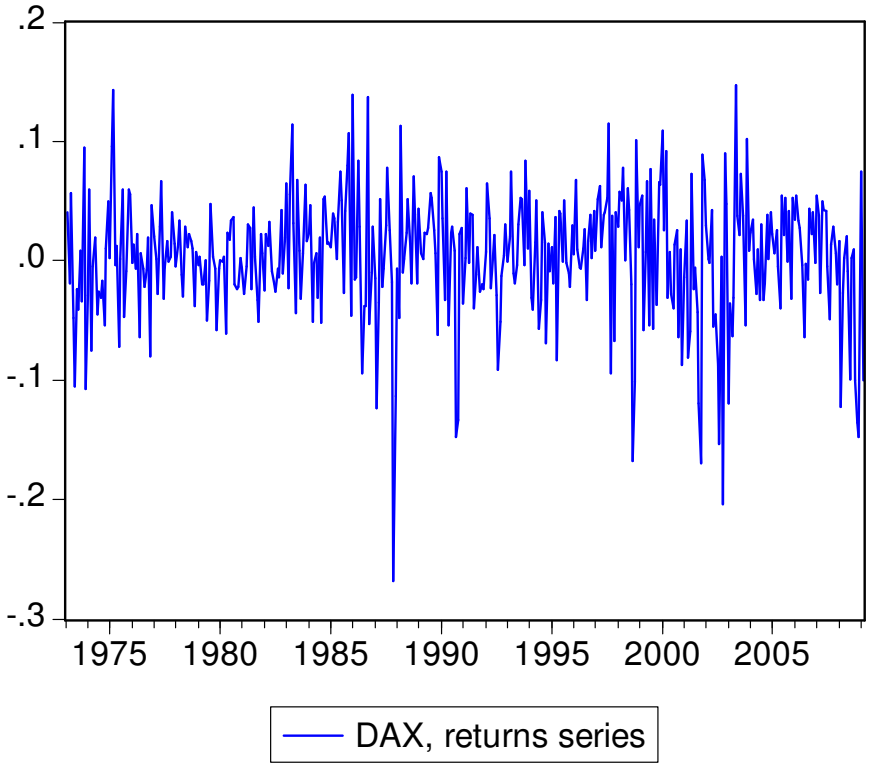


Figure 4.5. Price returns, Nikkei.

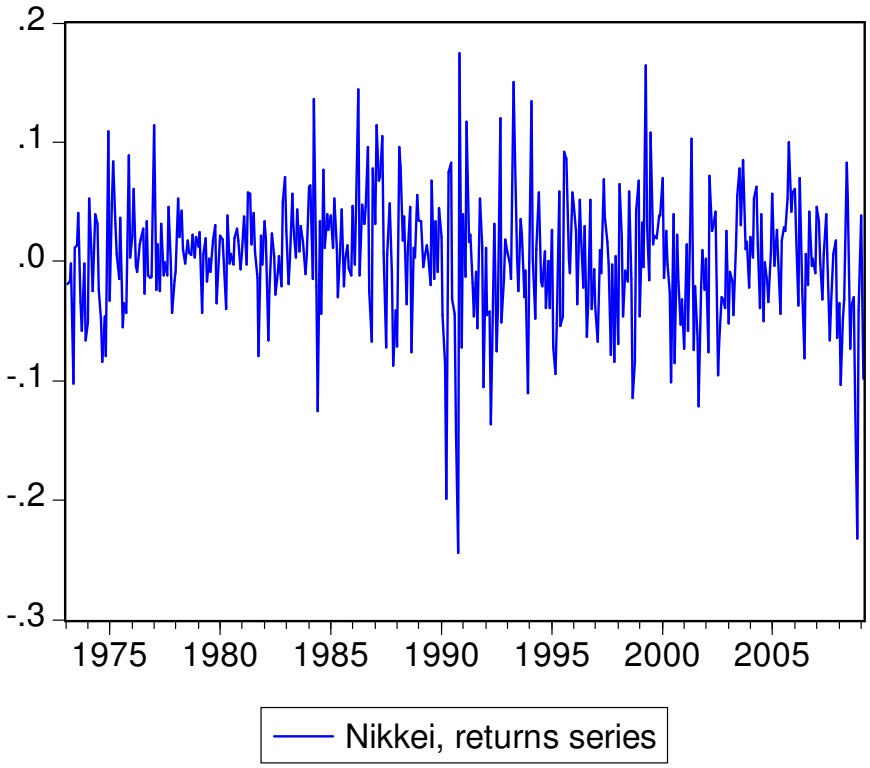
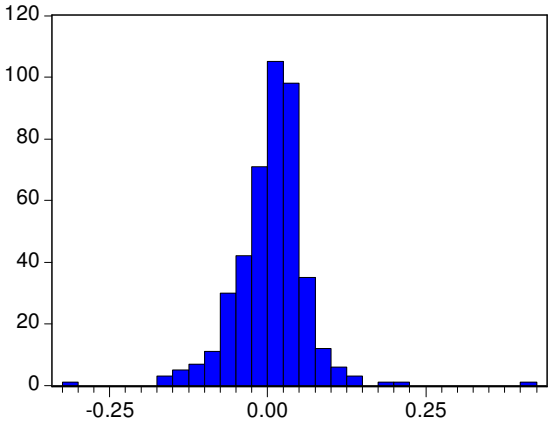
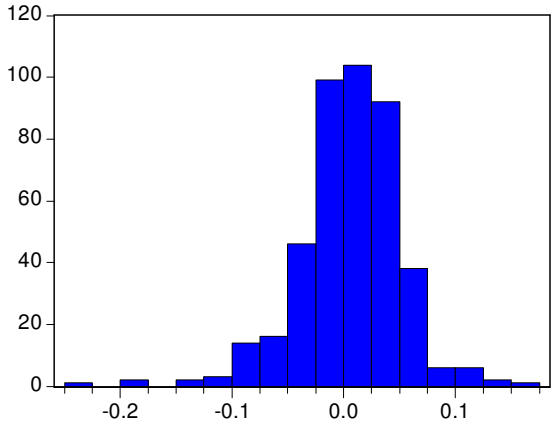


Figure 4.6. Returns histograms: FTSE, S&P, DAX, Nikkei.



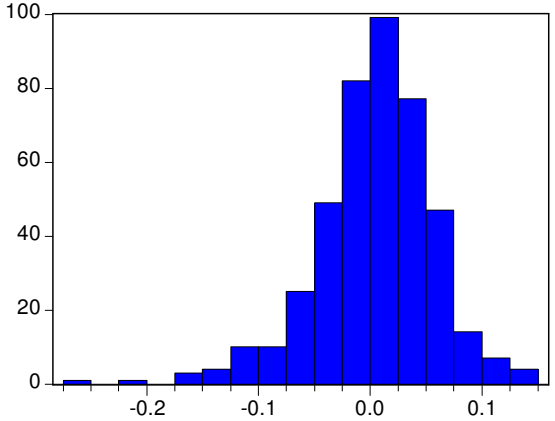
Series: FTSE, returns series
Sample 1973:02 2009:01
Observations 432

Mean	0.005914
Median	0.012367
Maximum	0.400545
Minimum	-0.318231
Std. Dev.	0.056036
Skewness	0.133535
Kurtosis	11.32726
Jarque-Bera	1249.464
Probability	0.000000



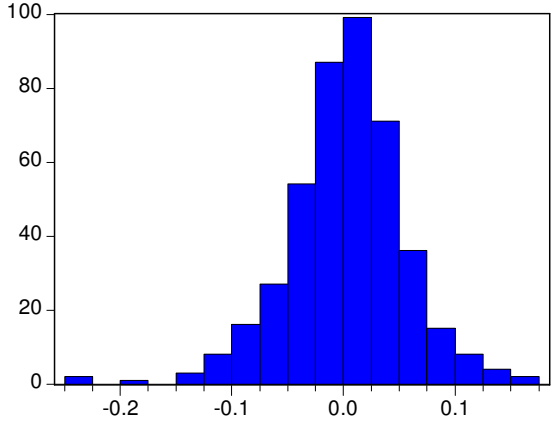
Series: S&P, returns series
Sample 1973:03 2009:02
Observations 432

Mean	0.004777
Median	0.007982
Maximum	0.158126
Minimum	-0.241021
Std. Dev.	0.045269
Skewness	-0.784885
Kurtosis	6.345435
Jarque-Bera	245.8101
Probability	0.000000



Series: DAX, returns series
Sample 1973:02 2009:02
Observations 433

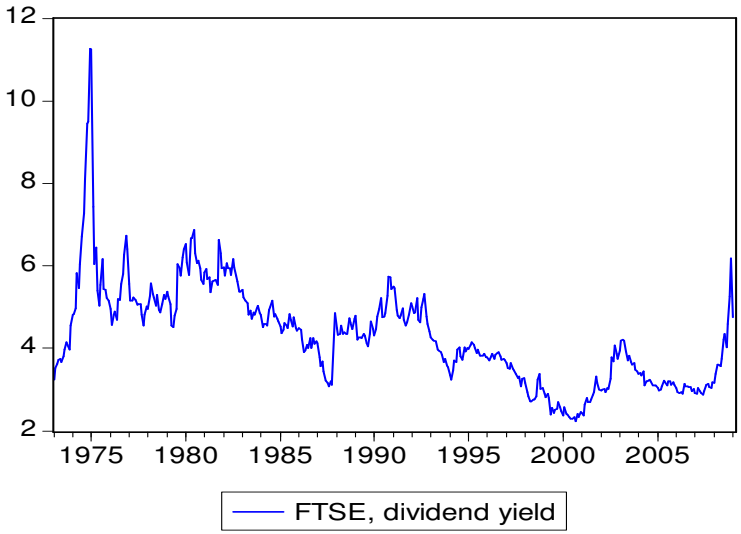
Mean	0.003717
Median	0.006693
Maximum	0.147495
Minimum	-0.268245
Std. Dev.	0.052426
Skewness	-0.810272
Kurtosis	5.567112
Jarque-Bera	166.2761
Probability	0.000000

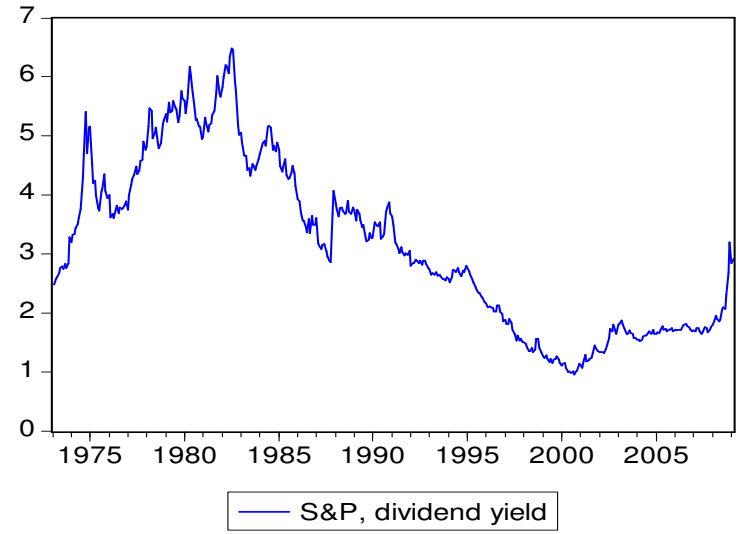


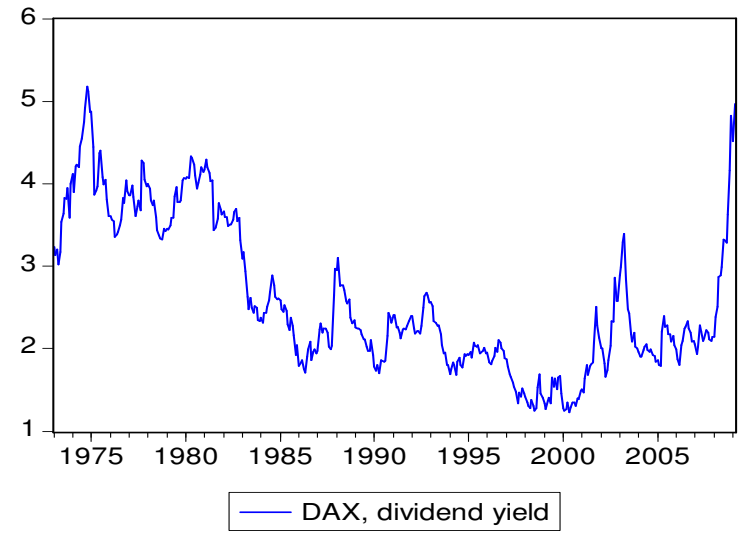
Series: Nikkei, returns series
Sample 1973:02 2009:02
Observations 433

Mean	0.002034
Median	0.004871
Maximum	0.174454
Minimum	-0.244309
Std. Dev.	0.053174
Skewness	-0.406589
Kurtosis	5.186912
Jarque-Bera	98.21600
Probability	0.000000

Figure 4.7. Dividend yield: FTSE, S&P, DAX, Nikkei.







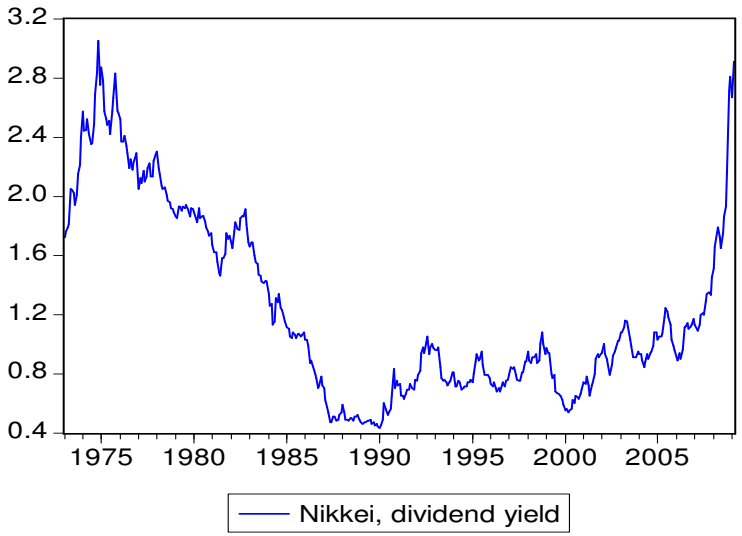
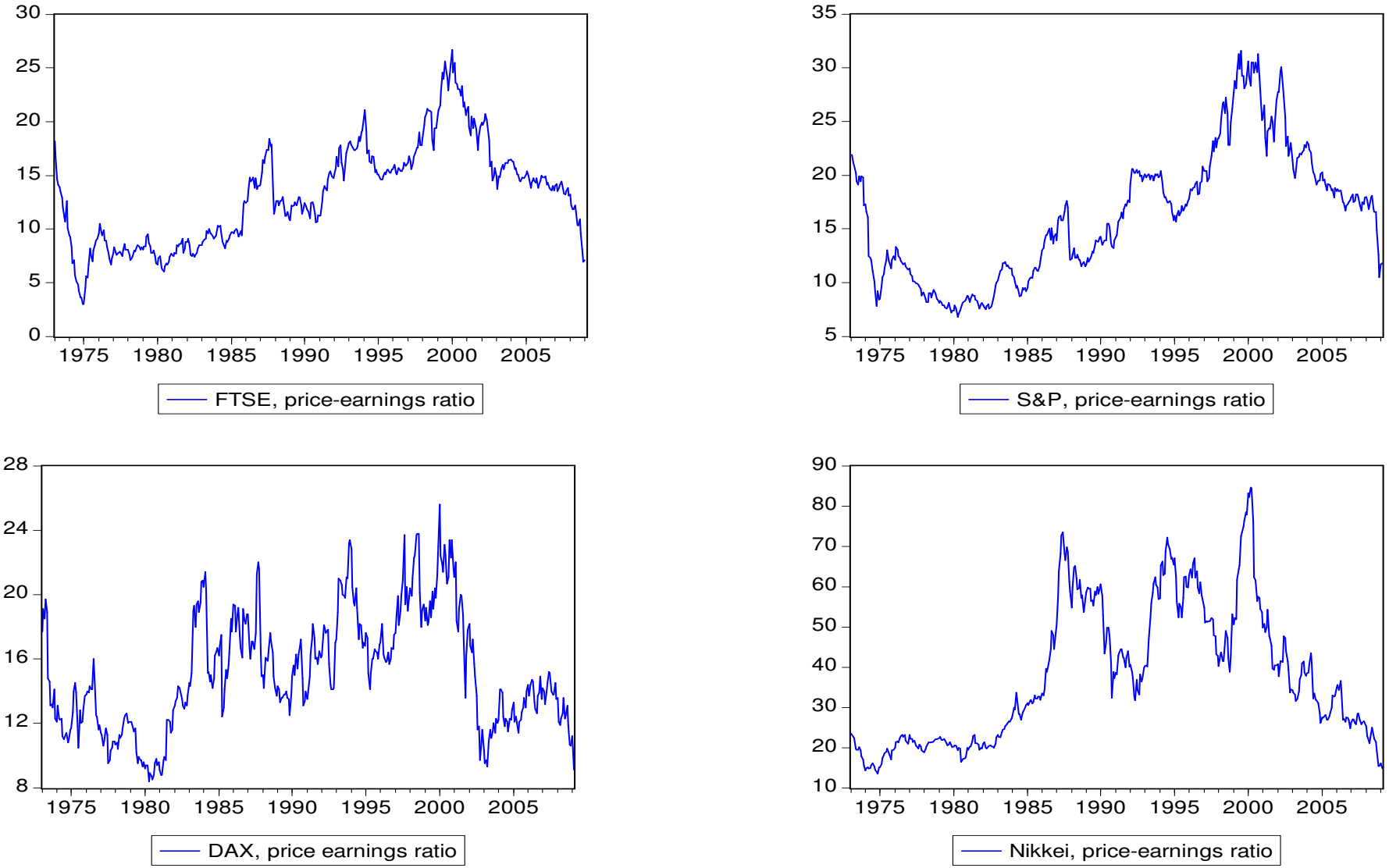


Figure 4.8. Price-earnings ratio: FTSE, S&P, DAX, Nikkei.



Unit root tests

Linear as well as non-linear unit root tests are performed on stock price returns, dividend yields and price-earnings ratios for each time-series data set. Furthermore, three types of data modification were performed on each series, hence, the tests are performed on logs of each time-series as the main data set, de-meaned logs and de-trended logs. The latter two adjustments to the data were performed in order to centre the long-run equilibrium around zero. The augmented Dickey-Fuller (ADF) test is applied as a linear stationarity test. Non-linear unit root tests performed in this chapter include the ESTAR non-linearity test by Kapetanios et al. (2003), the asymmetric ESTAR stationarity test by Sollis (2009), the LSTAR non-linearity and general STAR non-linearity tests by Pascalau (2007).

The augmented Dickey-Fuller (ADF) test was performed as a linear unit root test by testing the null hypothesis of unit root against the alternative of stationarity. Following the results of the ADF test, price returns for all four series were found to be stationary as well as de-trended dividend yield for FTSE series, and log price-earnings ratio, de-meaned and de-trended price-earnings ratios for DAX series; whereas the null hypothesis of unit root could not be rejected for the rest of dividend yield and price-earnings ratio series (Table 4.2).

Table 4.2. ADF unit root test result.

ADF critical values	**1 %**	**5 %**	**10 %**
	-3.4477	-2.8685	-2.5705

	FTSE	***S&P***	***DAX***	***Nikkei***
Price returns	-14.9869*	-14.1755*	-13.5948*	-13.5123*
Log prices	-1.8553	-0.7460	-0.9531	-1.5517
Dividend yield	-3.3907*	-1.2499	-1.1347	-0.0966
Log dividend yield	-2.4873	-1.0158	-1.4904	-0.9696
De-meaned dividend yield	-2.4873	-1.0158	-1.4904	-0.7154
De-trended dividend yield	-3.9341*	-1.8956	-0.9245	0.0317
Price-earnings ratio	-1.6204	-1.5232	-3.2549**	-1.7350
Log price-earnings ratio	-1.8553	-1.5048	-2.9471*	-1.2599
De-meaned price-earnings ratio	-1.8553	-1.5048	-2.9417*	-1.1340
De-trended price-earnings ratio	-2.3628	-2.3337	-2.9959*	-0.4797
Note: * indicates a statistically significant result of stationarity ** indicates a statistically significant stationarity at 5%, 10%				

The unit root test for non-linear ESTAR process (4.13) by Kapetanios et al. (2003) is characterised by testing the null hypothesis of unit root ($H_0: \beta = 0$) against the alternative of stationarity ($H_1: \beta < 0$) using a *t*-type statistic (4.7):

$$\Delta y_t = \beta y_{t-1}^3 + \varepsilon_t \tag{4.13}$$

$$t_{NL} = \hat{\beta} / s.e.(\hat{\beta}) \tag{4.14}$$

where $\hat{\beta}$ is the OLS estimate of β and $s.e.(\hat{\beta})$ is the standard error of $\hat{\beta}$. Asymptotic critical values of the t_{NL} statistic are given in Table 4.3.

ESTAR non-linearity tests by Kapetanios et al. (2003) suggests non-linear stationarity for de-meaned dividend yield and price-earnings ratio of the FTSE index, and de-

meaned price-income ratio and de-trended price-earnings ratio for the DAX index (Table 4.4).

Table 4.3. ESTAR non-linearity critical values.

Fractile (%)	***Raw data***	***De-meaned data***	***De-trended data***
1	-2.82	-3.48	-3.93
5	-2.22	-2.93	-3.40
10	-1.92	-2.66	-3.13

Table 4.4. ESTAR non-linearity unit root test results.

	FTSE	**S&P**	**DAX**	**Nikkei**
Log dividend yield	-1.1711	-0.6970	-0.3884	-0.4482
Log price-earnings ratio	-1.1530	-1.0652	-1.1724	-0.6931
De-meaned dividend yield	-3.6232*	-1.3839	-1.8163	-0.8699
De-meaned price-earnings ratio	-2.6301	-1.8136	-3.3911**	-1.3323
De-trended dividend yield	-5.2398*	-2.2500	1.1080	1.3497
De-trended price-earnings ratio	-2.2859	-2.6156	-3.7893**	0.5142
Note: *indicates a statistically significant result of stationarity. ** reject H_0 at 5%, 10%				

The unit root test proposed by Sollis (2009) is based on the test by Kapetanios et al. (2003) and tests the null hypothesis of unit root $(H_0: \beta = \delta = 0)$ against the alternative of stationary asymmetric ESTAR process (4.15) using specially tabulated critical values (Table 4.5):

$$\Delta y_t = \beta y_{t-1}^3 + \delta y_{t-1}^4 + \varepsilon_t \tag{4.15}$$

Table 4.5. Asymmetric STAR non-linearity critical values.

Fractile (%)	***Zero mean data***	***Non-zero mean data***	***Deterministic trend***
1	4.241	6.236	8.344
5	2.505	4.557	6.292
10	1.837	3.725	5.372

Table 4.6. Asymmetric STAR non-linearity unit root test results.

	FTSE	**S&P**	**DAX**	**Nikkei**
Log dividend yield	6.2301*	1.0976	0.1368	0.5038
Log price-earnings ratio	0.8279	0.9066	4.8183***	0.5038
De-meaned dividend yield	6.7930*	1.0752	2.1091	0.7078
De-meaned price-earnings ratio	3.4935	1.6421	6.8295*	0.7034
De-trended dividend yield	13.8694*	3.0476	3.5921	1.0075
De-trended price-earnings ratio	4.7059	3.5936	8.9656*	3.4536
Note: * indicates a statistically significant result of stationarity. ** reject H_0 at 5%, 10% *** reject H_0 at 1%				

The results of the asymmetric ESTAR stationarity test (Sollis, 2009) in the table above (Table 4.6) suggest that there is a presence of asymmetric ESTAR non-linearity in dividend yield, de-meaned and de-trended dividend yield for FTSE data, and in price-earnings ratio, de-meaned and de-trended price-earnings ratio of DAX series.

Similarly, the general STAR-type (4.16) and logistic STAR (LSTAR) (4.17) unit root tests (Pascalau, 2007) are described as follows:

$$\Delta y_t = \gamma y_{t-1}^2 + \beta y_{t-1}^3 + \delta y_{t-1}^4 + \varepsilon_t \tag{4.16}$$

$$\Delta y_t = \gamma y_{t-1}^2 + \delta y_{t-1}^4 + \varepsilon_t \tag{4.17}$$

Table 4.7. LSTAR non-linearity critical values.

Fractile (%)	*Raw data*	*De-meaned data*	*De-trended data*
1	6.40	5.06	3.73
5	4.51	3.42	2.46
10	3.67	2.66	1.90

Table 4.8. LSTAR non-linearity unit root test results.

	FTSE	**S&P**	**DAX**	**Nikkei**
Log dividend yield	5.9317**	1.0409	0.1348	1.5479
Log price-earnings ratio	0.8538	0.9385	4.7471**	0.6693
De-meaned dividend yield	5.9711*	0.5265	0.4805	1.1305
De-meaned price-earnings ratio	4.4617**	0.19845	0.3854	0.4536
De-trended dividend yield	15.7045*	3.2343**	3.9210*	3.7569*
De-trended price-earnings ratio	7.8908*	3.8516*	1.1178	1.9419
Note: *denotes a statistically significant result of stationarity. ** reject H_0 at 5%, 10%				

On the basis of the test by Pascalau (2007), LSTAR non-linearity is suggested for dividend yield, de-meaned and de-trended dividend yield, de-meaned and de-trended price-earnings ratio of FTSE data; S&P de-meaned and de-trended price-earnings ratio; DAX price-earnings ratio and de-trended dividend yield; and Nikkei de-trended dividend yield (Table 4.8).

Table 4.9. General STAR non-linearity critical values.

Fractile (%)	*Raw data*	*De-meaned data*	*De-trended data*
1	4.92	5.16	6.08
5	3.64	3.87	4.72
10	3.05	3.30	6.08

Table. 4.10. General STAR non-linearity unit root test results.

	FTSE	**S&P**	**DAX**	**Nikkei**
Log dividend yield	4.2219*	0.7432	0.0910	1.1125
Log price-earnings ratio	0.6755	0.7036	3.2674***	0.6163
De-meaned dividend yield	4.6281**	0.7733	1.5667	1.0386
De-meaned price-earnings ratio	3.0388	1.1967	4.5477**	0.8590
De-trended dividend yield	11.9216*	2.5831	3.0433	3.1580
De-trended price-earnings ratio	6.3906*	4.0890	5.9904****	1.3089
Note: * denotes a statistically significant result of stationarity. ** reject H_0 at 5%, 10% *** reject H_0 at 10% **** reject H_0 at 5%				

The test for general STAR non-linearity by Pascalau (2007) detected the presence of non-linearity in dividend yield, de-meaned dividend yield and de-meaned and de-trended price-earnings ratio for FTSE data; and price-earnings ratio, de-meaned and de-trended price-earnings ratio for DAX index (Table 4.10).

Linear and non-linear model estimation and forecasting

Following the unit root tests above, appropriate non-linear STAR models as well as linear alternatives were estimated for all series confirmed to be stationary. The linear benchmark models estimated in this chapter included the random walk model for stock price returns time-series for all four data sets, a simple linear regression for price returns with dividend yield as a regression variable, and a simple linear regression with price-earnings ratio as a regression variable (4.18).

$$y_t = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \cdots + \alpha_i x_{it} + u_t \tag{4.18}$$

Where the price returns, y_t, are regressed on the explanatory variable, x_{it}, which is either the dividend yield or price-earnings ratio.

A non-linear STAR model applied in this chapter follows a general form:

$$r_t = \pi_0 + \sum_{i=1}^{p} \pi_i y_{t-i} + \left(\theta_0 + \sum_{i=1}^{p} \theta_i y_{t-i}\right) F(s_{t-d}) + \varepsilon_t \tag{4.19}$$

where the dependent variable r_t is determined by the explanatory variable y_{t-i}, s_{t-d} is the transition variable, *d* is the delay parameter and ε_t is an error term. π_i and θ_i are the autoregressive components of the model. The transition function, $F(s_{t-d})$, is different for ESTAR (4.20) and LSTAR (4.21) specifications:

$$F(s_{t-d}) = 1 - exp(-\gamma(s_{t-d} - c)^2/\sigma^2(s_{t-d}))\ , \qquad \gamma > 0 \tag{4.20}$$

$$F(s_{t-d}) = \left(1 + exp(-\gamma\,(s_{t-d} - c)/\sigma(s_{t-d}))\right)^{-1}, \qquad \gamma > 0 \tag{4.21}$$

Moreover, the asymmetric ESTAR (AESTAR) model captures different speeds of adjustment, γ_1 and γ_2, following the indication function I_t:

$$S_t(\gamma_1, \gamma_2, z_{t-1}) = [1 + exp\{-\gamma_1^2 z_{t-1}^2 I_t - \gamma_2^2 z_{t-1}^2 (1 - I_t)\}]^{-1} - 0.5 \quad (4.22)$$

$$I_t = 1 \; if \; z_{t-1} > 0 \quad (4.23)$$

$$I_t = 0 \; if \; z_{t-1} \leq 0$$

Table 4.11 represents a list of estimated non-linear STAR models in this chapter for each time-series.

Table 4.11. List of STAR models for FTSE, S&P, DAX and Nikkei.

	STAR models
FTSE	ESTAR de-meaned dividend yield ESTAR de-trended dividend yield AESTAR log dividend yield AESTAR de-meaned dividend yield AESTAR de-trended dividend yield LSTAR log dividend yield LSTAR de-meaned dividend yield LSTAR de-trended dividend yield LSTAR de-meaned pe ratio LSTAR de-trended pe ratio
S&P	LSTAR de-trended dividend yield LSTAR de-trended price-earnings ratio
DAX	ESTAR de-meaned pe ratio ESTAR de-trended pe ratio AESTAR de-meaned price-earnings ratio AESTAR de-trended price-earnings ratio LSTAR de-trended dividend yield LSTAR log pe ratio
Nikkei	LSTAR de-trended dividend yield

The forecasting exercise is performed by incorporating the STAR models into the error-correction model. The specifications for each STAR model, ESTAR (4.24), LSTAR (4.25) and AESTAR (4.26) in particular, are as follows:

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 - exp(-\gamma(s_{t-d} - c)^2/\sigma^2(s_{t-d}))\big) + \varepsilon_t \quad (4.24)$$

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 + exp(-\gamma(s_{t-d} - c)/\sigma(s_{t-d}))\big)^{-1} + \varepsilon_t \quad (4.25)$$

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\left(1 + exp\big(-\gamma_1^2 s_{t-1}^2 I_t - \gamma_2^2 s_{t-1}^2 (1 - I_t)\big)\right)^{-1} + \varepsilon_t \quad (4.26)$$

Thus, further to model estimation, a recursive one-step-ahead out-of-sample forecast is carried out. For the purpose of a forecasting exercise the main sample of thirty six years of monthly data ranging from 1973:01 to 2009:02 with the total of 434 observations is split into an in-sample of eighteen years from 1973:01 to 1990:12, and an out-of-sample of eighteen years from 1991:01 to 2009:02.

Forecasting accuracy tests

In order to establish the most successful forecasting model, all forecasted series are assessed using forecasting accuracy tests including the standard statistical loss functions such as ME, MAE and RMSE, as well as the Diebold-Mariano test of equal forecast accuracy, the forecast encompassing test, combined forecast tests; and a trade rule

technique as an economic loss function test. Refer to Section 2.3 of Chapter 2 for detailed discussion and full methodology of forecast accuracy tests.

ME, MAE, and RMSE

Table 4.12 below includes the accuracy tests results for ME, MAE and RMSE; and a trade rule approach. It is evident from these results that the random walk model for each series is described with the lowest value of statistics, indicating consistent accuracy of forecasts. However, while most models considered produce a positive trade value, as a potential profitability indicator, the highest values within each series are produced by STAR-type models, with the exception of five forecasting series which generate negative trade rule result: linear dividend yield model for FTSE series, random walk for S&P, and random walk and both linear models for Nikkei. The only model to produce a positive trade value for Nikkei series is LSTAR de-trended dividend yield, which is also indicated by the best values of MAE and RMSE for the Japanese index.

Table 4.12. Forecasting accuracy tests results.

	ME	MAE	RMSE	Trade
FTSE				
ESTAR de-meaned dy	0.0003	0.0389	0.0547	0.0065
ESTAR de-trended dy	0.0007	0.0384	0.0540	0.0095
AESTAR log dy	-0.0121	0.0339	0.0442	0.0063
AESTAR de-meaned dy	-0.0116	0.0347	0.0446	0.0057
AESTAR de-trended dy	0.0048	0.0325	0.0413*	0.0038
LSTAR log dy	0.0000*	0.0389	0.0548	0.0060
LSTAR de-meaned dy	0.0000	0.0377	0.0511	0.0077
LSTAR de-meaned pe ratio	-0.0000	0.0386	0.0544	0.0077
LSTAR de-trended dy	0.0000	0.0382	0.0540	0.00991
LSTAR de-trended pe ratio	0.0000**	0.0375	0.0510	0.0104*
Random walk	-0.0039	0.0319*	0.0432**	0.0039
Linear dy	0.0070	0.0341	0.0438	-0.0036
Linear pe ratio	0.0032	0.0331**	0.0436	0.0005
S&P				
LSTAR de-trended dy	-0.0000*	0.0323**	0.0439	0.0078*
LSTAR de-trended pe ratio	-0.0000**	0.0328	0.0444	0.0071
Random walk	-0.0014	0.0320*	0.0434*	-0.0030
Linear dy	0.0027	0.0334	0.0442	0.0014
Linear pe ratio	0.0019	0.0329	0.0438**	0.0049
DAX				
ESTAR de-meaned pe ratio	0.0003	0.0384**	0.0518**	0.0066
ESTAR de-trended pe ratio	0.0009	0.0384	0.0520	0.0036
AESTAR de-meaned pe	0.0063	0.0443	0.0566	0.0113*
AESTAR de-trended pe	0.0046	0.0443	0.0562	0.0074
LSTAR de-trended dy	-0.0000**	0.0381*	0.0517*	0.0047
LSTAR log pe ratio	0.0000*	0.0384	0.0521	0.0061
Random walk	-0.0017	0.0411	0.0549	0.0032
Linear dy	-0.0020	0.0411	0.0550	0.0032
Linear pe ratio	-0.0009	0.0412	0.0551	0.0003
NIKKEI				
LSTAR de-trended dy	0.0000*	0.0381*	0.0510*	0.0074*
Random walk	-0.0075	0.0429	0.0554**	-0.0030
Linear dy	-0.0064	0.0428**	0.0554**	-0.0030
Linear pe ratio	-0.0066**	0.0428	0.0555	-0.0030
Note:	* indicates the best statistic ** indicates the second best statistic dy – dividend yield; pe –price-earnings			

Diebold-Mariano tests

The Diebold-Mariano test of equal forecasting accuracy (Table 4.13 – 4.16), where the hypothesis of equal forecast accuracy is tested using standard normal distribution critical values, produced insignificant test statistics for all competing forecasts suggesting that the differences in values of MEs between those forecasts are not statistically different, thus suggesting that it is not possible to draw valid conclusions on the basis of these tests. Similarly, the modified Diebold-Mariano test failed to identify any statistically significant differences between these values.

Table 4.13. Diebold-Mariano test results, FTSE.

FTSE	*DM statistic*	*DM modified*
RW/ linear dy	-0.0509	-0.0508
RW/ linear pe ratio	-0.0493	-0.0492
RW/ ESTAR de-meaned dy	0.0454	0.0453
RW/ ESTAR de-trended dy	0.0023	0.0023
RW/ AESTAR log dy	-0.0915	-0.0913
RW/ AESTAR de-meaned dy	-0.0767	-0.0765
RW/ AESTAR de-trended dy	0.1081	0.1078
RW/ LSTAR log dy	0.0366	0.0365
RW/ LSTAR de-meaned dy	0.0466	0.0464
RW/ LSTAR de-meaned pe ratio	0.1026	0.1023
RW/LSTAR de-trended dy	0.0295	0.0294
RW/ LSTAR de-trended pe ratio	0.0852	0.0850
Linear dy/ linear pe	0.0234	0.0234
Linear dy/ ESTAR de-meaned dy	0.1333	0.1330
Linear dy/ ESTAR de-trended dy	0.0499	0.0498
Linear dy/ AESTAR log dy	-0.0651	-0.0650
Linear dy/ AESTAR de-meaned dy	-0.0507	-0.0506
Linear dy/ AESTAR de-trended dy	0.1408	0.1405
Linear dy/ LSTAR log dy	0.1232	0.1229
Linear dy/ LSTAR de-meaned dy	0.1318	0.1315
Linear dy/ LSTAR de-meaned pe	0.1650	0.1646
Linear dy/ LSTAR de-trended dy	0.0858	0.0857
Linear dy/ LSTAR de-trended pe	0.1137	0.1134
Linear pe/ linear pe	-0.0234	-0.0234
Linear pe/ ESTAR de-meaned dy	0.0924	0.0922
Linear pe/ ESTAR de-trended dy	0.0321	0.0320
Linear pe/ AESTAR log dy	-0.0939	-0.0937
Linear pe/ AESTAR de-meaned dy	-0.0709	-0.0707
Linear pe/ AESTAR de-trended dy	0.1405	0.1402
Linear pe/ LSTAR log dy	0.0830	0.0828
Linear pe/ LSTAR de-meaned dy	0.0977	0.0974
Linear pe/ LSTAR de-meaned pe	0.1440	0.1437
Linear pe/ LSTAR de-trended dy	0.0624	0.0622
Linear pe/ LSTAR de-trended pe	0.1060	0.1058
* indicates statistical significance at 5%		

Table 4.14. Diebold-Mariano test results, DAX.

DAX	*DM statistic*	*DM modified*
RW/ linear dy	-0.0709	-0.0707
RW/ linear pe	-0.0971	-0.0969
RW/ ESTAR de-meaned pe	0.0427	0.0426
RW/ ESTAR de-trended pe	0.0259	0.0259
RW/ AESTAR de-meaned pe	0.1657	0.1653
RW/ AESTAR de-trended pe	0.1644	0.1640
RW/ LSTAR de-trended dy	0.0691	0.0690
RW/ LSTAR log pe	0.0333	0.0332
Linear dy/ ESTAR de-meaned pe	0.0477	0.0476
Linear dy/ ESTAR de-trended pe	0.0357	0.0356
Linear dy/ AESTAR de-meaned pe	0.1709	0.1705
Linear dy/ AESTAR de-trended pe	0.1710	0.1706
Linear dy/ LSTAR de-trended dy	0.0736	0.0734
Linear dy/ LSTAR log pe	0.0445	0.0444
Linear pe/ ESTAR de-meaned pe	0.0609	0.0607
Linear pe/ ESTAR de-trended pe	0.0497	0.0496
Linear pe/ AESTAR de-meaned pe	0.1639	0.1635
Linear pe/ AESTAR de-trended pe	0.1625	0.1621
Linear pe/ LSTAR de-trended dy	0.0822	0.0820
Linear pe/ LSTAR log pe	0.0611	0.0610

Table 4.15. Diebold-Mariano test results, S&P.

S&P	*DM statistic*	*DM modified*
RW/ linear dy	-0.1954	-0.1949
RW/ linear pe	-0.1343	-0.1340
RW/ LSTAR de-trended dy	0.0704	0.0703
RW/ LSTAR de-trended pe	0.0834	0.0832
Linear dy/ linear pe	0.1847	0.1843
Linear dy/ LSTAR de-trended dy	0.0918	0.0916
Linear dy/ LSTAR de-trended pe	0.1201	0.1198
Linear pe/ LSTAR de-trended dy	0.0813	0.0811
Linear pe/ de-trended pe	0.1006	0.1003

Table 4.16. Diebold-Mariano test results, Nikkei.

Nikkei	*DM statistic*	*DM modified*
RW/ linear dy	-0.0000	-0.0000
RW/ linear pe	-0.0242	-0.0241
RW/ LSTAR de-trended dy	0.1148	0.1145
Linear dy/ linear pe	-0.0538	-0.0537
Linear dy/ LSTAR de-trended dy	0.1106	0.1104
Linear pe/ linear dy	0.0538	0.0537
Linear pe/ LSTAR de-trended dy	0.1133	0.1131

Forecast encompassing test

For the forecast encompassing test, appropriate STAR models are to be considered against random walk and linear models respectively. In addition, two types of encompassing tests are applied whereby the first test assesses whether one model's forecast encompasses the forecast of the other model.

$$y_{t+s} = \alpha + \beta_1 f_{t,s}^{RW} + \beta_2 f_{t,s}^{STAR} + u_t \tag{4.27}$$

$$y_{t+s} = \alpha + \beta_1 f_{t,s}^{Linear} + \beta_2 f_{t,s}^{STAR} + u_t \tag{4.28}$$

where $f_{t,s}^{RW}$, $f_{t,s}^{Linear}$, $f_{t,s}^{STAR}$ are the forecasts obtained from a random walk model, linear regression and STAR model respectively. The null hypothesis of the first model encompassing the forecast of the second ($H_0: \beta_1 = 0$) is tested against the alternative of the first model forecast being encompassed by the second model ($H_1: \beta_2 > 0$).

The second test of the forecast encompassing uses the same hypotheses and determines whether forecast errors of one forecast can explain the forecasting errors of the other forecast.

$$y_{t+s} = \alpha + \beta_1 \left(f_{t,s}^{RW} - y_{t+s}\right) + \beta_2 \left(f_{t,s}^{STAR} - y_{t+s}\right) + u_t \tag{4.29}$$

$$y_{t+s} = \alpha + \beta_1 \left(f_{t,s}^{Linear} - y_{t+s}\right) + \beta_2 \left(f_{t,s}^{STAR} - y_{t+s}\right) + u_t \tag{4.30}$$

where $(f_{t,s}^{RW} - y_{t+s})$, $(f_{t,s}^{Linear} - y_{t+s})$, $(f_{t,s}^{STAR} - y_{t+s})$ are the forecasting errors from the random walk model, linear regression, and STAR model respectively.

Results in Tables 4.17 – 4.20 demonstrate statistical significance for most of the β_2 coefficients for all of the four series, thus implying that the STAR models are not encompassed by the linear alternatives and contain independent information for forecasting of the dependent variable. However, the β_2 coefficients are not significant for two models for FTSE index including LSTAR log dividend yield and LSTAR de-meaned dividend yield, suggesting that the both models are encompassed by the price-earnings ratio linear regression. In addition, ESTAR de-meaned and de-trended price-earnings ratio models for DAX index are seem to be encompassed by the linear alternatives. The forecast errors encompassing test, on the other hand, reveals significant β_1 coefficients for all series with only a few significant β_2 coefficients, including AESTAR de-trended dividend yield not being encompassed by the random walk model, and AESTAR log dividend yield and AESTAR de-meaned dividend yield not encompassed by the price-earnings linear regression for FTSE series; AESTAR de-meaned price-earnings ratio not encompassed by the random walk model, and AESTAR de-trended price-earnings ratio not encompassed by neither the random walk model nor dividend yield linear regression for the DAX index. The results of the forecasting errors encompassing test suggest that the forecasting errors from other STAR models are explained by the linear alternatives. Overall, the STAR models seem to encompass random walk and linear regression models with the exception of nine forecasts where both non-linear and linear models contain independent information for forecasting the price returns series. However, according to the second forecast encompassing test, it

seems that the STAR models forecasting errors are mostly explained by the linear alternatives.

Table 4.17. Forecast encompassing test results, DAX.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
DAX				
RW/ ESTAR de-meaned pe ratio	-0.1894	1.7479	-191.0603*	-7.6227
RW/ ESTAR de-trended pe ratio	-0.6849	1.1246	-133.2827*	-5.0151
RW/ AESTAR de-meaned pe ratio	-2.3542*	6.3269*	-281.4411*	3.0653*
RW/ AESTAR de-trended pe ratio	-4.1589*	7.4034*	-368.1000*	7.5680*
RW/ LSTAR de-trended dy	-0.5541	2.3114*	-159.2983*	-3.8452
RW/ LSTAR log pe ratio	-0.7397	1.0783	-106.6783*	-4.8778
Linear dy/ ESTAR de-meaned pe ratio	0.0481	1.7687	-79.8332*	-11.6764
Linear dy/ ESTAR de-trended pe ratio	-0.6354	1.1427	-51.5763*	-7.7702
Linear dy/ AESTAR de-meaned pe ratio	-2.0484*	6.2126*	-102.1599*	1.1466
Linear dy/ AESTAR de-trended pe ratio	-3.6694*	7.1054*	-129.1773*	5.3742*
Linear dy/ LSTAR de-trended dy	-0.0352	2.2667*	-67.2414*	-9.4780
Linear dy/ LSTAR log pe ratio	-0.6993	1.1004	-41.1691*	-8.1833
Linear pe/ ESTAR de-meaned pe ratio	-1.2434	1.7797	-60.1919*	-5.7994
Linear pe/ ESTAR de-trended pe ratio	-1.9080	1.7406	-40.5962*	-1.3071
Linear pe/ AESTAR de-meaned pe ratio	4.0102*	7.1074*	-155.8963*	-19.6729
Linear pe/ AESTAR de-trended pe ratio	1.7662	6.0578*	-138.4680*	-11.1558
Linear pe/ LSTAR de-trended dy	-0.5382	1.9750*	-64.4797*	-11.2415
Linear pe/ LSTAR log pe ratio	-1.7124	1.4636	-32.8459*	-4.1130

Note : * statistically significant at 5%
RW – random walk; dy – dividend yield; pe – price-earnings

Table 4.18. Forecast encompassing test results, FTSE.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
FTSE				
RW/ ESTAR de-meaned dy	0.4950	2.5312*	-134.8471*	-1.4139
RW/ ESTAR de-trended dy	1.4402	2.3565*	-241.4808*	-7.6584
RW/ AESTAR log dy	1.7543	3.0338*	-434.3917*	-7.4635
RW/ AESTAR de-meaned dy	1.9400	3.6377*	-453.9381*	-7.0945
RW/ AESTAR de-trended dy	-2.2942	6.9683*	-211.9156*	5.7580*
RW/ LSTAR log dy	0.5001	2.1874*	-139.4642*	-1.3610
RW/ LSTAR de-meaned dy	0.4585	2.1426*	-121.2643*	-1.3012
RW/ LSTAR de-meaned pe ratio	1.8523	3.6254*	-166.3697*	-7.6874
RW/ LSTAR de-trended dy	1.6673	3.0554*	-256.4971*	-7.2494
RW/ LSTAR de-trended pe ratio	1.5309	4.6814*	-335.8900*	-4.0968
Linear dy/ ESTAR de-meaned dy	-0.6484	2.5686*	-13.5383*	-5.4864
Linear dy/ ESTAR de-trended dy	-0.7167	1.9944*	-25.9425*	1.6933
Linear dy/ AESTAR log dy	1.0846	2,6965*	-58.3128*	-6.7971
Linear dy/ AESTAR de-meaned dy	1.0923	3.2515*	-60.4339*	-5.9897
Linear dy/ AESTAR de-trended dy	6.1199*	9.3377*	-50.4793*	-23.3327
Linear dy/ LSTAR log dy	-0.4638	2.1816*	-14.3159*	-5.1683
Linear dy/ LSTAR de-meaned dy	-0.3276	2.1204*	-12.0460*	-7.5843
Linear dy/ LSTAR de-meaned pe ratio	-0.5951	3.1525*	-16.2422*	-5.5821
Linear dy/ LSTAR de-trended dy	-1.1740	2.8109*	-28.6574*	2.3281
Linear dy/ LSTAR de-trended pe ratio	-0.0495	4.4013*	-42.9094*	-2.4896
Linear pe/ ESTAR de-meaned dy	-0.2355	2.3188*	-34.3979*	-12.8588
Linear pe/ ESTAR de-trended dy	-1.4226	2.1609*	-42.4192*	-1.8923
Linear pe/ AESTAR log dy	-4.6297*	5.1483*	-120.2249*	16.4189*
Linear pe/ AESTAR de-meaned dy	-5.0127*	5.8230*	-116.7414*	13.9952*
Linear pe/ AESTAR de-trended dy	-2.3142*	6.8524*	-41.6223*	-5.7271
Linear pe/ LSTAR log dy	-0.3677	1.9591	-35.2051*	-12.4487
Linear pe/ LSTAR de-meaned dy	-0.3765	1.9215	-31.3458*	-13.6364
Linear pe/ LSTAR de-meaned pe ratio	-1.3328	3.2495*	-29.2816*	-5.3969
Linear pe/ LSTAR de-trended dy	-1.2779	2.7039*	-46.4709*	-2.8864
Linear pe/ LSTAR de-trended pe ratio	-1.6684	4.6366*	-65.6892*	-0.6592
Note : * statistically significant at 5% RW – random walk; dy – dividend yield; pe – price-earnings				

Table 4.19. Forecast encompassing test results, S&P.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
S&P				
RW/ LSTAR de-trended dividend yield	-1.0534	5.1553*	-248.4476*	0.7505
RW/ LSTAR de-trended pe ratio	-0.3894	4.0980*	-147.3711*	-2.0113
Linear dy/ LSTAR de-trended dy	-0.5272	4.8332*	-48.5960*	-7.5068
Linear dy/ LSTAR de-trended pe ratio	-0.0635	3.7991*	-33.9624*	-14.0695
Linear pe/ LSTAR de-trended dy	-0.1819	4.8326*	-69.0613*	-7.1047
Linear pe/ LSTAR de-trended pe ratio	0.2361	3.8398*	-47.4371*	-12.7020
Note : * statistically significant at 5% RW – random walk; dy – dividend yield; pe – price-earnings				

Table 4.20. Forecast encompassing test results, Nikkei.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
Nikkei				
RW/ LSTAR de-trended dy	-1.0015	5.8361*	-240.1669*	0.8731
Linear dy/ LSTAR de-trended dy	-0.0766	5.4924*	-205.8363*	-5.0758
Linear pe/ LSTAR de-trended dy	-0.9466	5.5894*	-169.7870*	-2.6617
Note : * statistically significant at 5% RW – random walk; dy – dividend yield; pe – price-earnings				

Forecasting accuracy tests for a combined forecast

Statistical loss function tests were performed on forecast combinations consisting of random walk, linear regression and STAR models. Following results of forecast encompassing tests, it is expected for forecast combinations to demonstrate lower values of ME, MAE and RMSE statistics. Overall these expectations are confirmed. In the case of individual forecasts, the statistics delivered by random walk and linear regressions seemed to dominate over the STAR models' results. Although by a marginal amount, combined forecasts, on the other hand, seem to produce better statistics than

individual linear alternatives. Moreover, all the forecast combinations generate positive values for the trade rule approach.

Table 4.21. Combination forecasts test results.

	ME	MAE	RMSE	Trade
FTSE				
RW/linear dy/ ESTAR de-meaned dy	0.0003	0.0319	0.0424	0.0081
RW/ linear dy/ ESTAR de-trended dy	-0.0013	0.0318	0.0427	0.0094**
RW/ linear dy/ AESTAR log dy	0.0005*	0.0317	0.0424	0.0107
RW/ linear dy/ AESTAR de-meaned dy	0.0005	0.0316	0.0421	0.0119
RW/ linear dy/ AESTAR de-trended dy	0.0053	0.0315	0.0407	0.0084
RW/ linear dy/ LSTAR log dy	0.0001	0.0320	0.0425	0.0088
RW/ linear dy/ LSTAR de-meaned dy	0.0002	0.0320	0.0426	0.0088
RW/ linear dy/ LSTAR de-meaned pe ratio	-0.0008	0.0315	0.0421	0.0083
RW/ linear dy/ LSTAR de-trended dy	-0.0018	0.0312	0.0424	0.0096*
RW/ linear dy/ LSTAR de-trended pe ratio	-0.0004	0.0311**	0.0427	0.0082
RW/ linear pe/ ESTAR de-meaned dy	0.0004	0.0320	0.0424	0.0078
RW/ linear pe/ ESTAR de-trended dy	-0.0024	0.0317	0.0427	0.0088
RW/ linear pe/ AESTAR log dy	-0.0051	0.0314	0.0424	0.0078
RW/ linear pe/ AESTAR de-meaned dy	-0.0054	0.0312	0.0420	0.0130
RW/ linear pe/ AESTAR de-trended dy	0.0030	0.0309	0.0399	0.0128
RW/ linear pe/ LSTAR log dy	6.54E-05	0.0321	0.0425	0.0078
RW/ linear pe/ LSTAR de-meaned dy	6.52E-05	0.0321	0.0426	0.0082
RW/ linear pe/ LSTAR de-meaned pe ratio	-0.0017	0.0314	0.0420**	0.0076
RW/ linear pe/ LSTAR de-trended dy	-0.0023**	0.0311**	0.0424	0.0091
RW/ linear pe/ LSTAR de-trended pe ratio	-0.0023**	0.0309*	0.0411*	0.0087
S&P				
RW/linear dy/ LSTAR de-trended dy	-0.0006*	0.0304*	0.0410*	0.0096*
RW/ linear dy/ LSTAR de-trended pe ratio	-0.0003**	0.0312	0.0417**	0.0086
RW/ linear pe/ LSTAR de-trended dy	0.0001	0.0305**	0.0410*	0.0096*
RW/ linear pe/ LSTAR de-trended pe ratio	0.0002	0.0312	0.0417**	0.0080
DAX				
RW/ linear dy/ ESTAR de-meaned pe ratio	0.0004	0.0414	0.0541**	0.0047
RW/ linear dy/ ESTAR de-trended pe ratio	0.0005	0.0412**	0.0542	0.0060
RW/ linear dy/ AESTAR de-meaned pe ratio	0.0057	0.0395	0.0509	0.0128
RW/ linear dy/ AESTAR de-trended pe ratio	0.0064	0.0391	0.0501	0.0108
RW/linear dy/ LSTAR de-trended dy	0.0002	0.0411*	0.0540*	0.0070
RW/ linear dy/ LSTAR log pe ratio	0.0005	0.0413	0.0543	0.0065
RW/ linear pe/ ESTAR de-meaned pe ratio	-0.0007**	0.0414	0.0541**	0.0059
RW/ linear pe/ ESTAR de-trended pe ratio	-0.0008*	0.0412**	0.0542	0.0078
RW/ linear pe/ AESTAR de-meaned pe ratio	0.0124	0.0389	0.0498	0.0142*
RW/ linear pe/ AESTAR de-trended pe ratio	0.0087	0.0395	0.0501	0.0082
RW/ linear pe/ LSTAR de-trended dy	-8.97E-05	0.0411*	0.0540*	0.0061
RW/ linear pe/ LSTAR log pe ratio	-0.0008*	0.0413	0.0543	0.0064

Nikkei				
RW/ linear dy/ LSTAR de-trended dy	-0.0009**	0.0396*	0.0510*	0.0074
RW/ linear pe/ LSTAR de-trended dy	-0.0028*	0.0397**	0.0512**	0.0050
Note :	* indicates the best statistic ** indicated the second best statistic RW – random walk; dy – dividend yield; pe – price-earnings			

Taking into consideration all of the above tests of forecasting accuracy it seems that while linear models produce the best comparison statistics, STAR models generate the highest profits according to the trade rule test. Moreover, combined forecasts seem to outperform both linear and non-linear models individually.

The combinations for the FTSE series of random walk, dividend yield linear regression and LSTAR de-trended dividend yield and a combination of random walk, price-earnings ratio linear regression and LSTAR de-trended dividend yield, both produce the best combination of test statistics including the trade rule profit. However, according to the forecast errors encompassing test, the forecasting errors for both LSTAR models are encompassed by dividend yield linear regression, suggesting that the combination of random walk and linear regression might be responsible for superior statistics. Similarly, a combination of random walk, dividend yield linear regression and LSTAR de-trended dividend yield for DAX series generates the best statistical combination, while the forecasting errors of the LSTAR model are encompassed by price-earnings ratio linear regression. The combination of the random walk model, linear dividend yield regression and STAR de-trended dividend yield; and a combination of a random walk model, linear price-earnings ratio regression and LSTAR de-trended dividend yield for S&P returns, both generate good statistics and the highest value for the trade rule. Moreover, according to the encompassing test, the LSTAR model is not encompassed by either linear alternatives. Out of two forecast combinations for the

Nikkei index, the random walk, dividend yield linear regression and LSTAR de-trended dividend yield combination, seem to be superior to the other combination, however the forecasting errors of the LSTAR model in this forecasting arrangement are indicated to be encompassed by the random walk model, while containing independent information according to the simple forecast encompassing test. On the basis of results of the forecast encompassing test, combinations containing the STAR model encompassed in one of the linear alternatives were dismissed.

For the FTSE index LSTAR de-trended price-earnings ratio produced the highest trade value of 0.0104, closely followed by its combined forecast with random walk and dividend yield linear regression with trade value of 0.0096. The random walk model forecast produces the best ME, MAE and RMSE statistics. While LSTAR de-meaned price-earnings ratio and a combination of random walk, price-earnings ratio linear regression and LSTAR de-meaned price-earnings ratio both generate forecasts with relatively low statistics and reasonable positive trade values.

Combinations of random walk, dividend yield linear regression and LSTAR de-trended dividend yield, and random walk, price-earnings linear regression and LSTAR de-trended dividend yield forecasts for the S&P index generate very good low statistics as well as the highest value for the trading rule test of 0.0096. The best individual forecasting model for S&P is LSTAR de-trended dividend yield, while the random walk model is the only forecast to produce negative value for the trade rule for the US series.

For the DAX series AESTAR de-meaned price-earnings ratio model is characterised by good statistical results and possesses the highest value of the trade test among the individual forecasts. Both the random walk and dividend yield linear regression while having the lowest statistics, do not demonstrate strong trade test value. The best forecast

for the DAX index appears to be a combined forecast of random walk, price-earnings ratio linear regression and ESTAR de-trended price-earnings ratio with the best statistics for this series and the highest trade test value, followed by other two combinations of random walk, price-earnings ratio linear regression with LSTAR de-trended dividend yield, and LSTAR log price-earnings ratio respectively.

The Nikkei index was the only series to have all individual linear models produce negative trading rule test results. A combined random walk, dividend yield linear regression and LSTAR de-trended dividend yield generated very promising results, however the LSTAR forecast appears to be encompassed by the random walk model both in forecast combination and individually. However, the same model produces the best forecasting model for Nikkei series is the combination with random walk and price-earnings ratio linear regression generating relatively low statistics and a positive trade rule test value. Table 4.22 summarises all the best forecasting models for each of the four series with the results of statistical results for each forecast.

Table 4.22. Test statistics results for best forecasting models.

	ME	MAE	RMSE	Trade
FTSE				
LSTAR de-meaned pe ratio	0.0000	0.0386	0.0544	0.0077
RW/ linear pe/ LSTAR de-meaned pe ratio	-0.0017	0.0314	0.0420	0.0076
S&P				
RW/linear dy/LSTAR de-trended dy	-0.0006	0.0304	0.0410	0.0096
RW/linear pe/ LSTAR de-trended dy	0.0001	0.0305	0.0410	0.0096
DAX				
RW/linear pe/ ESTAR de-trended pe ratio	-0.0008	0.0412	0.0542	0.0078
Nikkei				
RW/ linear pe/ LSTAR de-trended dy	-0.0028	0.0397	0.0512	0.0050
Note:	RW – random walk dy – dividend yield pe – price-earnings			

4.5. Empirical results for 3-, 6-, and 12-month returns

Introduction

While Fair and Shiller (1990) suggested that changes in economic structure and changes in the behaviour of data dynamics are more evident in long-horizon data, Montgomery et al. (1998) demonstrated that forecasting models performed on less frequent data series displayed smoother trends and generated better forecasting performance while still capturing cyclical and trend characteristics of the data. Hence, expanding the topic of long-horizon returns predictability and building on research by Montgomery et al. (1998) in an attempt to investigate whether non-linear models could be utilised to

generate more efficient forecasts in long-horizon frequency data, this section of the chapter will consider extending the investigation of monthly returns forecast to three, six and twelve month forecasts.

The methodological approach to the investigation of long-horizon price returns will be based on the methodology applied in the empirical study of monthly returns using a dividend yield and price-earnings ratio. The time-series for long-horizon returns is designed as a straightforward buy-and-hold strategy, where the stock is assumed to be held for three, six, or twelve months before selling. The strategy is repeated recursively for the duration of the data set. The stock price return at the end of each period is forecasted using the predictive variables, namely the dividend yield and price-earnings ratio, for the same period as opposed to values from the previous period in the monthly forecasting framework. In addition to non-linear STAR models, linear alternatives, specifically the random walk model and simple linear regression, will be estimated to provide comparative benchmarks of forecasting performance.

Unit root tests, model estimation and forecasting

The linear benchmark in the form of a random walk model and a linear regression with either dividend yield or price-earnings ratio as the determinant variable, as well as STAR models, were estimated for each long-horizon period of three, six and twelve months. The choice of appropriate STAR models for this section is based on the non-linear unit root tests performed in the previous section (Section 4.4) and selected at a limitation of a 5% level of significance (Table 4.23).

Further to model estimation, a recursive one-step-ahead out-of-sample forecast is carried out. The in-sample and out-of-sample periods are the same as in Section 4.4. Thus, the main sample of 36 years of monthly data ranging from 1973:01 to 2009:02 with a total of 434 observations is split into an in-sample of eighteen years from 1973:01 to 1990:12, and an out-of-sample of eighteen years from 1991:01 to 2009:02.

Table 4.23. List of estimated STAR models.

	STAR models
FTSE	AESTAR log dividend yield AESTAR de-meaned dividend yield AESTAR de-trended dividend yield ESTAR de-meaned dividend yield ESTAR de-trended dividend yield LSTAR de-meaned dividend yield LSTAR de-trended dividend yield
S&P	LSTAR de-trended price-earnings ratio
DAX	AESTAR de-meaned price-earnings ratio AESTAR de-trended price-earnings ratio LSTAR de-trended dividend yield
Nikkei	LSTAR de-trended dividend yield

Forecasting accuracy tests

Results of the standard forecasting accuracy tests on the basis of forecast error magnitude (Tables 4.24 – 4.26) suggest that the lowest, and thus the best, statistics are generated by the STAR-type models, in particular the asymmetric ESTAR (AESTAR) model. Whereas, the highest value of a trading rule test is produced consistently by the random walk model for all three holding periods.

The standard and modified Diebold-Mariano tests of equal forecasting accuracy for long-horizon data reveal significant statistical differences between MSEs of most random walk and STAR models for all time periods and for all series, with the exception of the Nikkei index, where all test statistics were found to be insignificant(Tables 4.27 – 4.29).

Forecast encompassing tests (Tables 4.30 – 4.32) were performed on the forecasts and forecasting errors of a random walk model, linear regressions with dividend yield and price-earnings ratio as determinants, and a STAR model for all indices over three long-horizon periods. The results suggest statistical significance of the linear models forecasts and forecasting errors for all series across all time periods, with the exception of the linear price-earnings ratio model for FTSE at three and twelve months holding period in a combination with AESTAR log dividend yield; both linear regressions for Nikkei at all holding periods, and the random walk model at three and six months periods. Moreover, the Nikkei index dividend yield linear regression is characterised by an insignificant coefficient for forecasting errors at six and twelve month periods; and insignificant coefficients for linear price-earnings regression for all time periods. These results suggest that for most of the series a random walk model and linear regressions contain independent information required in forecasting long-horizon returns series.

Non-linear models demonstrate significant results in the forecast encompassing tests for all series, especially for S&P and DAX, across all horizons. While AESTAR models seem to perform best for FTSE series, results for S&P and DAX series forecasts demonstrate consistent significant performance of AESTAR and LSTAR models, suggesting that these non-linear forecasts contain independent forecasting information in addition to linear alternatives. The forecast encompassing test results for Nikkei

index forecasts, on the other hand, demonstrate an improved performance of non-linear models over linear alternatives in terms of informational content.

Table 4.24. Forecasting accuracy tests results for a three month holding period.

3 month holding period	**ME**	**MAE**	**RMSE**	**Trade**
FTSE				
AESTAR log dividend yield	0.0001*	0.0013*	0.0021*	-0.0072
AESTAR de-meaned dividend yield	-0.0008	0.0040	0.0047	0.0019
AESTAR de-trended dividend yield	-0.0001**	0.0023**	0.0027**	0.0014
ESTAR de-meaned dividend yield	-0.0041	0.0043	0.0051	0.0024
ESTAR de-trended dividend yield	-0.0002	0.0033	0.0037	0.0024
LSTAR de-meaned dividend yield	-0.0072	0.0073	0.0085	0.0016
LSTAR de-trended dividend yield	-0.0020	0.0035	0.0041	0.0011
Random walk	0.0098	0.0103	0.0103	0.0028*
Linear dividend yield	0.0008	0.0039	0.0050	0.0019
Linear price-earnings ratio	0.0007	0.0054	0.0061	0.0019
S&P				
LSTAR de-trended price-earnings ratio	0.0127	0.0128**	0.0138*	0.0024
Random walk	0.0236	0.0263	0.0264	0.0095*
Linear dividend yield	0.0066*	0.0142	0.0163	0.0095*
Linear price-earnings ratio	0.0073**	0.0126*	0.0138**	0.0094
DAX				
AESTAR de-meaned price-earnings ratio	0.0015*	0.0029*	0.0035*	-0.0134
AESTAR de-trended price-earnings ratio	0.0016**	0.0029**	0.0035**	0.0128
LSTAR de-trended dividend yield	0.0016	0.0043	0.0052	0.0127
Random walk	0.0189	0.0217	0.0219	0.0131*
Linear dividend yield	0.0066	0.0142	0.0163	0.0131*
Linear price-earnings ratio	0.0150	0.0231	0.0242	0.0131*
NIKKEI				
LSTAR de-trended dividend yield	-0.0022*	0.0034*	0.0064*	0.0151*
Random walk	0.0122	0.0122	0.0125**	0.0151
Linear dividend yield	0.0055	0.0116**	0.0137	0.0151
Linear price-earnings ratio	0.0038**	0.0137	0.0156	0.0151
Note:	* indicates the best statistic ** indicates the second best statistic In the case of trade rule: * the largest positive value.			

Table 4.25. Forecasting accuracy tests results for a six month holding period.

6 month holding period	**ME**	**MAE**	**RMSE**	**Trade**
FTSE				
AESTAR log dividend yield	-0.0005	0.0038	0.0044*	0.0018
AESTAR de-meaned dividend yield	-0.0005	0.0038	0.0044**	0.0018
AESTAR de-trended dividend yield	0.0000*	0.0023*	0.0028	0.0016
ESTAR de-meaned dividend yield	-0.0035	0.0036	0.0043	0.0024
ESTAR de-trended dividend yield	-0.0003	0.0031**	0.0036	0.0007
LSTAR de-meaned dividend yield	-0.0025	0.0109	0.0163	0.0017
LSTAR de-trended dividend yield	-0.0008	0.0031	0.0036	0.0010
Random walk	0.0090	0.0104	0.0104	0.0028*
Linear dividend yield	0.0005	0.0038	0.0048	0.0019
Linear price-earnings ratio	0.0004**	0.0051	0.0057	0.0017
S&P				
LSTAR de-trended price-earnings ratio	0.0161	0.0161	0.0171	0.0024
Random walk	0.0238	0.0266	0.0266	0.0096*
Linear dividend yield	0.0010*	0.0070*	0.0083*	0.0070
Linear price-earnings ratio	0.0070**	0.0119**	0.0130**	0.0095
DAX				
AESTAR de-meaned price-earnings ratio	0.0013*	0.0030*	0.0035*	0.0129
AESTAR de-trended price-earnings ratio	0.0013**	0.0030**	0.0035**	0.0129
LSTAR de-trended dividend yield	-0.0162	0.0268	0.1579	0.0129
Random walk	0.0189	0.0219	0.0221	0.0132*
Linear dividend yield	0.0062	0.0130	0.0148	0.0132*
Linear price-earnings ratio	0.0149	0.0229	0.0240	0.0132*
NIKKEI				
LSTAR de-trended dividend yield	-0.0009*	0.0031*	0.0058*	0.0151*
Random walk	0.0124	0.0126	0.0129	0.0150
Linear dividend yield	0.0052	0.0108**	0.0128**	0.0150
Linear price-earnings ratio	0.0036**	0.0133	0.0151	0.0150
Note:	* indicates the best statistic ** indicates the second best statistic In the case of trade rule : * the largest positive value.			

Table 4.26 Forecasting accuracy tests results for a twelve month holding period.

12 month holding period	**ME**	**MAE**	**RMSE**	**Trade**
FTSE				
AESTAR log dividend yield	0.0004	0.0011*	0.0013*	0.0023
AESTAR de-meaned dividend yield	-0.0003	0.0037	0.0043	0.0018
AESTAR de-trended dividend yield	0.0003	0.0022**	0.0027**	0.0017
ESTAR de-meaned dividend yield	-0.0029	0.0030	0.0036	0.0023
ESTAR de-trended dividend yield	-0.0001**	0.0031	0.0035	0.0007
LSTAR de-meaned dividend yield	-0.0007	0.0115	0.0185	-0.0079
LSTAR de-trended dividend yield	-0.0003	0.0029	0.0033	0.0010
Random walk	0.0100	0.0106	0.0106	0.0029*
Linear dividend yield	0.0001	0.0037	0.0046	0.0018
Linear price-earnings ratio	-0.0000*	0.0046	0.0051	0.0015
S&P				
LSTAR de-trended price-earnings ratio	0.0157	0.0157	0.0168	0.0029
Random walk	0.0240	0.0269	0.0270	0.0098*
Linear dividend yield	0.0005*	0.0069*	0.0082*	0.0071
Linear price-earnings ratio	0.0065**	0.0113**	0.0124**	0.0096
DAX				
AESTAR de-meaned price-earnings ratio	0.0014*	0.0029**	0.0034*	0.0131
AESTAR de-trended price-earnings ratio	0.0014**	0.0029*	0.0035**	0.0131
LSTAR de-trended dividend yield	0.0047	0.0061	0.0073	0.0130
Random walk	0.0191	0.0221	0.0223	0.0135*
Linear dividend yield	0.0057	0.0111	0.0124	0.0135*
Linear price-earnings ratio	0.0148	0.0226	0.0237	0.0135*
NIKKEI				
LSTAR de-trended dividend yield	-0.0005*	0.0021*	0.0035*	0.0151*
Random walk	0.0127	0.0130	0.0133	0.0151
Linear dividend yield	0.0048	0.0095**	0.0112**	0.0151
Linear price-earnings ratio	0.0032**	0.0126	0.0144	0.0151
Note:	* indicates the best statistic ** indicates the second best statistic In the case of trade rule : * the largest positive value.			

Table 4.27. Diebold-Mariano test results for a three month holding period.

UK, 3 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.8445	1.8402
RW/ linear price-earnings ratio	1.9014	1.8970
RW/ ESTAR de-meaned dividend yield	1.7922	1.7880
RW/ ESTAR de-trended dividend yield	1.6428	1.6390
RW/ AESTAR log dividend yield	3.3273*	3.3196*
RW/ AESTAR de-meaned dividend yield	2.0124*	2.0077*
RW/ AESTAR de-trended dividend yield	3.3320*	3.3243*
RW/LSTAR de-meaned dividend yield	0.6680	0.6665
RW/ LSTAR de-trended dividend yield	1.9200	1.9155
Linear dy/ linear pe	0.2384	0.2378
Linear dy/ ESTAR de-meaned dividend yield	0.2072	0.2068
Linear dy/ ESTAR de-trended dividend yield	0.1414	0.1411
Linear dy/ AESTAR log dividend yield	0.6581	0.6565
Linear dy/ AESTAR de-meaned dividend yield	0.2864	0.2857
Linear dy/ AESTAR de-trended dividend yield	0.5340	0.5328
Linear dy/ LSTAR de-meaned dividend yield	-0.3196	-0.3189
Linear dy/ LSTAR de-trended dividend yield	0.2067	0.2062
Linear pe/ ESTAR de-meaned dividend yield	-0.0281	-0.0281
Linear pe/ ESTAR de-trended dividend yield	-0.1626	-0.1623
Linear pe/ AESTAR log dividend yield	0.6979	0.6962
Linear pe/ AESTAR de-meaned dividend yield	0.4048	0.4038
Linear pe/ AESTAR de-trended dividend yield	0.3684	0.3675
Linear pe/ LSTAR de-meaned dividend yield	-0.7139	-0.7123
Linear pe/ LSTAR de-trended dividend yield	-0.0614	-0.0612

S&P, 3 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.1851	1.1824
RW/ linear price-earnings ratio	2.7383*	2.7320*
RW/ LSTAR de-trended price-earnings ratio	2.3404*	2.3350*
Linear dy/ linear price-earnings ratio	0.395	0.3940
Linear dy/ LSTAR de-trended price-earnings ratio	0.2403	0.2397
Linear pe/ LSTAR de-trended price-earnings ratio	-0.3200	-0.3192

DAX, 3 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.0284	1.0260
RW/ linear price-earnings ratio	0.2171	0.2166
RW/ AESTAR de-meaned price-earnings ratio	3.6886*	3.6800*
RW/ AESTAR de-trended price-earnings ratio	3.6781*	3.6696*
RW/ LSTAR de-trended dividend yield	3.7282*	3.7196*
Linear dy/ linear price-earnings ratio	-0.9138	-0.9117
Linear dy/ AESTAR de-meaned pe ratio	0.5020	0.5008
Linear dy/ AESTAR de-trended pe ratio	0.5005	0.4993
Linear dy/ LSTAR de-trended dividend yield	0.4735	0.4724
Linear pe/ AESTAR de-meaned pe ratio	1.0219	1.0196
Linear pe/ AESTAR de-trended pe	1.0219	1.0196
Linear pe/ LSTAR de-trended dividend yield	0.9926	0.9903

Nikkei, 3 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	0.5291	0.5278
RW/ linear price-earnings ratio	0.3186	0.3179
RW/ LSTAR de-trended dividend yield	1.0143	1.0120
Linear dy/ linear pe	-0.5402	-0.5389
Linear dy/ LSTAR de-trended dividend yield	0.6289	0.6274
Linear pe/ LSTAR de-trended dividend yield	0.7585	0.7568
Note: * indicates statistical significance at 5%.		

Table 4.28. Diebold-Mariano test results for a six month holding period.

UK, 6 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.8720	1.8676*
RW/ linear price-earnings ratio	2.2043*	2.1991*
RW/ ESTAR de-meaned dividend yield	2.3251*	2.3197*
RW/ ESTAR de-trended dividend yield	2.0522*	2.0474*
RW/ AESTAR log dividend yield	3.9833*	3.9741*
RW/ AESTAR de-meaned dividend yield	2.3104*	2.3050*
RW/ AESTAR de-trended dividend yield	3.3909*	3.3831*
RW/LSTAR de-meaned dividend yield	0.7554	0.7537
RW/ LSTAR de-trended dividend yield	2.4664*	2.4607*
Linear dy/ linear pe	0.2706	0.2700
Linear dy/ ESTAR de-meaned dividend yield	0.3310	0.3302
Linear dy/ ESTAR de-trended dividend yield	0.2271	0.2266
Linear dy/ AESTAR log dividend yield	0.3211	0.3204
Linear dy/ AESTAR de-meaned dividend yield	0.3211	0.3204
Linear dy/ AESTAR de-trended dividend yield	0.5215	0.5202
Linear dy/ LSTAR de-meaned dividend yield	-0.2714	-0.2708
Linear dy/ LSTAR de-trended dividend yield	0.2739	0.2733
Linear pe/ ESTAR de-meaned dividend yield	0.1661	0.1657
Linear pe/ ESTAR de-trended dividend yield	-0.0771	-0.0769
Linear pe/ AESTAR log dividend yield	0.4000	0.3990
Linear pe/ AESTAR de-meaned dividend yield	0.3977	0.3968
Linear pe/ AESTAR de-trended dividend yield	0.4166	0.4157
Linear pe/ LSTAR de-meaned dividend yield	-0.6816	-0.6800
Linear pe/ LSTAR de-trended dividend yield	-0.0294	-0.0293

S&P, 6 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	6.0470*	6.0330*
RW/ linear price-earnings ratio	3.0503*	3.0432*
RW/ LSTAR de-trended price-earnings ratio	2.4560*	2.4503*
Linear dy/ linear price-earnings ratio	-0.2882	-0.2875
Linear dy/ LSTAR de-trended price-earnings ratio	-0.4942	-0.4931
Linear pe/ LSTAR de-trended price-earnings ratio	-0.4295	-0.4285

DAX, 6 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.2966	1.2936
RW/ linear price-earnings ratio	0.2341	0.2335
RW/ AESTAR de-meaned price-earnings ratio	3.6400*	3.6316*
RW/ AESTAR de-trended price-earnings ratio	3.6363*	3.6279*
RW/ LSTAR de-trended dividend yield	3.6477*	3.6393*
Linear dy/ linear price-earnings ratio	-0.9138	-0.9117
Linear dy/ AESTAR de-meaned pe ratio	0.5503	0.5490
Linear dy/ AESTAR de-trended pe ratio	0.5477	0.5464
Linear dy/ LSTAR de-trended dividend yield	0.5182	0.5170
Linear pe/ AESTAR de-meaned pe ratio	1.0072	1.0049
Linear pe/ AESTAR de-trended pe	1.0072	1.0049
Linear pe/ LSTAR de-trended dividend yield	0.9889	0.9866

Nikkei, 6 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	0.6606	0.6590
RW/ linear price-earnings ratio	0.3851	0.3842
RW/ LSTAR de-trended dividend yield	0.9863	0.9841
Linear dy/ linear pe	-0.6097	-0.6083
Linear dy/ LSTAR de-trended dividend yield	0.6026	0.6012
Linear pe/ LSTAR de-trended dividend yield	0.6929	0.6913
Note: * indicates statistical significance at 5%.		

Table 4.29. Diebold-Mariano test results for a twelve month holding period

UK, 12 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.8136	1.8093
RW/ linear price-earnings ratio	2.6322*	2.6261*
RW/ ESTAR de-meaned dividend yield	2.9805*	2.9736*
RW/ ESTAR de-trended dividend yield	2.0223*	2.0176*
RW/ AESTAR log dividend yield	4.4419*	4.4316*
RW/ AESTAR de-meaned dividend yield	2.6716*	2.6654*
RW/ AESTAR de-trended dividend yield	3.5992*	3.5908*
RW/LSTAR de-meaned dividend yield	0.9681	0.9659
RW/ LSTAR de-trended dividend yield	2.6867*	2.6804*
Linear dy/ linear pe	0.3364	0.3356
Linear dy/ ESTAR de-meaned dividend yield	0.4536	0.4526
Linear dy/ ESTAR de-trended dividend yield	0.2296	0.2291
Linear dy/ AESTAR log dividend yield	0.6733	0.6718
Linear dy/ AESTAR de-meaned dividend yield	0.3791	0.3782
Linear dy/ AESTAR de-trended dividend yield	0.5359	0.5347
Linear dy/ LSTAR de-meaned dividend yield	-0.1764	-0.1760
Linear dy/ LSTAR de-trended dividend yield	0.3027	0.3020
Linear pe/ ESTAR de-meaned dividend yield	0.2464	0.2459
Linear pe/ ESTAR de-trended dividend yield	-0.1130	-0.1127
Linear pe/ AESTAR log dividend yield	0.7087	0.7071
Linear pe/ AESTAR de-meaned dividend yield	0.3752	0.3743
Linear pe/ AESTAR de-trended dividend yield	0.4837	0.4826
Linear pe/ LSTAR de-meaned dividend yield	-0.5720	-0.5707
Linear pe/ LSTAR de-trended dividend yield	-0.0442	-0.0441

S&P, 12 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	6.6963*	6.6807*
RW/ linear price-earnings ratio	3.2307*	3.2232*
RW/ LSTAR de-trended price-earnings ratio	2.9141*	2.9073*
Linear dy/ linear price-earnings ratio	-0.2250	-0.2244
Linear dy/ LSTAR de-trended price-earnings ratio	-42.800*	-42.7008*
Linear pe/ LSTAR de-trended price-earnings ratio	-0.3756	-0.3747

DAX, 12 month holding period	*DM statistic*	*DM modified*
RW/ linear dividend yield	1.6666	1.6628
RW/ linear price-earnings ratio	0.2667	0.2661
RW/ AESTAR de-meaned price-earnings ratio	3.5242*	3.5161*
RW/ AESTAR de-trended price-earnings ratio	3.5242*	3.5161*
RW/ LSTAR de-trended dividend yield	3.4038*	3.3959*
Linear dy/ linear price-earnings ratio	-1.0149	-1.0125
Linear dy/ AESTAR de-meaned pe ratio	0.6396	0.6381
Linear dy/ AESTAR de-trended pe ratio	0.6358	0.6343
Linear dy/ LSTAR de-trended dividend yield	0.5923	0.5909
Linear pe/ AESTAR de-meaned pe ratio	0.9927	0.9904
Linear pe/ AESTAR de-trended pe	0.9890	0.9867
Linear pe/ LSTAR de-trended dividend yield	0.9776	0.9753

Nikkei, 12 month holding period	DM statistic	DM modified
RW/ linear dividend yield	0.9027	0.9006
RW/ linear price-earnings ratio	0.4941	0.4929
RW/ LSTAR de-trended dividend yield	1.5596	1.5560
Linear dy/ linear pe	-0.7081	-0.7065
Linear dy/ LSTAR de-trended dividend yield	0.5919	0.5905
Linear pe/ LSTAR de-trended dividend yield	0.7508	0.7491
Note: * indicates statistical significance at 5%.		

Table 4.30. Forecast encompassing test for a three month holding period.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
FTSE, 3 month holding period				
RW/ ESTAR de-meaned dividend yield	21.4447*	3.6684*	3.4209*	-6.5781
RW/ ESTAR de-trended dividend yield	20.4517*	3.5416*	2.5763*	-7.3832
RW/ AESTAR log dividend yield	31.0697*	8.4510*	9.4589*	0.2853
RW/ AESTAR de-meaned dividend yield	40.3024*	15.6497*	17.3163*	9.5077*
RW/ AESTAR de-trended dividend yield	23.8373*	1.8466	6.9826*	-5.5789
RW/ LSTAR de-meaned dividend yield	22.4105*	3.3577*	6.1550*	-3.7878
RW/ LSTAR de-trended dividend yield	15.7205*	-5.6840	4.8633*	-14.7307
Linear dy/ ESTAR de-meaned dividend yield	28.7347*	-16.8660	23.8247*	-21.3634
Linear dy/ ESTAR de-trended dividend yield	29.4429*	-18.3275	24.4527*	-23.0248
Linear dy/ AESTAR log dividend yield	29.3283*	-12.5341	28.4099*	-18.0151
Linear dy/ AESTAR de-meaned dividend yield	34.4606*	-16.9364	33.2304*	-19.2401
Linear dy/ AESTAR de-trended dividend yield	23.6860*	-9.6321	19.0467*	-13.7937
Linear dy/ LSTAR de-meaned dividend yield	24.0300*	-11.8437	19.8312*	-14.1212
Linear dy/ LSTAR de-trended dividend yield	16.3470*	-12.3400	12.2832*	-16.8797
Linear pe/ ESTAR de-meaned dividend yield	15.0532*	-19.9163	14.5767*	-24.2417
Linear pe/ ESTAR de-trended dividend yield	16.3117*	-22.1914	15.6152*	-26.5719
Linear pe/ AESTAR log dividend yield	0.75060	-0.4545	4.5297*	-7.5410
Linear pe/ AESTAR de-meaned dividend yield	-13.7904*	14.2894*	-13.3576*	12.9593*
Linear pe/ AESTAR de-trended dividend yield	7.2545*	-9.4826	7.6317*	-14.2367
Linear pe/ LSTAR de-meaned dividend yield	9.22387*	-13.1059	7.1769*	-13.4429
Linear pe/ LSTAR de-trended dividend yield	6.81167*	-18.8641	5.3846*	-22.6186
S&P, 3 month holding period				
RW/ LSTAR de-trended price-earnings ratio	7.8357*	4.4561*	-7.5815*	16.0446*
Linear dy/ LSTAR de-trended price-earnings ratio	8.6308*	31.5879*	2.3790*	12.4414*
Linear pe/ LSTAR de-trended price-earnings ratio	29.9012*	51.2135*	18.0172*	26.2280*
DAX, 3 month holding period				
RW/ AESTAR de-meaned price-earnings ratio	22.0262*	63.7839*	5.7380*	20.5242*
RW/ AESTAR de-trended price-earnings ratio	21.0902*	64.7524*	5.9527*	22.1843*
RW/ LSTAR de-trended dividend yield	36.0556*	38.7904*	-8.9642*	-3.7014
Linear dy/ AESTAR de-meaned pe ratio	-6.22457*	62.2631*	-7.9954*	28.8331*
Linear dy/ AESTAR de-trended pe ratio	-6.16879*	65.1334*	-7.9168*	30.5160*
Linear dy/ LSTAR de-trended dividend yield	-12.4466*	29.9361*	-10.6958*	12.4755*
Linear pe/ AESTAR de-meaned pe ratio	-19.3174*	100.2960*	-22.1308*	48.3467*
Linear pe/ AESTAR de-trended pe ratio	-18.3400*	101.4949*	-21.0942*	49.1424*
Linear pe/ LSTAR de-trended dividend yield	-11.4457*	29.9657*	-11.8742*	12.9938*
NIKKEI, 3 month holding period				
RW/ LSTAR de-trended dividend yield	1.7852	9.2241*	-19.9695*	-5.0492
Linear dy/ LSTAR de-trended dividend yield	-1.8830	7.3130*	-2.2343*	2.8523*
Linear pe/ LSTAR de-trended dividend yield	-0.7049	8.2190*	-0.5460	1.7830
Note :	* significant at 5%; RW – random walk; dy – dividend yield; pe – price-earnings.			

Table 4.31. Forecast encompassing test for a six month holding period.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
FTSE, 6 month holding period				
RW/ ESTAR de-meaned dividend yield	22.2342*	3.7998*	3.6369*	-6.4768
RW/ ESTAR de-trended dividend yield	16.2063*	-3.2009	4.0247*	-12.2514
RW/ AESTAR log dividend yield	39.8684*	15.6508*	17.5109*	9.7649*
RW/ AESTAR de-meaned dividend yield	39.7505*	15.5703*	17.2353*	9.6188*
RW/ AESTAR de-trended dividend yield	24.8001*	2.7911*	7.3192*	-4.3960
RW/ LSTAR de-meaned dividend yield	21.8398*	3.0435*	5.7019*	-4.3045
RW/ LSTAR de-trended dividend yield	15.2724*	-6.7206	4.8082*	-16.4615
Linear dy/ ESTAR de-meaned dividend yield	27.8023*	-14.1554	22.5885*	-18.6677
Linear dy/ ESTAR de-trended dividend yield	18.3909*	-11.1889	13.9253*	-15.2541
Linear dy/ AESTAR log dividend yield	33.0973*	-15.0027	32.0888*	-17.2022
Linear dy/ AESTAR de-meaned dividend yield	33.1069*	-15.0047	32.0912*	-17.2040
Linear dy/ AESTAR de-trended dividend yield	25.4880*	-9.3721	20.8540*	-13.2211
Linear dy/ LSTAR de-meaned dividend yield	24.7650*	-11.6003	20.3543*	-13.9173
Linear dy/ LSTAR de-trended dividend yield	14.6761*	-10.5727	10.3383*	-15.4868
Linear pe/ ESTAR de-meaned dividend yield	12.8157*	-16.7817	12.7670*	-21.8794
Linear pe/ ESTAR de-trended dividend yield	10.0098*	-19.6267	8.5520*	-22.9769
Linear pe/ AESTAR log dividend yield	-15.2823	15.9981*	-14.8242*	14.4398*
Linear pe/ AESTAR de-meaned dividend yield	-15.0042	15.7140*	-14.5604*	14.1762*
Linear pe/ AESTAR de-trended dividend yield	6.9830*	-8.4606	7.6365*	-13.0551
Linear pe/ LSTAR de-meaned dividend yield	9.9264*	-13.9869	7.8304*	-14.3826
Linear pe/ LSTAR de-trended dividend yield	6.2704*	-19.5185	4.8628*	-24.0491
S&P, 6 month holding period				
RW/ LSTAR de-trended price-earnings ratio	7.3357*	5.5356*	-7.5225*	16.6420*
Linear dy/ LSTAR de-trended price-earnings ratio	23.6058*	34.1870*	18.5604*	22.1656*
Linear pe/ LSTAR de-trended price-earnings ratio	29.2106*	51.8031*	18.0827*	26.9989*
DAX, 6 month holding period				
RW/ AESTAR de-meaned price-earnings ratio	20.2291*	60.7634*	4.4671*	18.8852*
RW/ AESTAR de-trended price-earnings ratio	19.4477*	61.9013*	4.7312*	20.4825*
RW/ LSTAR de-trended dividend yield	35.5893*	39.7696*	-9.5794	-4.04672
Linear dy/ AESTAR de-meaned pe ratio	-5.7620*	59.5773*	-7.5924*	28.2812*
Linear dy/ AESTAR de-trended pe ratio	-5.7605*	62.3681*	-7.5650*	29.9759*
Linear dy/ LSTAR de-trended dividend yield	-11.8435*	28.8922*	-10.4480*	12.4541*
Linear pe/ AESTAR de-meaned pe ratio	-18.8853*	98.6283*	-21.6262*	47.3911*
Linear pe/ AESTAR de-trended pe ratio	-17.9988*	99.9993*	-20.6914*	48.2963*
Linear pe/ LSTAR de-trended dividend yield	-13.2009*	32.8941*	-13.1978*	14.4983*
NIKKEI, 6 month holding period				
RW/ LSTAR de-trended dividend yield	1.2380	7.2703*	-21.1985*	-6.2539
Linear dy/ LSTAR de-trended dividend yield	-0.1536	4.9730*	-1.5427	1.0267
Linear pe/ LSTAR de-trended dividend yield	-0.0328	6.3715*	-0.1393	-0.0553
Note :	* significant at 5%; RW – random walk; dy – dividend yield; pe – price-earnings.			

Table 4.32. Forecast encompassing test for a twelve month holding period.

	Forecasting encompassing		**Forecasting errors encompassing**	
	t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
FTSE, 12 month holding period				
RW/ ESTAR de-meaned dividend yield	22.0040*	3.4140*	3.7128*	-6.6043
RW/ ESTAR de-trended dividend yield	15.6190*	-3.3469	3.6542*	-12.4904
RW/ AESTAR log dividend yield	30.9257*	9.5043*	7.4963*	-0.8863
RW/ AESTAR de-meaned dividend yield	37.9346*	15.3303*	16.1557*	9.3017*
RW/ AESTAR de-trended dividend yield	25.3393*	3.9243*	7.5516*	-2.9869
RW/ LSTAR de-meaned dividend yield	19.7866*	1.7599*	4.6998*	-5.8168
RW/ LSTAR de-trended dividend yield	14.5415*	-5.9115	3.8770*	-16.0272
Linear dy/ ESTAR de-meaned dividend yield	27.1283*	-10.4976	21.2830*	-14.6224
Linear dy/ ESTAR de-trended dividend yield	19.5538*	-9.8390	14.6450*	-13.5880
Linear dy/ AESTAR log dividend yield	28.2820*	-8.3032	26.8400*	-14.1830
Linear dy/ AESTAR de-meaned dividend yield	32.7294*	-12.7069	31.7213*	-14.7288
Linear dy/ AESTAR de-trended dividend yield	27.6981*	-8.0654	23.0998*	-11.4729
Linear dy/ LSTAR de-meaned dividend yield	25.3347*	-10.4724	20.5357*	-12.8082
Linear dy/ LSTAR de-trended dividend yield	16.0780*	-8.9715	11.1677*	-13.5533
Linear pe/ ESTAR de-meaned dividend yield	10.3755*	-13.4042	10.2606*	-18.7459
Linear pe/ ESTAR de-trended dividend yield	10.6121*	-19.8314	8.9945*	-23.3013
Linear pe/ AESTAR log dividend yield	0.2298	0.6045	7.8501*	-11.1803
Linear pe/ AESTAR de-meaned dividend yield	-12.1355	12.9494*	-11.6421*	11.4258*
Linear pe/ AESTAR de-trended dividend yield	6.1478*	-6.5309	7.1944*	-11.1519
Linear pe/ LSTAR de-meaned dividend yield	11.0671*	-15.5469	8.8883*	-16.2980
Linear pe/ LSTAR de-trended dividend yield	7.4057*	-19.5067	5.9024*	-24.3077
S&P, 12 month holding period				
RW/ LSTAR de-trended price-earnings ratio	8.5906*	5.1625*	-6.7776*	13.7280*
Linear dy/ LSTAR de-trended price-earnings ratio	21.2295*	28.1110*	16.4845*	17.2084*
Linear pe/ LSTAR de-trended price-earnings ratio	26.5920*	44.1517*	15.3737*	21.3363*
DAX, 12 month holding period				
RW/ AESTAR de-meaned price-earnings ratio	16.5358*	57.9356*	2.6347*	17.6765*
RW/ AESTAR de-trended price-earnings ratio	16.2452*	61.3289*	3.3402*	20.0362*
RW/ LSTAR de-trended dividend yield	29.1100*	38.9549*	-8.0891*	-1.7836
Linear dy/ AESTAR de-meaned pe ratio	-5.1076*	56.8411*	-6.7979*	28.4178*
Linear dy/ AESTAR de-trended pe ratio	-5.3338*	61.1570*	-6.9824*	31.1575*
Linear dy/ LSTAR de-trended dividend yield	-7.5867*	25.5573*	-7.7694*	11.3989*
Linear pe/ AESTAR de-meaned pe ratio	-14.5169*	86.9323*	-16.8217*	41.1735*
Linear pe/ AESTAR de-trended pe ratio	-15.0160*	94.5118*	-17.3269*	45.3590*
Linear pe/ LSTAR de-trended dividend yield	-12.6846*	35.9116*	-13.1555*	15.9334*
NIKKEI, 12 month holding period				
RW/ LSTAR de-trended dividend yield	3.0636*	14.1445*	-22.6537*	-6.3218
Linear dy/ LSTAR de-trended dividend yield	-1.4011	11.0539*	-1.1947	0.1453
Linear pe/ LSTAR de-trended dividend yield	-1.6253	12.8942*	0.1052	-0.6939
Note :	* significant at 5%; RW – random walk; dy – dividend yield; pe – price-earnings.			

Combined forecast

On the basis of forecast encompassing test results, the most successful, in terms of the information content across all long-horizon period, the single non-linear model is chosen for each of the indices in order to incorporate these models into a combined forecast. The combination will include a random walk model to account for a random component of the stock price time-series and a preferred STAR model. Hence, the following STAR models were chosen for each series across three long-horizon holding periods.

Table 4.33. Combined forecast models.

Series	**Non-linear model**
FTSE	AESTAR de-meaned dividend yield
S&P	LSTAR de-trended price-earnings ratio
DAX	AESTAR de-meaned price-earnings ratio
Nikkei	LSTAR de-trended dividend yield

In addition, the forecast encompassing test for the DAX series revealed two non-linear models to have the same informational content, thus the final decision was based on the results of the standard error magnitude tests.

The equal weighting method of forecast combination takes the following form:

$$f_c(y) = \frac{1}{k}\sum_{i=1}^{k} f_i \tag{4.31}$$

where $f_c(y)$ is the combined forecast of individual forecasts $f_1, f_2 \ldots f_k$, and k is the number of forecasts. Similarly, the combined forecast methodology employed in this section will apply the following regression:

$$y_{t+s} = \beta_0 f_{t,s}^{RW} + (1 - \beta_0) f_{t,s}^{STAR} + u_t \tag{4.32}$$

where $f_{t,s}^{RW}$ is the random walk model forecast, $f_{t,s}^{STAR}$ is the STAR model forecast and u_t is the error term.

Figures below (Figure 4.9 – 4.12) represent plots of the actual and fitted values, along with the residuals, of the combined twelve month models for each of the four time-series. All the combined models seem to fit the data rather well, with the Nikkei index having the smallest variation in the residuals throughout with the exception of extreme values at the end of the sample.

Figure 4.9. FTSE combined model, 12 month holding period.

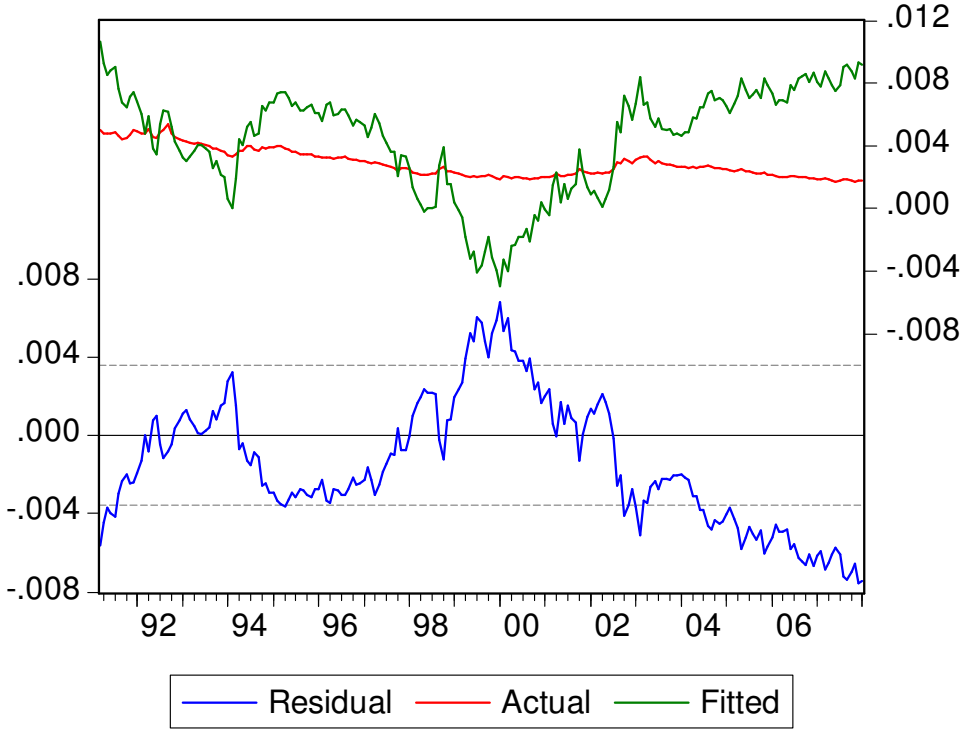


Figure 4.10. S&P combined model, 12 month holding period.

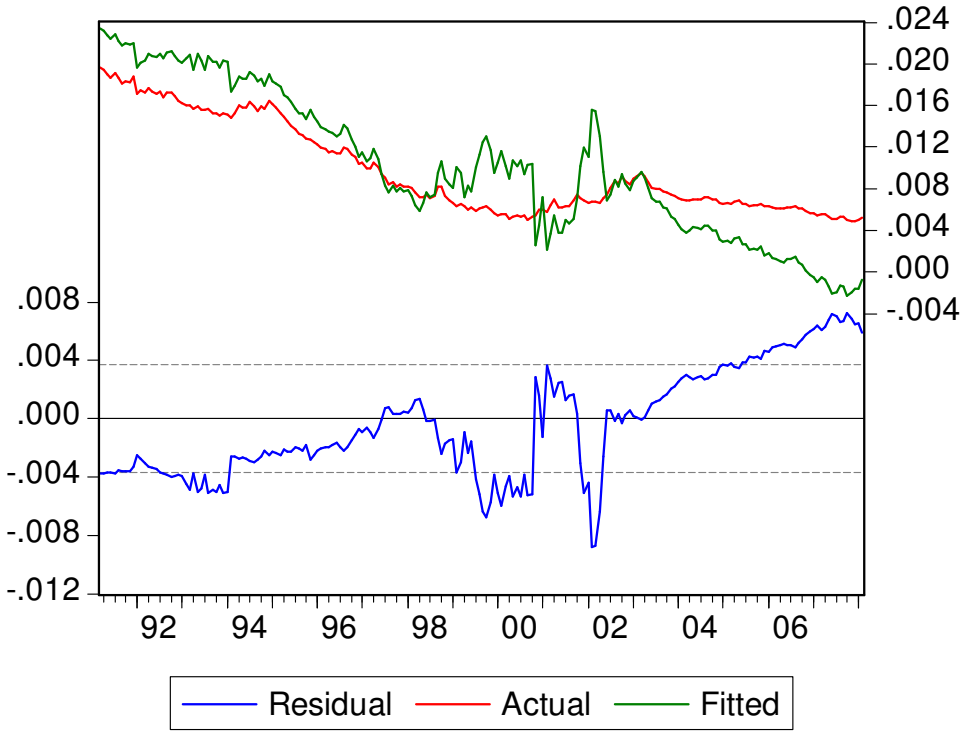


Figure 4.11. DAX combined model, 12 month holding period

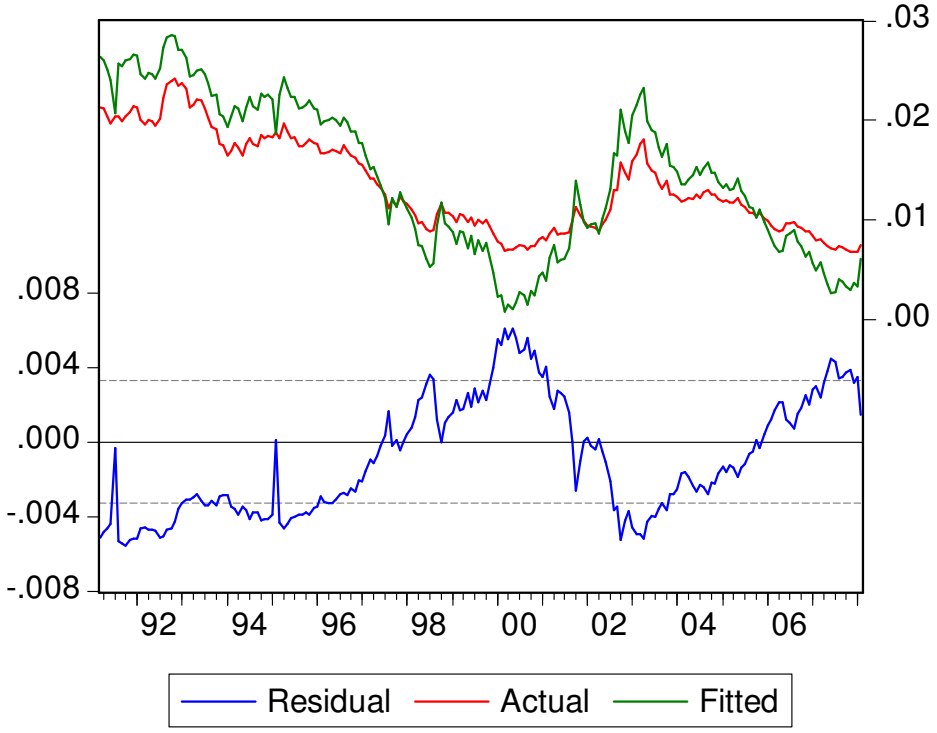
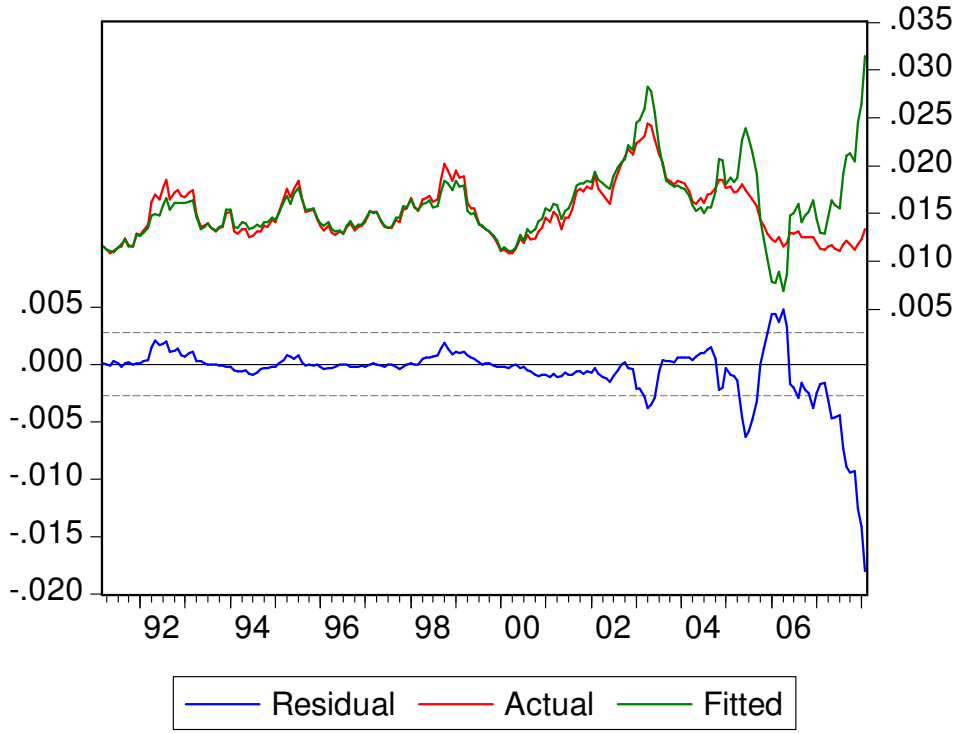


Figure 4.12. Nikkei combined model, 12 month holding period.



The results of ME, MAE, RMSE and trade rule tests of forecasting accuracy of combined forecasts for each long-horizon period are presented in the table below (Table 4.34)

Table 4.34. Combined forecast accuracy tests.

	ME	**MAE**	**RMSE**	**Trade**
FTSE – AESTAR de-meaned dy				
3 month holding period	-0.001769	0.003163	0.003869	0.002373
6 month holding period	-0.001620	0.003104	0.003722	0.002341
12 month holding period	-0.001462*	0.003003*	0.003560*	0.002383*
S&P – LSTAR de-trended pe				
3 month holding period	-0.000493	0.003574	0.004086	0.008284
6 month holding period	-0.000458	0.003405	0.003908	0.008508
12 month holding period	-0.000408*	0.003150*	0.003688*	0.009013*
DAX – AESTAR de-meaned pe				
3 month holding period	-0.000910	0.002912	0.003294	0.013147
6 month holding period	-0.000916	0.002908	0.003297	0.013238
12 month holding period	-0.000906*	0.002880*	0.003261*	0.013482*
Nikkei – LSTAR de-trended dy				
3 month holding period	-0.001927	0.002116	0.004402	0.015139*
6 month holding period	-0.001975	0.002158	0.004647	0.015102
12 month holding period	-0.000697*	0.001366*	0.002710*	0.015130
Note: * indicates the best statistic dy – dividend yield pe – price-earnings				

According to the results of the forecasting accuracy tests, the longest holding period of twelve months seems to produce the best forecast for all series, as well as generating the

highest value of the trade rule test, with the exception of the Nikkei index where the highest trade rule value is produced by the three month period combination forecast.

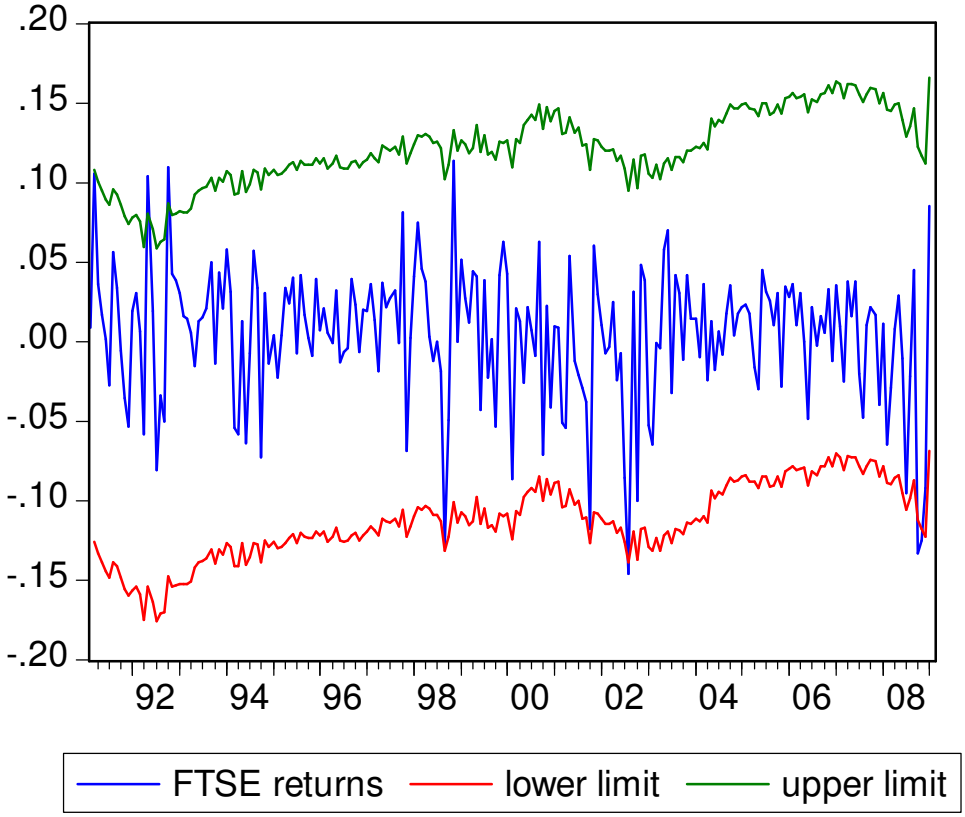
Overall, the combination of a random walk model and a STAR model over the long-horizon period of three, six and twelve months seem to produce better statistics and trade rule results compared to all the linear and non-linear forecasts considered in this chapter, including forecasts of monthly returns (Section 4.4), and long-horizon individual forecasts. Moreover, the twelve month holding period combined forecasts performs best out of all the long-horizon forecast combinations, and hence can be suggested as a superior forecast model.

4.6. Interval forecasts for monthly data[7]

The methodology for carrying out an out-of-sample interval forecast for monthly data is similar to one used on daily data in Chapter 2 and based on the technique suggested by Christoffersen (1998). The interval prediction barriers are set as upper and lower limits with certain probability and level of confidence. This study will apply standard distribution *t*-statistic value at 95% level of confidence. Figure 4.13 illustrates the interval forecast upper and lower barriers on the example of FTSE returns series.

[7] The main objective of this thesis is an investigation of point forecasting with non-linear models and does not include a thorough examination into interval forecasts. The subject of interval forecasts is an important area of time-series research that lacks extensive empirical examination in the literature. I would like to thank my examiners for their valuable comments and recommendations for further research within the field of forecasting.

Figure 4.13.Interval forecast of FTSE AESTAR log dividend yield.



The goodness of fit of the forecast is tested using a success ratio of the indicator variable, I_t, which determines how accurate the interval forecast values are.

$$I_t = \begin{cases} 1, & if \;\; y_t \in \left[L_{t|t-1}(p), U_{t|t-1}(p)\right] \\ 0, & if \;\; y_t \notin \left[L_{t|t-1}(p), U_{t|t-1}(p)\right] \end{cases} \tag{4.33}$$

Where, $L_{t|t-1}(p)$ and $U_{t|t-1}(p)$ are lower and upper limits respectively, for a given interval forecast, $\left(L_{t|t-1}(p), U_{t|t-1}(p)\right)$ for time *t*, made at time *t-1*, with the coverage probability, p, for a time-series of a random variable, y_t. Thus, zero value is assigned to every forecasted value outside the prediction barriers, whereas unity value is assigned to every forecast within the range of the set interval.

Table 4.35 contains the success ratio of linear and non-linear forecasts estimated for monthly data considered in this chapter. The results suggest that most interval forecasts performed in this section surpassed the limit required by the 95% confidence interval. Overall, the general forecasting performance is higher compared to the results of the daily interval forecast results. These results confirm the suggestion that while daily data is characterised by a large number of extreme values described by the fat tails of the normal distribution, monthly data is less noisy and characterised by well defined trends and patters. Moreover, similar to the results of the point forecasts of monthly returns carried out in this chapter, interval forecast results suggest stronger statistics for non-linear models in the form of the success ratio compared to the linear benchmarks. These results confirm the suggestion that non-linear models demonstrate superiority to linear models in producing out-of-sample forecasts for the long-horizon data.

Table 4.35. Interval forecast success ratio results.

	Success ratio
FTSE	
AESTAR log dividend yield	0.9767*
AESTAR de-meaned dividend yield	0.9907*
AESTAR de-trended dividend yield	0.9861*
ESTAR de-meaned dividend yield	0.9814*
ESTAR de-trended dividend yield	0.9722*
LSTAR de-meaned dividend yield	0.9814*
LSTAR de-trended dividend yield	0.9722*
Random walk	0.9768*
Linear dividend yield	0.9722*
Linear price-earnings ratio	0.9629*
S&P	
LSTAR de-trended price-earnings ratio	0.9585*
Random walk	0.9400
Linear dividend yield	0.9400
Linear price-earnings ratio	0.9400
DAX	
AESTAR de-meaned price-earnings ratio	0.9814*
AESTAR de-trended price-earnings ratio	0.9814*
LSTAR de-trended dividend yield	0.9585*
Random walk	0.9308
Linear dividend yield	0.9308
Linear price-earnings ratio	0.9262
Nikkei	
LSTAR de-trended dividend yield	0.9493
Random walk	0.9354
Linear dividend yield	0.9354
Linear price-earnings ratio	0.9354
Note: * indicated statistical significance at 95% level of confidence	

4.7. Conclusion

This chapter has concentrated on the subject of long-horizon predictability of stock returns using the dividend-price ratio, or dividend yield, and price-earnings ratio. Based on the idea of the seminal research by Campbell and Shiller (1987) further research into this topic included using the relationship between stock returns and dividend yield in econometric forecasting by applying non-linear models (e.g. McMillan and Speight, 2006; McMillan, 2007). As a result, this study applied an error-correction methodology using non-linear STAR-type models in order to carry out a forecasting exercise of monthly price returns. In addition, the non-linear forecasts were compared to linear benchmarks in the form of random walk and simple linear regression models.

This chapter considered time-series data of monthly financial indices of four major economies including the FTSE All-Share of the UK, S&P index of the US, German DAX30 Performance and Japanese Nikkei 225 Stock Average index. The data obtained covered a period of thirty six years from January 1973 to February 2009 and included time-series of dividend yields and price-earnings ratio for each index. Descriptive statistics revealed that monthly data, expectedly, is less volatile in comparison to the daily data in Chapter 3, and naturally shares similar patterns observed with the daily frequency data. All data series were additionally de-meaned and de-trended in order to centre the long-run equilibrium around zero. The ADF unit root test confirmed linear stationarity for all price returns, FTSE dividend yield and de-trended dividend yield, as well as price-earnings ratio, log price-earnings ratio, de-meaned, and de-trended price-earnings ratio for the DAX series. Non-linear unit root tests confirmed ESTAR and LSTAR non-stationarity for FTSE and DAX series and LSTAR-type stationarity for

S&P and Nikkei indices. A random walk model, and dividend yield and price-earnings linear regressions, as well as the appropriate STAR-type models, were estimated for all stationary time-series. The forecasting process consisted of an out-of-sample one-step ahead error-correction model procedure. Forecasting accuracy testing included ME, MAE, RMSE and trading rule style tests, as well as Diebold-Mariano tests of equal forecasting accuracy and the forecasting encompassing testing procedure. Furthermore, combination forecasts for each series included random walk, linear regression and STAR models, which in turn were assesses by the same forecasting accuracy tests.

From the empirical results obtained in the first part of the empirical chapter, Section 4.4, it is apparent that while non-linear STAR models have slight advantage in terms of forecasting accuracy, which, whilst providing an appealing topic for an academic purpose, might seem as only a marginal superiority in terms of practical applications. However, evidence obtained in this chapter demonstrates the presence of stock returns predictability through the dividend yield and price-earnings ratio with no clear preference for either one of these variables. Moreover, similar to the empirical findings by numerous researchers who found patterns of non-linear mean reversion and STAR models to provide an adequate fit for the data (Kanas, 2005; Rapach et al., 2005; Bali et al., 2008), this paper also confirms capability of the STAR-type models to generate sufficiently accurate out-of-sample forecasts.

The second empirical part of the chapter, Section 4.5, reviewed the non-linear error-correction methodology in the context of long-horizon forecasting by applying a buy-and-hold strategy for periods of three, six and twelve months. The results of the forecasting exercise seem to be similar to the suggestions by Fair and Shiller (1990) and Montgomery et al. (1998), in the sense that the long-horizon data displays smoother

trends and produces better forecasts. While the results obtained in Section 4.5 seem to confirm the findings by Montgomery et al. (1998) of long-horizon data to produce better forecasts, the overall investigation extends previous work by considering an out-of-sample forecasting exercise, as opposed to in-sample predictions. Thus, the recursive one-step out-of-sample forecasting error-correction methodology is applied to long-horizon stock prices data using a random walk model, linear regression with dividend yield and price-earnings ratio as determinants, and STAR-type models. The forecasts are then assessed using forecasting tests of forecast error magnitude, tests of equal forecasting accuracy and forecast encompassing tests. While all models seemed to produce reasonable adequate forecasts, STAR models proved to perform better comparing to monthly forecast results, with the asymmetric ESTAR model in particular being favoured for the FTSE index. However, the most considerable improvement in forecasting ability of the non-linear models followed a combined forecast approach. A combination of a random walk model and a STAR model for each series across three, six and twelve months of the holding period were assessed using the same forecasting accuracy testing procedures to reveal significant improvement over monthly forecasts obtained in Section 4.4 in terms of forecasting accuracy. Moreover, the notion of long-horizon forecasts performing better comparing to higher frequency forecasts is confirmed by the twelve months combination forecast which appears to generate the best overall forecast for all four price series.

The combination of a random walk model and a STAR-type model appears to be the best choice in terms of forecasting performance. This phenomenon could be explained due to the STAR model capturing the cyclical nature of the stock market characterised by asymmetric adjustments, while the random walk model accounts for periods of calm

in the financial market when the tranquil state is best described by a random process instead of deterministic trend. These results would seem to be most appropriate for market participants and policy-makers concerned with long-horizon predictions of the financial market.

Chapter 5
House price returns forecasting

5.1. Introduction

This chapter investigates the application of the present value model approach to forecasting of house prices using price-income ratio with STAR-type models. The previous chapter, Chapter 4, applied the non-linear error-correction approach to investigate long-horizon predictability of stock returns using the dividend yield and price-earnings ratio. Cointegration methodology has been used by numerous studies (Campbell and Shiller, 1987, 1988a; Evans, 1991; Enders and Siklos, 2001; Bohl, 2003; Brooks and Katsaris, 2003; Bohl and Siklos, 2004; Psaradakis et al., 2004; Kanas, 2005; Cochrane, 2008) in order to examine the presence of a long-run relationship between stock prices and dividends and hence to test the validity of the present value model. The current study is concerned with predictability of financial assets and forecasting applications of non-linear models. Hence, in order to extend an investigation further this chapter intends to apply non-linear methodology and present value model procedure to forecasting house prices. This chapter will discuss the application of the above stock market situation to a housing market. In other words, it is an attempt to apply modelling and analysis of the relationship between stock prices and dividends to the possible relationship between real house prices and real income. It can be noted that behaviours of both markets are similar in nature, including the presence of bubbles, for instance.

However, there are some differences, mainly the time frame, as the housing market is less adjustable than the stock market where trade is taking place on an every-minute basis. This particular difference between the markets will also be reflected in forecasting horizons.

The empirical version of the present value model was introduced in the seminal paper by Campbell and Shiller (1987). The model states that current stock prices are discounted values of future dividends where the discount rate is the required rate of return. According to Campbell and Shiller (1987), if the present value model holds, stock prices and dividends should cointegrate. This relationship between stock prices and dividends is also implied by the efficient market hypothesis (EMH), however, in such way that return stock predictability can be interpreted as an indication of market inefficiency. The relationship between stock prices and dividends can be examined using a cointegration approach where prices and dividends cointegrate assuming either constant discount rate or allowing for a time-varying discount rate. The unexpected significant rise in stock prices and subsequent fall in late the 1990s and early 2000s have raised new interest in the present value model and a re-examination of relationships between stock prices and dividends, as the simple constant discount rate model did not seem to hold. Nasseh and Strauss (2004) have suggested that low dividend payouts combined with record-high stock prices are an indication of stock price overreaction. They have pointed out that stocks were overvalued by 43% during the late 1990s, and further suggested that such stock price overvaluation can be explained by a break in dividend payments in the mid 1990s. It has been pointed out that failure of the present value model can be explained due to the presence of constant discount rate and rational expectations. However, allowing for a time-varying, as

opposed to constant, discount rate and including a component designed to capture speculative bubbles into the model has resulted in inconclusive empirical findings. Furthermore, Kanas (2005) also suggested non-linearities in the relationship between stock prices and dividends as a possible reason for the present value model to fail. Kanas (2005) attempted to use non-linear extensions of the present value model in order to investigate whether such models are superior in explaining stock prices as a function of dividends. The study had investigated monthly real stock index prices and real dividends for the UK, US, Japan and Germany and found significant evidence of the presence of non-linear cointegration for all considered countries. Kanas (2005) suggests that application of a non-linear approach improves the present value model in its ability to explain the relationship between the stock prices and dividends. Suggestion of the presence of non-linearities in the stock market and non-linear adjustment dynamics within mean reversion relationship between the stock prices and dividends encouraged research into possible explanations of these dynamics. The presence of non-linearities in the stock market has been attributed to the presence of bubbles (Evans, 1991; McMillan, 2001; Bohl, 2003; Brooks and Katsaris, 2003; Kilian and Taylor, 2003; Psaradakis et al., 2004), transaction costs (Kilian and Taylor, 2003; McMillan, 2005; Bali et al., 2008), and interaction between traders (McMillan, 2003; McMillan, 2005; McMillan and Speight, 2006). However, empirical evidence and studies of a long-run equilibrium relationship, or cointegration, between stock prices and dividends are mixed and unclear. In addition, McMillan (2005) points out that most of these studies are focused on US data.

Black et al. (2006) claim that housing markets received rather limited attention in academic literature compared to financial markets. However, despite the main attention

being focused on the financial market, there have been a number of studies to address the issue of housing markets. Many researchers highlighted the importance of the effects of housing prices on the whole economy and hence the importance of understanding and predicting house price dynamics (Case and Shiller, 1989; Brown et al., 1997; Muellbauer and Murphy, 1997; Crawford and Fratatoni, 2003; Turner, 2003; Fraser et al., 2008; Miles, 2008; Koetter and Poghosyan, 2009). Case et al. (2001) found changes in housing market wealth to have stronger effect on consumption compared to changes in the stock market wealth. The findings were sustained throughout the data from fourteen countries as well as a panel of US states, and found to be robust for different model specifications. Moreover, Koetter and Poghosyan (2009) point out that the policy makers do take into account property prices as being one of the indicators of the financial market's susceptibility since imbalances in the housing market can lead to instability in the financial sector due to banks acting as mortgage lenders. Consequently, while an increase in house prices might increase the value of real estate in the bank's possession and thus improve bank capital, and decrease the probability of mortgage borrowers defaulting on appreciated assets, the same house price increase and consequent lower perceived risk might also bring instability to banks by encouraging lendings to higher risk real estate at a lower interest rate.

Despite Case and Shiller (2004) finding that homeowners treat their properties as an investment, Black et al. (2006) suggest that most house purchases are motivated by a consumption rather than an investment decision. High transaction costs and legal regulations prevent professional speculators to freely trade on the housing market. Thus, while the financial market is characterised with a large number of professional traders and market arbitragers, the real estate is not associated with professional speculators due

to limits to arbitrage and properties on the housing market being mostly bought and sold by private homeowners. Moreover, limits to arbitrage and an inelastic supply of housing contribute to slow mean reversion and thus long periods of adjustments of mispricing.

Similarly to the financial market situation, bursts and booms observed in the housing market have encouraged investigations into house prices deviating from their fundamental values. The present value model approach was applied to mean-reverting house prices. Studies have found a number of variables to determine house prices including income levels, construction costs and elasticity of supply (Case and Mayer, 1996; Case and Shiller, 2004; Black et al., 2005). In addition, demographic trends, shifts in employment and financial regulations have been shown to affect levels of house prices (Case and Mayer, 1996). Furthermore, researchers have recognised the presence of non-linearities and asymmetries in house prices and thus proposed possible explanations to these dynamics including presence of bubbles and transaction costs (Hall et al., 1997; Holly and Jones, 1997; Crawford and Fratantoni, 2003; Case and Shiller, 2004; Coleman et al., 2008; Goodman and Thibideau, 2008; Miles, 2008).

The rest of this chapter is organised as follows. Section 5.2 reviews appropriate literature on the topic of applications of the present value model in forecasting and issues concerning the housing market; Section 5.3 introduces the methodology used to investigate the relationship between real house prices and real income, with empirical results presented in Section 5.4, which will concentrate on outlining and comparing forecasts drawn from the estimated models. Section 5.5 will conclude the chapter.

5.2. Literature review

Introduction to the housing market

The UK housing market has increased dramatically since the Second World War and has one of the largest owner-occupier rates in the world (Brown et al., 1997). Holly and Jones (1997) explain the excess demand for housing in the post-war period due to extensive decline in supply as a result of aerial bombing. In addition, less strict post-war lending policies appear to have encouraged the demand. Similarly, in much later periods, Brown et al. (1997) identified financial market deregulation and removal of mortgage constraints as main reasons for the housing market experiencing structural changes which brought about considerable price rises above consumer disposable incomes in the early 1970s, early 1980s and late 1980s. Hence, financial deregulation and availability of mortgage in the early 1980s encouraged a rise in the demand for housing. However, further monetary policies resulted in the rise of the mortgage interest rates during the 1990s, which weakened the housing market as a result of economic recession (Pain and Westaway, 1997). Thus, Goodman and Thibodeau (2008) attribute the fall of house ownership in 1980 in the US due to a rise in real interest rates. Muellbauer and Murphy (1997) suggested income growth in the early 1980s and income growth expectations, as well as the financial liberalisation of the 1980s, amongst the factors that contributed towards the UK house price boom in the late 1980s. Whereas, the subsequent burst in the 1990s was accompanied by weakened income growth and growth expectations, reversal of demographic trends, reintroduction of a property tax and stricter mortgage lending criteria. Muellbauer and Murphy (1997), therefore, point out the importance of understanding the UK housing market to the

government and policy makers as it can be a potential factor in causing macroeconomic instability. For further in-depth examination Case and Shiller (2004) offer an extensive review of historical patterns of housing prices in regional data for the US during periods from the 1980s to 2002.

In the early models the supply and demand based approach was not sufficient to successfully forecast such rapid changes in housing prices movements. Hence, these substantial changes in the housing market initiated new research interest in modelling and forecasting of the housing market. In attempts to improve the house prices modelling an asset market approach proved to be the most promising.

Some researchers approached the topic of housing market and patterns in house prices in terms of the so called standard urban model, which is a regression model of house prices against a set of locational attributes and amenities. Thus, Case and Mayer (1996) investigated the appreciation of housing in the Boston metropolitan area in terms of the effects local amenities such as employment, demographics, rate of new construction, as well as location and quality of schools, have on patterns in house prices. The study used a simple model of price determination with different sets of locational characteristics across series of different jurisdictions within the area while making an assumption of a fixed household income. Case and Mayer (1996) found that while shifts in employment and demographics have a significant effect on the housing market, it is, nonetheless, very slow to adjust and respond to such changes. In addition, Abraham and Hendershott (1996) suggest price trends to be localised phenomena. The researchers did not explore the issue of supply restrictions. The effects of supply restrictions might manifest itself in enhancing the localised phenomena described by Abraham and Hendershott (1996).

Furthermore, to improve modelling methodologies research studies concentrated on investigating determinants of the house prices and correct identification of the fundamental values while allowing for a time-varying approach. In addition, house prices are characterised by strong autocorrelation patterns and mean reversion (Case and Shiller, 1989; Gillen et al., 2001; Gu, 2002; Capozza et al., 2004; Zandi and Chen, 2006). Gao et al. (2009) demonstrated the presence of positive correlation between fundamental house prices and household income, and negative correlation between house prices and mortgage costs. Moreover, Gao et al. (2009) did not consider forecasting exercise, however the research found that the longer housing market remains overvalued or undervalued, the stronger is the likelihood of reversion to the equilibrium. Mikhed and Zemčík (2009) found construction costs and income to be the main determinants of house prices. Mikhed and Zemčík (2009) based their investigation of the present value model approach and comprised a number of economic variables to form fundamental price levels including real house rent, mortgage rate, personal income, building costs, stock market wealth and population.

Fraser et al. (2009), while recognising income as one of the main determinants of house prices, pointed out that an assumption of a constant, as opposed to time-varying, relationship between house prices and income is highly unlikely. Similarly, Brown et al. (1997) applied the Time Varying Coefficient (TVC) approach to quarterly data of house prices, disposable incomes, a demographic variable and the nominal user cost of housing for the period between 1968 and 1992. The results confirmed the TVC methodology to outperform an alternative constant parameter procedure in generating a house prices forecast. However, the study did not include the period of the housing

market crash in 1992, with researchers intending to extend the investigation to further analyse the performance of their approach beyond this date.

Black et al. (2005) analysed UK house prices in a sample period from 1973:04 until 2004:03 in order to define fundamental prices and investigate their connection to the inflationary process. As a result, they found that in this period actual house prices deviate from their fundamental prices. In addition, speculative activities do not seem to be the core cause of these deviations. Instead, over-sensitivity to expectations about fundamentals seemed to be the main reason for such behaviour. Case and Shiller (2004) found income to explain behaviour of US house prices for the majority of the data considered. Moreover, it appeared that in the states where income and house prices were highly correlated, inclusion of additional fundamental factors to the regression gained little explanatory power. Including these factors in regressions for the data where income had less explanatory power, on the other hand, significantly improved the R^2 value.

Black et al. (2006) modelled fundamental values of the UK house prices using time-varying present value model excluding an explosive rational bubble caused by non-fundamental factors as the reason for deviation of actual prices from their fundamentals. Instead, researchers find intrinsic and momentum price dynamics to be main determinants of price deviation. Fraser et al. (2008) also applied a time-varying present value model approach to determine the fundamental values of housing prices in New Zealand based on the real disposable income between 1970 and 2005. The real disposable income was used in the study in order to capture the income of households after taxes and inflation in contrast to previous studies which used equilibrium models with real income, real employment and real interest rates. Fraser et al. (2008) also based

their model on empirical framework proposed by Black et al. (2006). Fraser et al. (2008) defined the intrinsic bubble as the bubble component related to fundamentals. The results have shown deviations from the fundamental prices, however, most of housing market overvaluation was due to price dynamics and not an overreaction to changes in fundamentals. Following their previous study, Fraser et al. (2009) investigated the relationship between house prices and income in terms of responsiveness of the prices to temporary and permanent shock to the income. As a result, researchers found New Zealand's house prices to be higher than suggested by the deterministic component, thus suggesting the temporary component to be responsible for such price overreaction. The UK data, on the other hand, revealed that an increase in housing prices was stimulated by permanent deterministic components. In addition, US house prices seemed to be more responsive to fluctuations in temporary or cyclical price components. As a result, Fraser et al. (2009) concluded that there was no consistent global pattern in the behaviour of house prices as a response to permanent and temporary income shocks, suggesting a closely considered and tailored approach to this phenomenon by policy makers.

Evident inadequacy of simple linear models to produce accurate results have naturally integrated into academic studies concentrating on identifying and modelling non-linearities and asymmetries present in house prices. The study by Holly and Jones (1997) investigated one of the most extensive data sets of UK housing prices from 1939 to 1994. The researchers considered the possibility of non-linear adjustment in house prices dynamics by applying an asymmetric error-correction model and found the real income to be the main determinant of the real house prices in an asymmetric cointegrating relationship. Periods of disequilibrium between the two variables are

attributed to periods related to significant demographic changes and regulations on lending and mortgages. Miles (2008) suggests that symmetry in rising and falling markets imposed by linear models seems to be inappropriate for successful forecasting of housing prices due to the specific nature of the housing market often characterised by booms and bursts. Miles (2008) suggests a non-linear approach to forecasting housing prices, however, a Markov-switching model in particular was found to perform poorly in out-of-sample forecasting. Hence, the study applies the generalised autoregressive (GAR) model. Results showed the GAR model to perform better in out-of-sample forecast than ARMA and GARCH models when modelling high volatility house prices. Gao et al. (2009) recognised asymmetric patterns in house prices whereby price increases are very rapid while price declines are characterised by much slower speed. Gao et al. (2009) also distinguished between two types of behaviour displayed by the house prices: cyclical or volatile and non-cyclical or tame; and found cyclical markets to be characterised by larger autoregressive coefficients compared to non-cyclical markets. In addition, Gao et al. (2009) attached an important value to regional variability in house prices dynamics. Crawford and Fratantoni (2003) recognised the importance of specific boom-burst dynamics of the housing market and proposed the use of non-linear models in forecasting house prices. Black et al. (2006) also confirmed the cyclical nature of housing market dynamics.

Presence of non-linearities and asymmetric dynamics of house prices behaviour have been observed in the housing market and attributed to various factors. Case and Shiller (2004) also confirmed downward stickiness of house prices following the survey on homebuyers in the US carried out in 2003. In addition to sticky downward house prices, Case and Shiller (2004) also pointed out the existence of sellers' reserve prices thus

offering an ample explanation of the situation when the house prices do not fall immediately subsequent to an excess in supply. Moreover, supply constraints seem to be very significant in house price reactions to any market changes. For instance, Himmelberg et al. (2005) found that changes in house market behaviour could be seen as a local phenomenon with different US cities reacting differently to changes in fundamentals, which can be attributed to differences in elasticity of housing supply. Similarly, Goodman and Thibodeau (2008) point out the observed house prices increases in certain areas of the US were caused by the inelastic supply of housing and hence the average increase in the national aggregate house prices masks different rates of price appreciation for different regions for the country. They also found that the rate of house prices appreciation was directly affected by and was very sensitive to the housing supply elasticity. Restricted supply may indeed increase housing prices, however, Case and Shiller (2004) point out that building and construction companies, assuming these are driven by profit maximising strategies, seem to respond clearly to rises in demand and thus prices by increasing supply, permitting any required regulations. In addition, Case and Shiller (2004) also suggested that changes in employment might have positive as well as negative impact on demand for housing. Thus an increase in employment in certain areas will naturally be accompanied by increased demand for housing followed by a rise in prices, however, high housing costs, on the other hand, may make it difficult to attract employees thus decreasing growth of employment for that region.

According to McQuinn and O'Reilly (2008), in the period between 1995 and 2005 the house prices for new Irish homes increased by 260 percent following a very successful performance of the Irish economy. Their investigation into demand for real estate as

being determined by the borrowing constraints indeed revealed a long-term cointegrating relationship between actual house price levels and the fundamental levels determined by the average amount of individual borrowing. McQuinn and O'Reilly (2008) suggested that the level of borrowing by individuals from financial institutions depends on levels of personal disposable income and interest rates.

Case and Shiller (2004) noted that during a housing bubble homebuyers consider high housing prices affordable due to assumption that the purchase will be compensated as with further price increases. As researchers found out, the belief of further house price increases seems to be especially relevant to first time buyers as they become anxious that it will become more difficult to afford properties later on. This adds to the little perceived risk by homeowners treating real estate as an investment.

Koetter and Poghosyan (2009) criticised previous studies for not taking into account of regional differences when applying housing price indicators in modelling, such as regional differences in financing schemes and tax laws, as well as evident immobility of real estate assets. According to Koetter and Poghosyan (2009), due to differences in these regional regulations the traditional approach to the analysis of real estate markets lacks the entity of objective comparability in cross-country studies. Koetter and Poghosyan (2009) found that while the German housing market was not characterised by rapid price increases the house prices nonetheless deviated from the fundamentals, displaying low speed of adjustment when compared to the results of similar studies on US data.

Fraser et al. (2008) pointed out that limited arbitrage opportunities in the housing market may lead to any mispricing to have an effect for prolonged periods of time.

Muellbauer and Murphy (1997) recognised the importance of transaction costs and boom and burst dynamics of housing markets, and thus presence of non-linearities.

Housing market bubbles versus changes in fundamentals

Large movements in house prices have prompted speculation of the presence of bubbles in the housing market. As it was mentioned in the previous chapter, the concept of bubbles in financial markets has been considered somewhat controversial due to the inability to correctly identify such bubbles. Moreover, the presence of bubbles implies the existence of market inefficiencies. However, despite the debate, an ample number of research studies have been carried out into the investigation of rational bubbles.

Stiglitz (1990) provides a comprehensive definition of a bubble in terms that the bubble exists when the *only* reason for the high price today is that investors believe the selling price of the asset will be high tomorrow and there seems to be no justification of such a price rise in terms of fundamental values of the asset. Furthermore, Black et al. (2006) indicated three types of market bubbles in the context of a housing market: momentum, explosive and intrinsic. While momentum investors' behaviour is motivated purely by price, where a price rise or a price fall is expected to be followed by further price rise or fall respectively, and is treated as evidence against market rationality, rational bubbles and intrinsic bubbles are treated as supporting evidence of rationality. Black et al. (2006) describes explosive rational bubbles to cause price divergence from fundamentals to be driven by extraneous factors, whereas intrinsic rational bubbles

trigger price deviations due to exogenous factors. Furthermore, in the contrast to explosive bubbles, intrinsic bubbles periodically revert to the fundamental equilibrium value.

Hall et al. (1997) applied a two-state Markov process to UK real house prices in order to identify bubble-like behaviour of house prices during known housing booms in 1971-1974, 1977-1979 and 1986-1989. The Markov process estimated in the study was characterised by unknown transition probabilities. The process correctly identified periods associated with housing booms and the probability of bubbles bursting according to their size. Deterministic components of real house prices distinguished in the study included real personal income, the owner-occupied stock of housing, and the mortgage rate of interest. Garino and Sarno (2004) provided empirical evidence of the presence of bubbles in quarterly UK house price data in the period between 1983 and 2002. The study identified two explosive bubbles in the late 1980s and in the late 1990s which is consistent with the house price bubble hypothesis and observed housing booms. In addition, the researchers noted that the latter bubble appeared to extend up to the end of the sample period of 2002. Black et al. (2006) used a time-varying present value approach to investigate the relationship between actual and fundamental housing prices in the UK. Results of the study revealed the presence of a rational bubble caused by non-fundamental factors.

Black et al. (2006) after an investigation of UK quarterly housing data between 1973:04 and 2004:03 found house prices to be overvalued by almost 25 percent at the end of the estimation period with an intrinsic bubble and fundamental components equally contributing to these price dynamics. Goodman and Thibodeau (2008) investigated the presence of speculative bubbles in the US housing market in terms of the extent the

prices will have to be above fundamentals in order to comprise a speculative bubble. This approach inevitably required estimation of fundamental values, which in this particular study were distinguished into long-run equilibrium prices and short-run deviations which move to correct the long-run equilibrium. Applying the rule of a 30 percent increase over the fundamental values as an indication of presence of a speculative bubble, the researchers found that only 25 sets out of 84 considered regions could have been described as surpassing this threshold, suggesting that only specific individual regions were characterised by a speculative housing bubble rather than the country's housing market as a whole.

Coleman et al. (2008) investigated housing bubble dynamics of a sharp rise and subsequent fall in US house prices over the period of 1998 to 2008. The results suggested economic fundamentals to explain the house prices dynamics for the period prior 2003. However, the easy availability of loan products seemed to have encouraged increased consumption levels and rates of house ownership, thus instigating fruitful conditions for occurrence of a bubble. Researchers also found support of supply constraints to have an influence of house price movements. Similarly, Wheaton and Nechayev (2008) found an economic fundamental such as population, income growth and decline in interest rates to explain the increase in house prices between 1998 and 2005. Coleman et al. (2008) also pointed out a regime shift in early 2004 which significantly affected lending patterns with the record increase of lending volume. This was brought about by a combination of political, regulatory and economic factors, and seemed to have reduced the importance of fundamentals in determining the house prices and for the prices to display bubble characteristics.

However, Stiglitz (1990) pointed out that the difficulty in identifying the presence of a bubble also lies in establishing whether the terminal price taken as the fundamental price level is indeed determined by the fundamental factors or is characterised by the reminiscence of another bubble. In other words, it is very challenging to distinguish between movements of a bubble and misspecification of a fundamental model. Himmelberg et al. (2005) investigated the US housing market highlighting difficulties when assessing whether rapid growth in house prices is the result of changes in fundamentals or presence of a bubble. Himmelberg at el. (2005) points out that high price growth is not necessarily an indication of the house prices to be overvalued. Thus, though they did find evidence of a housing bubble in the US data at the end of 2004, the results did not reveal excessive price increases over the fundamental prices. According to Himmelberg et al. (2005), a fall in house prices could be initiated by changes in economic fundamentals such as, for instance, a negative shock to the economy or a decline in economic growth, as well as increased sensitivity to increases to mortgage rates as a result of an unanticipated rise in interest rates. Moreover, Stiglitz (1990) suggests that there is no need to interpret a decline in asset price as the breaking of a bubble. Stiglitz (1990) demonstrates the argument on an example of crude oil prices being dependent on a speculative element of the future possibility of a development of a petroleum substitute which might reduce the demand and, hence, the value of oil. Thus, a sudden decline in the price of an asset could arise due to an occurrence of new information relevant to the future developments of the asset. However, there is a plausibility of a sharp decline in prices to be attributed to the breaking of a bubble when no presence of such new information has occurred. In addition, Stiglitz (1990) proposes that the presence of speculative bubbles could be supported by the fact that no other

evidence of the cause behind the market crashes of October 1987 and October 1989 have been found to explain these events. Similarly, Stiglitz (1990) questions the rationale of interpretation of the booms of the 1920s and the crashes in 1929. Correspondingly, White (1990) explains the 1920s events as a boom and burst of a speculative bubble by systematically excluding other alternative explanations.

As pointed out by Case and Shiller (2004), diminished demand for housing may cause the fall in housing prises thus resulting in the burst of the housing bubble. However, Case and Shiller (2004) found changes in fundamentals to explain much of the price increase in the housing market. Abraham and Hendershott (1996) found that the larger the bubble grows the more likely it is to burst. Abraham and Hendershott (1996) constructed their model to include a proxy for the bubble tendency to burst. The proxy was formed to account for the differences occurring between the actual house prices and the prices dictated by the fundamentals. Abraham and Hendershott (1996) found using such proxy to be useful in explaining large cyclical movements in house price levels. The researchers found that the inflation of the real cost of construction, real income growth and changes in real after-tax interest rates as determinants of real house price appreciation explained nearly half of historical fluctuations in the inflation of the real house prices. However, the model used by Abraham and Hendershott (1996) failed to explain prolonged cycles in house prices for some of the regions considered in the study suggesting the presence of bubbles. Abraham and Hendershott (1996) pointed out that it is extremely difficult to distinguish between whether changes in house price are caused by fundamentals or bubbles. Complicating the issue even further by comprising that bubble-like behaviour could also be a result of model misspecification.

Following an unprecedented rise in house prices in Dublin, Roche (2001) applied a regime-switching model to investigate whether the cause for such an excess increase in demand for housing was brought about by speculative bubbles or changes in economic fundamentals. In order to do that Roche (2001) divided the house prices into fundamental and non-fundamental components and used one of the methods of calculating non-fundamental house prices based on a standard asset-pricing model. Most of the models used in the study could be rejected in favour of the regime-switching model, which seemed to produce some evidence supportive of the presence of speculative bubbles.

Mikhed and Zemčík (2009) carried out a research into the determinants of house prices using the present value model approach and confirmed the presence of a bubble in the US housing market prior to 2006, which appears to be reverting to fundamentals after two years. Meese and Wallace (1994) also investigated house prices in the context of the present value model and found that while the relationship between actual and fundamental prices was rejected for the short-run data, it persisted in the long-run, thus recognising the presence of informational asymmetry in housing prices. Meese and Wallace (1994) attributed asymmetric adjustment to the presence of large transaction costs, which suggests that the utility gains from acting upon movements in the housing market will have to exceed transaction costs in order for the trade to take place. Meese and Wallace (1994) reject the proposal of bubbles as one of the reasons for failure of the short-run present value relationship in the housing market on the basis of absence of any empirically compelling evidence of the presence of bubbles. Similarly, Black et al. (2006) pointed out that the process of equilibrium correction of house prices will be prolonged by the limit to arbitrage. According to Black et al. (2006), limit to arbitrage

and hence longer reaction of housing price to revert to long-term equilibrium are due to high transaction costs, and due to heterogeneity and illiquidity characteristics of housing. Black et al. (2005) also highlighted the importance of identifying and understanding housing bubbles, since these have an impact on inflation, thus pointing out that studies in this area are very much significant for policy makers.

Forecasting of the housing market

Many researchers highlighted the importance of housing prices to the whole economy (Case and Shiller, 1989; Brown et al., 1997; Muellbauer and Murphy, 1997; Crawford and Fratatoni, 2003; Turner, 2003; Fraser et al., 2008; Miles, 2008; Koetter and Poghosyan, 2009). Brown et al. (1997) pointed out that movements in the UK housing market affect general price levels and consumer expenditure. Similarly, Garino and Sarno (2004) recommended further extended research into theoretical work concerning the house prices behaviour due to its importance in practical implications and the impact of public policy on housing markets and thus standards of living and patterns of saving and borrowing.

Case and Shiller (1989) carried out research into the efficiency of the housing market, and found it to appear inefficient where house prices did not follow the random walk model, however did reveal substantial persistence in price changes. Hence, Case and Shiller (1989) suggested the house prices to be forecastable. Muellbauer and Murphy (1997) provided empirical evidence of both house prices and relative rates of return in

housing to be forecastable, thus refuting the hypothesis of efficiency of the housing market. Koetter and Poghosyan (2009) argue that in an efficient and frictionless market, properties in the real estate would reflect economic cycles and be priced according to the demand and supply which in turn will be determined by economic fundamentals. However, this relationship between house prices and macroeconomic fundamentals does not seem to hold. Koetter and Poghosyan (2009) name three main reasons for its failure, including the fact that real estate represents a non-standardised asset characterised by regional differences. In addition, the absence of principal trading centres causes imperfect information and thus lack of transparency and high transaction costs. Finally, Koetter and Poghosyan (2009) point out sluggish response to changes in supply as a result of construction times and limited land availability. Black et al. (2006) proposed that the occurring price inefficiency in housing market is due to limited arbitrage which results in prolonged periods of adjustment back to the fundamental equilibrium thus causing pricing inefficiencies. Crawford and Fratantoni (2003) pointed out that while price changes of an individual real estate are sometimes difficult to predict, the house prices as an aggregate, on the other hand, are forecastable.

Fraser et al. (2008) emphasised the importance of understanding and correctly identifying movements of the housing sector, since dramatic changes in the housing market have a greater effect on the whole economy comparing to financial stock market movements, thus recognising causes of house price movements is vital for policy makers. Fraser et al. (2008) point out that, for instance, a rapid increase in housing wealth will instigate increased consumption and aggregate demand which in turn put pressures on maintaining the inflation band. Thus, the intervention will not be required in the case of a price increase caused by the changes in fundamental values. A

speculative bubble, on the other hand, will signify central bank intervention in terms of attempts to control inflation and potential economic slowdown which can be brought by the eventual burst of the bubble. Similarly, Miles (2008) underlines the importance of accurate forecasts of changes in house prices. Miles (2008) suggests that a significant increase in house price over the last decade had an effect on many economic components, including increased consumption and growth of secondary mortgage ownership. Moreover, these dynamics appear to be not only on a local or national level, but to be a global phenomenon. Similarly to Fraser et al. (2008), Black et al. (2006) supports the importance of understanding the housing market due to its considerable wealth effect which has been shown to be greater than that of financial assets. Moreover, Black et al. (2006) points out that due to housing assets forming a major part of household portfolios, the housing market crashes were observed to have more severe effects on the whole economy than stock market crashes and are characterised by longer recovery periods. Ortalo-Magné and Rady (2006) find that changes in income affect house prices and housing transactions. The researchers investigate house price dynamics in the context of credit constraint and limits of down payment which is an especially significant factor for young and first time buyers. These restrictions to entering the housing market by young households in turn can affect the housing market as a whole.

Crawford and Fratantoni (2003) pointed out the importance of accurate modelling and forecasting of house prices for pricing mortgage credit risk. Similarly, Rosen et al. (1984) illustrated that volatility of house prices adds risk to household portfolios and thus discourages homeownership of naturally risk averse investors. According to Rosen et al. (1984), this phenomenon is evident during increases in house prices in the 1970s without subsequently significant changes in the number of homeowners. Englund et al.

(2002) pointed out the cost of homeownership will account for a larger portion of portfolios of an especially younger household with presumably lower incomes thus imposing an additional risk. Similarly, Turner (2003) found that low or moderate income households and first-time buyers are more sensitive to house prices volatility, which could be due to high income households' ability to diversify their housing investment portfolios to a greater extent. Turner (2003) demonstrated a significant negative effect of investment risk which can arise from house prices volatility on homeownership and housing demand. Miles (2008) denotes growing financial sophistication of consumers, lenders and pension funds, which in turn encourages increased demand for accurate forecasts of house prices. This is especially valid in banking, as the housing market had a great effect on the financial market in a recent time period and since the probability of default and mortgage prepayment is determined by the volatility of house prices. Thus, more accurate forecasts will assist in management of prepayment risk on mortgage backed securities. Moreover, Koetter and Poghosyan (2009) point out that the policy makers do take into account property prices as being one of the indicators of financial market susceptibility since imbalances in the housing market can lead to instability in the financial sector due to the banks acting as mortgage lenders. Consequently while an increase in house prices might increase the value of real estate in the bank's possession and thus improve bank capital, and decrease the probability of mortgage borrowers defaulting on appreciated assets, the same house price increase and consequent lower perceived risk might also bring instability to banks by encouraging lendings to higher risk real estate at a lower interest rate.

Das et al. (2009) found a large-scale Bayesian Vector Autoregressive (BVAR) model to outperform linear alternatives in forecasting annualised real house price growth rates for

South Africa. According to researchers, linear models have failed to produce favourable forecasts because these failed to recognise non-linearities present in the data and did not take into account of asymmetries in house prices dynamics. Crawford and Fratantoni (2003) proposed non-linear regime-switching models as the most suitable for forecasting housing markets which are known to be prone to boom and busts. Different regimes under regime-switching models can accommodate different behaviour exhibited by the housing market under different economic conditions. The study found the regime-switching model to fit the house prices data better than linear alternatives, however, the simple ARMA model outperformed regime-switching in the out-of-sample forecasting exercise.

Conclusion

The review of the housing market literature suggests rather an understated amount of research into the dynamics of the market and reasons behind it, as well as limited studies of the house prices forecasting exercise, while it is difficult to underestimate the importance of investigations into these issues. Thus, Black et al. (2006) recommended further investigation into the causes of changes in fundamental values as an essential research with the intention of the facilitation of a better understanding of house prices dynamics by policy makers, while Miles (2008) points out a significant importance of accurate forecasts of house prices. Moreover, Das et al. (2009) pointed out the importance of forecasts of the house prices inflation to policy makers since changes in

house prices have direct effect on overall consumption, inflation and investment as houses make up a large proportion of a households' wealth. Moreover, Das et al. (2009) suggests the recent credit crunch in the US imminently resulted in the economic recession, initiated by the burst of the housing bubble.

The importance of investigating the housing market on a global level is evident from an investigation carried out by Beltratti and Morana (2010), where researchers applied a large scale macroeconomic model to investigate linkages between house prices and macroeconomic variables, and global factors determining international house prices in the data for the G-7 countries. The research revealed that indeed the global economic shocks have a major effect on fluctuations of the international housing prices with supply shocks having a larger consequence on price levels compared to demand variations. In addition, Beltratti and Morana (2010) found that the international macroeconomic and financial shocks can be construed by those of the US, suggesting the importance of the US market's influence on the global economy. Beltratti and Morana (2010) also found that while both stock market shocks have significant effects on macroeconomy, the housing market price shocks have far greater effect than the stock market shocks. Beltratti and Morana (2010) pointed out that, according to their results, international housing markets appear to be interconnected and there is a possibility of speculative behaviour. House prices, at the same time, are also affected by the supply side individual to each country, which strongly suggests rational pricing as opposed to fads. However, the explicit investigation into the international housing market speculation and its effects is yet to be carried out. Beltratti and Morana (2010) concluded that policy makers should take into account the international business cycle

when assessing macroeconomic risks and suggested a further investigation into instability of the international banking market in the light of the current credit crunch.

While the housing market shares some characteristics with the financial market, it also possesses very specific characteristics which result in unique reactions of the market to major changes. One of these features of the housing market is the price trends being a localised phenomenon, producing different price reactions in different regions (Abraham and Hendershott, 1996; Goodman and Thibideau, 2008; Koetter and Poghosyan, 2009). Consequently, basing their assumptions on the historical record, Case and Shiller (2004) suggested that a severe nationwide crash in housing prices is highly unlikely due to localised trends of house price movements implying the lack of synchrony in the response to regional markets. Thus, the lack of synchronous response of the aggregate housing market will diminish the severity of effects on the economy following the eventual burst of the housing bubble. However, Gao et al. (2009) pointed out that a decline in a housing market can lead to a considerable amount of mortgage defaults due to borrowers' equity diminishing in value. Gao et al. (2009) attribute the mortgage melt-down in 2007 to such a decline in the housing market.

Moreover, Case and Shiller (2004) approached investigation to the housing prices dynamics in explaining market bubbles and bursts somewhat differently by conducting a survey amongst a random sample of homebuyers in 2003. The aim of the survey was to focus on homebuyers' expectations, their understanding of the housing market and hence their behaviour in response. Despite receiving a lower response rate in 2003 comparing to the survey conducted previously in 1988, the researchers were able to draw certain conclusions about expectations and perceptions of US homebuyers. The lack of references to quantitative or professional based evidence by the respondents

illustrated the degree of amateurism amongst buyers and sellers in the housing market. Regardless of a small number of possibly professional market speculators, the majority of homebuyers are owner-occupiers treating the real estate as a long-term investment. Moreover, the survey revealed that whilst the behaviour and investment decisions of homebuyers seemed to be based on their exaggerated expectations, emotional excitement about local real estate and casual word of mouth, majority of survey participants did not believe the housing market was driven by psychology.

From the literature review it is evident that while the housing market is an important part of the economy and its dynamics are greatly considered by the policy makers, academic research has yet to offer ample investigations into house prices behaviour and adequate forecasting methodology. Moreover, studies by Englund et al. (2002), Turner (2003), Case and Shiller (2004) and Ortalo-Magné and Radly (2006) confirmed the importance of movements in house prices on an individual homeowners level due to the property investment forming the majority of household portfolios. Hence, it is difficult to comprehend that such complex dynamics of the market are associated with large number of highly amateur investors due to most homebuyers lacking the professional speculative qualifications.

5.3. Methodology

The following methodology is based on the approach introduced by Black et al. (2005), which in turn has the present value model at its foundation proposed by Campbell and

Shiller (1987; 1988a; 1988b). Campbell and Shiller's (1987) original present value model was constructed for two random variables, Y_t and y_t, where Y_t is a linear function of the present discounted value of its expected future values, y_t (5.1).

$$Y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t y_{t+i} + c \tag{5.1}$$

The present value equation contains a coefficient, c, proportionality coefficient, θ, and the constant discount factor, δ.

The model used by Black et al. (2005) relates the real house prices to the expected value of discounted future real disposable income. Similar to the principals of the original present value model the model used by Black et al. (2005) intends to capture the size of deviations of real house prices from their fundamentals. Black et al. (2005) assume the expected value of future real disposable income discounted at the real discount rate as a proxy of the fundamental residential property value.

There are various methods of determining fundamental values for house prices. Some researchers, for instance, compare real house prices to disposable income. Thus, Muellbauer and Murphy (1997) suggested a measure of affordability in the form of price to income ratio to model booms and bursts in the housing market. Black et al. (2005) also based their study of fundamental prices being linked to the affordability concept, which relates to the perception of wealth and based on real wages, employment rates and real interest rates. In Black's et al. (2005) case the affordability proxies are real disposable income and real interest rates.

Hence, based on the present value model, the following approach relates the real house prices to the expected value of discounted future real disposable income:

$$P_t = E_t \sum_{i=0}^{\infty} \delta^i Q_{t+i} \tag{5.2}$$

where P_t is the real level of house prices, E_t is the expectations operator, δ^i is the required rate of return, or discount rate, and Q_{t+i} is the real disposable household income in the period between *t* and *i*.

By dropping expectations, equation (5.2) can be written as following:

$$P_t = \sum_{i=0}^{\infty} \delta^i \, Q_{t+i} \tag{5.3}$$

Realised discount rates, ρ_t, or the real return, is defined in Black et al. (2005) as:

$$(1 + \rho_{t+1}) = (P_{t+1} + Q_t)/P_t \tag{5.4}$$

The above expression (5.4) can be re-written as following:

$$r_t = ln\big(1 + exp(q_t - p_{t+1})\big) + p_{t+1} - p_t \tag{5.5}$$

where r_t is defined as $ln(1 + \rho)$ and the term $(q_t - p_t)$ is the income-price ratio.

Furthermore, following the work of Campbell and Shiller (1988a, 1988b), the time-varying discount rate can be introduced using a first-order Taylor's approximation for the first term of the equation (5.5) resulting in:

$$r_t = -(p_t - q_{t-1}) + \mu(p_{t+1} - q_t) + \Delta q_t + k \tag{5.6}$$

Where k and μ are linearisation constants:

$$\mu = 1/\big(1 + exp(\overline{q - p})\big) \tag{5.7}$$

$$k = -ln\mu - (1 - \mu) \cdot (\overline{q - p}) \tag{5.8}$$

where $(\overline{q - p})$ is the sample mean of $(q - p)$ about which the linearisation was taken.

Equation (5.6) contains terms p_t and q_t which in practice tend to be *I(1)*, hence to ensure stationarity the equation is re-written so that it contains pq_t which is log price-income ratio $(p_t - q_{t-1})$.

$$pq_t = k + \mu pq_{t+1} + \Delta q_t - r_t \tag{5.9}$$

Further, repeating substitutions for $pq_{t+1}, pq_{t+2}, \ldots$ on the right hand side of the above equation, it can be re-expressed as follows:

$$pq_t = \frac{k(1-\mu^i)}{(1-\mu)} + \sum_{j=0}^{i-1} \mu^j r_{t+j} + \mu^i pq_{t+i} \tag{5.10}$$

After letting $j \to \infty$, limiting last term to zero and taking conditional expectations the equation becomes:

$$pq_t = \frac{k-f}{(1-\mu)} + \sum_{j=0}^{\infty} \mu^j E_t \Delta q_{t+j} - \alpha \sum_{j=0}^{\infty} \mu^j \, E_t \sigma_{t+j}^2 \tag{5.11}$$

where f is the constant real-risk free component of real required returns.

The final equation for the ratio which measures fundamental house prices is modelled as below:

$$pq_t^* = \frac{k-f}{1-\mu} + (e_2' - \alpha e_3')A(I-\mu A)^{-1} z_t \tag{5.12}$$

where $e_2' z_t = E_t \Delta q_{t+i}$ and $e_3' z_t = E_t \sigma_{t+1}$.

The equation for fundamental price-income ratio (5.12) can be re-written in simplified form.

$$pq_t^* = (k/1-\mu) + E_t \sum_{i=1}^{\infty} \mu^i \left(\Delta q_{t+i} - r_{t+j}\right) \tag{5.13}$$

According to Black et al (2005), actual and fundamental prices deviations can be assessed by simply testing $pq_t = pq_t^*$, i.e. equations (5.12) and (5.13).

Providing stationarity of changes in real income and stationarity of the discount rate, the relationship modelled by the above equation (5.12) implies that log prices and log income are cointegrated with a cointegration vector of $[1, -1]$. If this relationship holds, the price-income ratio should be stationary. Hence, the statistical analysis of the equation (5.13) will involve testing the log price-income ratio for the presence of stationarity.

5.4. Empirical results

This study analysed quarterly real house prices from thirteen UK regions, including the UK as a whole, and a price series for the UK, and quarterly real disposable income data over the period of thirty years from 1974:01 to 2004:04. The full list of regions used in this paper is given in Table 5.1 below. In addition, following the methodology in Section 5.3, the price-income ratio for each region has been calculated using a technique based on methodology by Black et al. (2005).

Table 5.1. List of regions.

	Region
1	North
2	Yorkshire and Humberside
3	North West
4	East Midlands
5	West Midlands
6	East Anglia
7	Outer South East
8	Outer Metropolitan London
9	London
10	South West
11	Wales
12	Scotland
13	Northern Ireland
14	UK

Descriptive statistics

The following diagrams (Figure 5.1 – 5.2) represent the time-series plot and histogram with descriptive statistics for the UK house prices as one of the considered series.[8] House price booms in 1980s and 1990s, and a dramatic increase in 2000s are clearly seen in the time-series pattern (Figure 5.1). The income growth and financial deregulation with easy availability of mortgages had resulted in a house price boom in the early 1980s. The next housing boom in the late 1980s was followed by a burst in the 1990s due to rises in interest rates and stricter mortgage criteria following the economic

[8] All the estimations and statistical calculations are performed using the EViews 3.1 software.

recession. The late 1990s and early 2000s show signs of recovery followed by a rapid price increase well into the end of the period. Moreover, it is evident that the overall price level has risen dramatically over past thirty years.

Figure 5.1. Real house prices, UK.

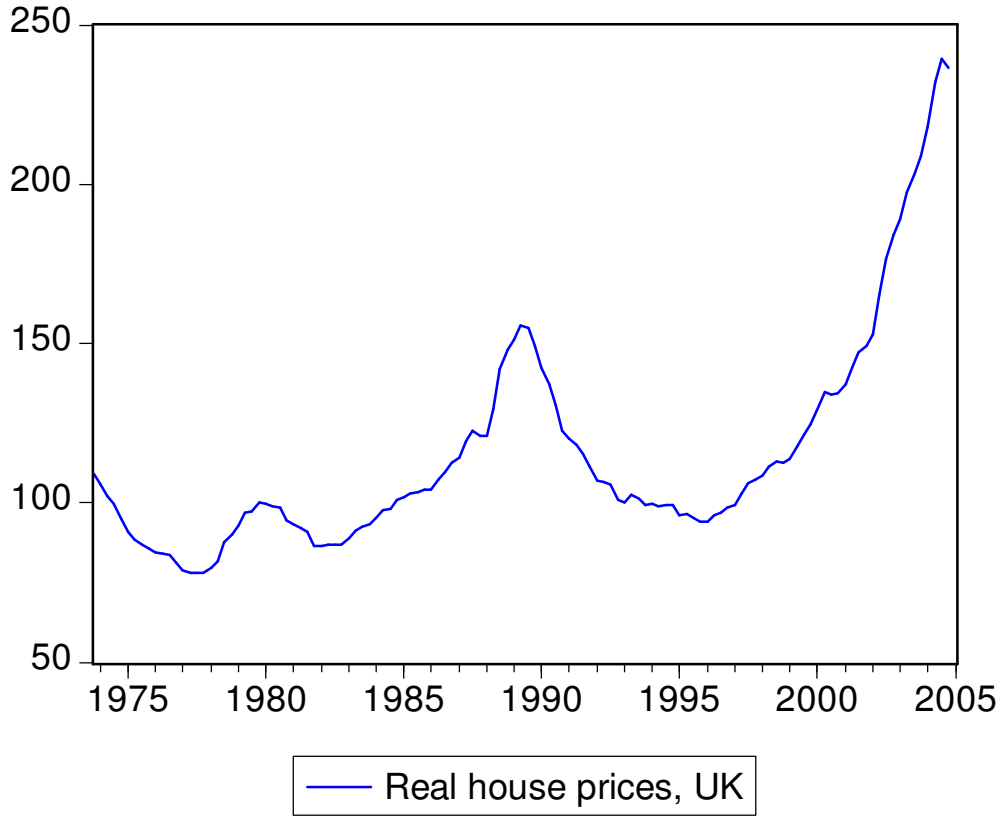
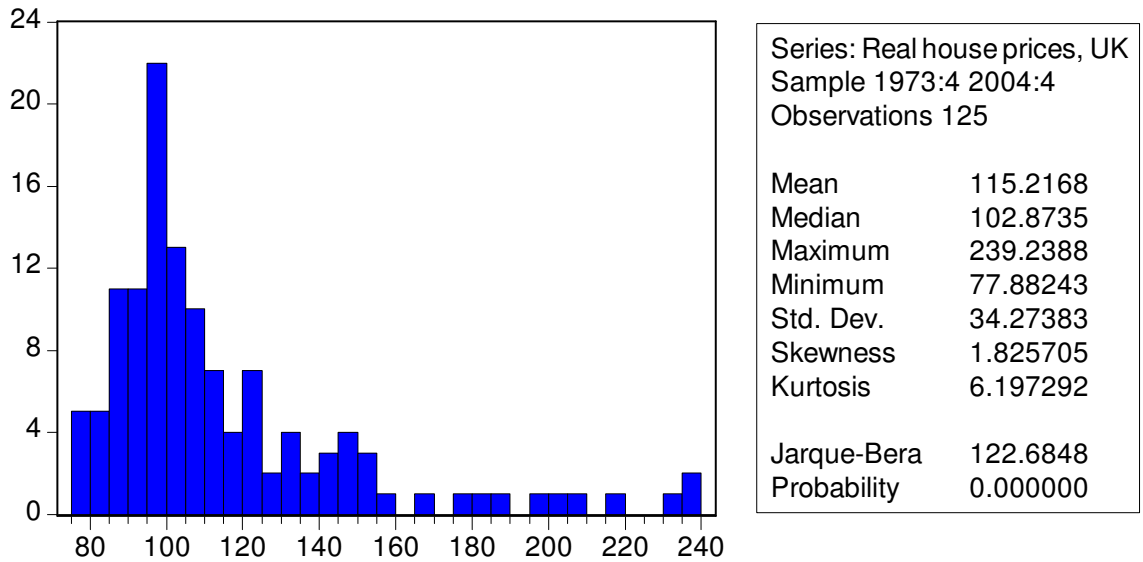


Figure 5.2. Histogram of real house prices, UK.



The histogram provides a good insight into the shape of the data distribution, whereas skewness and kurtosis indicate the symmetry and thickness of the tails of a distribution respectively. High kurtosis in the UK price series indicates a presence of fewer extreme values and more moderately sized observations, which is consistent with the time-series where extreme values are observed towards the end of the sample. Distribution for the UK series seems to be skewed to the right and have the tail of the distribution thicker than normal, supporting the assumptions made on the basis of kurtosis that the main concentration of the distribution is focused around lower observation values. According to the results of the Jarcque-Bera statistic, the hypothesis of normality was rejected for UK data at 5% significance level. The UK data has relatively high standard deviation compared to the other regions. However, it has to be taken into account that this data is likely to be non-stationary.

Figure 5.3 represents the time-series of the price-income ratio for UK data series, estimated using equation (5.13) from the methodology section (Section 5.3). Figure 5.4 represents the histogram and descriptive statistics for the price-income ratio.

Figure 5.3. Price-income ratio, UK.

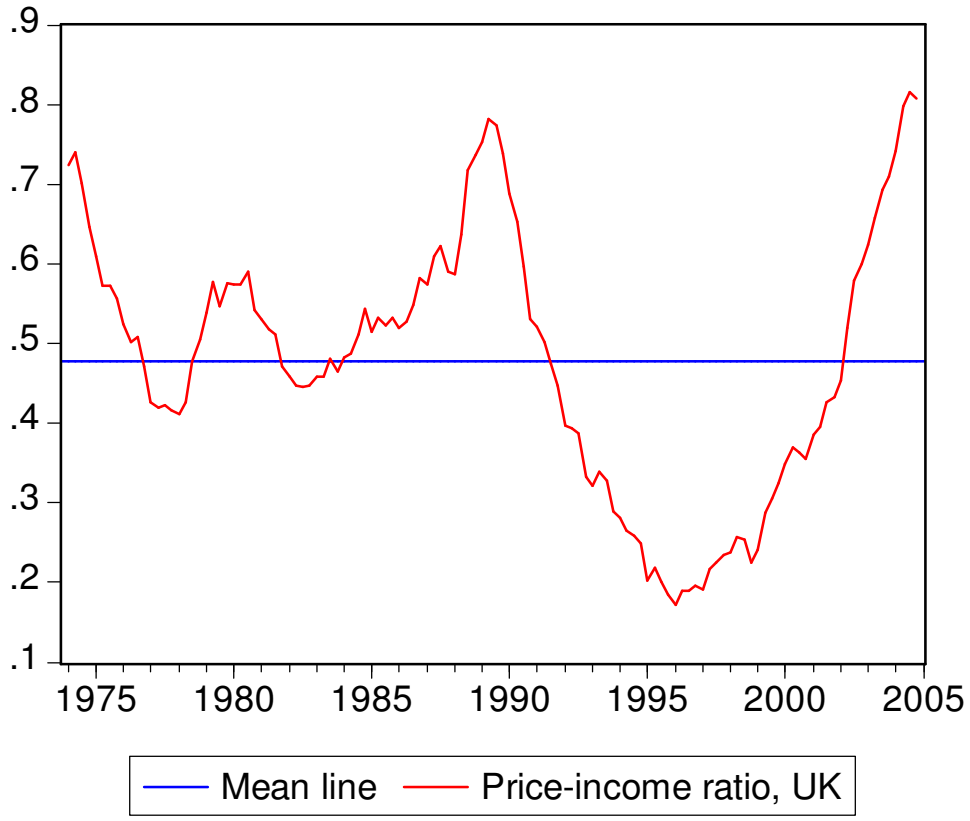
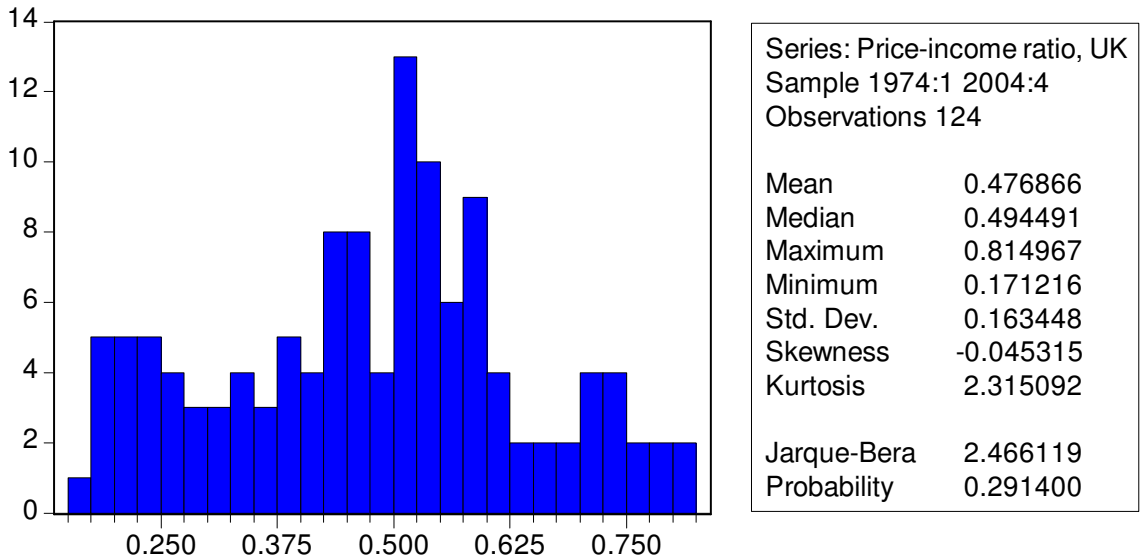


Figure 5.4. Histogram of price-income ratio, UK.



Consistent with the house prices time-series, the price-income ratio diagram demonstrates large increases during the boom in the 1980s due to high income growth, and rapid decline in the 1990s followed by the rise in interest rates and economic recession. In addition, a similar pattern of recovery could be seen towards the 2000s. The distribution of the UK price-income ratio seems to approximate to normal distribution as according to the Jarcque-Bera statistic, the hypothesis of normality could not be rejected for the UK price-income ratio at 5% level of significance. However, as it is evident from the diagram and the value of skewness, the lower tail of the distribution is thicker than that of a normal distribution. Negative skewness also suggests the main body of the distribution to be concentrated on the right of the diagram implying few low values.

Nonetheless, descriptive statistics give merely a brief description of the data in order to give a researcher an idea of its basic characteristics. Further examination of the data considers issues of a different nature.

Unit root tests

The augmented Dickey-Fuller (ADF) test was used as a linear unit root test on house prices, logs of house price, house prices returns, price-income ratio, logs of price-income ratio and price-income ratio returns. The test was performed including an intercept and one lagged difference. Prices, price-income ratios and logs of price and price-income ratios were found to be non-stationary, while the null hypothesis of unit root was rejected for house prices returns and price-income ratio returns at 5% level of significance (Table 5.2).

Non-linear unit root tests performed on the data included tests for presence of general STAR-type non-stationarity by Pascalau (2007), asymmetric ESTAR stationarity (Sollis, 2009), ESTAR stationarity (Kapetanios et al., 2003), and LSTAR stationarity (Pascalau, 2007). The results of these tests are presented in Table 5.3.

The procedure developed by Kapetanios et al. (2003) is based on a specific ESTAR model where the *t*-type test procedure involves testing a first-order Taylor series approximated to the ESTAR model.

$$\Delta y_t = \beta y_{t-1}^3 + \varepsilon_t \tag{5.14}$$

Where variable y_t is substituted with the price-income ratio described in the above methodology, and ε_t is the error term. The cubed coefficient y_{t-1}^3 contained in the above equation, is the main analytical indicator used in assessing stationarity using

methodology by Kapetanios et al. (2003). The null hypothesis of unit root ($H_0: \beta = 0$) is tested against the alternative of ESTAR stationarity ($H_1: \beta < 0$). The significant negative value of the coefficient indicates that the ESTAR stationarity holds.

$$t_{NL} = \hat{\beta}/s.e.(\hat{\beta}) \tag{5.15}$$

The *t*-statistic above (5.15), where $\hat{\beta}$ is the OLS estimate of β and $s.e.(\hat{\beta})$ is the standard error of $\hat{\beta}$, tests the null hypothesis of $\beta = 0$ against $\beta < 0$. Asymptotic critical value of the t_{NL} statistic for the type of data used in this chapter is -2.22 for 5% level of significance for the data that was neither de-meaned nor de-trended (Kapetanios et al., 2003). The procedure was carried for all thirteen regions, including the UK as a whole, and the results of testing the null hypothesis of non-stationarity against the alternative hypothesis of stationarity are presented in Table 5.3.

General STAR-type stationarity test (5.16) developed by Pascalau (2007) where the null hypothesis of unit root ($H_0: \gamma = \beta = \delta = 0$) is tested against the presence of ESTAR or LSTAR stationarity ($H_1: \gamma + \beta + \delta < 0$) is based on the work by Kapetanios et al. (2003). Critical value at 5% level of significance for neither de-meaned nor de-trended data equals 3.64.

$$\Delta y_t = \gamma y_{t-1}^2 + \beta y_{t-1}^3 + \delta y_{t-1}^4 + \varepsilon_t \tag{5.16}$$

However, the rejection of the null hypothesis of unit root in the above test cannot distinguish between ESTAR and LSTAR stationarity. Thus, Pascalau (2007) proposed an additional test for the logistic smooth transition (LSTAR) process non-stationarity (5.17), with the critical value for the untreated data of 4.51 at 5% level of significance.

$$\Delta y_t = \gamma y_{t-1}^2 + \delta y_{t-1}^4 + \varepsilon_t \tag{5.17}$$

The unit root test developed by Sollis (2009) allows for asymmetry within ESTAR-type non-linear dynamics. The null hypothesis of unit root is tested against the alternative of asymmetric ESTAR (AESTAR) non-linear stationarity as the regression coefficients are equal to zero ($H_0: \beta = \delta = 0$). The critical value used in this chapter at 5% level of significance is 2.505.

$$\Delta y_t = \beta y_{t-1}^3 + \delta y_{t-1}^4 + \varepsilon_t \tag{5.18}$$

Table 5.2. ADF test with intercept and one lagged difference.

	Region	**House prices**	**Log house prices**	**House prices returns**	**Price-income ratio**	**Log price-income ratio**	**Price-income ratio returns**
1	North	2.4294	1.7280	-5.0543*	-1.1415	-0.5545	-10.2670*
2	Yorkshire and Humberside	0.0649	-0.0685	-4.3442*	-1.7780	-1.5534	-9.1173*
3	North West	1.5559	0.7254	-4.2486*	-1.2505	-1.3665	-6.6173*
4	East Midlands	0.5243	0.2022	-4.1623*	-1.5767	-1.2055	-6.1407*
5	West Midlands	0.6515	0.4804	-4.9310*	-1.6066	-1.3691	-6.1539*
6	East Anglia	-0.0491	-0.1802	-4.3220*	-1.3319	-1.0651	-6.2818*
7	Outer South East	-0.6144	-0.5367	-4.0220*	-1.3352	-1.0476	-5.4070*
8	Outer Metropolitan London	-0.3138	-0.6934	-3.7598*	-1.5207	-1.6024	-4.3783*
9	London	0.3390	-0.1257	-3.7354*	-1.0524	-0.9750	-5.0673*
10	South West	0.3606	0.0710	-4.7207*	-1.1328	-1.0436	-5.7897*
11	Wales	1.2376	0.8777	-3.8893*	-1.5124	-1.0809	-8.0248*
12	Scotland	1.8195	1.2440	-6.9703*	-1.8075	-1.3982	-8.5925*
13	Northern Ireland	2.8502	1.6901	-7.1757*	-1.0516	-0.9935	-9.5680*
14	UK	0.0175	-0.1698	-4.3201*	-1.4738	-1.1819	-5.4021*
Note : * indicates stationarity at 5 % level of significance (critical value: -2.8685)							

Table 5.3. Non-linear unit root tests results for price-income ratio.

	Region	**General STAR**	**AESTAR**	**ESTAR**	**LSTAR**
1	North	33.4391*	50.3027*	-5.1934*	47.5075*
2	Yorkshire and Humberside	13.0923*	19.7171*	-3.9195*	19.1865*
3	North West	3.8306*	4.6023*	-1.5369	3.8818
4	East Midlands	2.1547	1.7460	-1.0147	1.3622
5	West Midlands	0.3971	0.6003	-0.8664	0.5952
6	East Anglia	3.9056*	4.0681*	-1.3880	3.2319
7	Outer South East	3.0185	2.9671*	-1.1700	2.3043
8	Outer Metropolitan London	0.2225	0.3364	-0.7826	0.3366
9	London	5.1861*	5.0877*	-1.5274	4.0100
10	South West	0.4185	0.5800	-0.8648	0.5513
11	Wales	4.1689*	6.1912*	-1.9815	5.7016*
12	Scotland	5.3266*	6.9585*	-1.6350	5.6755*
13	Northern Ireland	4.5107*	6.1256*	-1.5744	5.3359*
14	UK	0.4378	0.5039	-0.7339	0.4415
	Critical values at 5% level of significance	3.64	2.505	-2.22	4.51
Note : * indicates stationarity at 5 % level of significance					

The results in Table 5.3 suggest that at 5% level of significance series for nine out of fourteen series display non-linear stationarity. Most regions that reveal asymmetric ESTAR (AESTAR), ESTAR or LSTAR dynamics are also confirmed to have STAR-type stationarity by the general STAR test; for the exception of the Outer South East (Region 7) which is specified to be following asymmetric ESTAR stationarity but not

confirmed by the general STAR test. Following the results of non-linear stationarity tests displayed in the table above, the appropriate STAR models are estimated for each specific region.

Linear and non-linear model estimation and forecasting

Following the unit root tests performed in this chapter, appropriate non-linear STAR-type models were estimated for the nine out of fourteen series to have exhibited stationarity. The random walk model of house price returns and simple regression of house price returns with price-income ratio as a determinant variable were estimated as a linear benchmark comparable to the non-linear estimation results.

The linear regression for house price returns, y_t, using the price-income ratio as explanatory variable, x_{it}, and a random error term, u_t, as follows:

$$y_t = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \cdots + \alpha_i x_{it} + u_t \tag{5.19}$$

Moreover, this chapter applies a forecasting exercise using non-linear STAR-type models as an error-correction term within the error-correction framework. Thus, the error-correction methodology takes on following forms for ESTAR (5.20), LSTAR (5.21) and AESTAR (5.22) models:

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 - exp(-\gamma(s_{t-d} - c)^2/\sigma^2(s_{t-d}))\big) + \varepsilon_t \quad (5.20)$$

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\big(1 + exp(-\gamma(s_{t-d} - c)/\sigma(s_{t-d}))\big)^{-1} + \varepsilon_t \quad (5.21)$$

$$r_t = (\pi_0 + \pi_1 s_{t-1}) + (\theta_0 + \theta_1 s_{t-1})\Big(1 + exp\big(-\gamma_1^2 s_{t-1}^2 I_t - \gamma_2^2 s_{t-1}^2 (1 - I_t)\big)\Big)^{-1} + \varepsilon_t \quad (5.22)$$

where s_{t-d} is a transition variable within the transition function $F(s_{t-d})$, π_i and θ_i are the autoregressive components of the model, d is the delay parameter, γ, $\gamma_1$and γ_2 are different speeds of adjustment, and ε_t is an error term. In addition, the indication function for AESTAR model depends on the sign of the transition variable:

$$I_t = 1 \; if \; s_{t-1} > 0 \quad (5.23)$$
$$I_t = 0 \; if \; s_{t-1} \leq 0$$

Thus, further to model estimation, a forecasting exercise was performed in the form of a recursive one-step ahead out-of-sample forecast. The main sample of 30 years of quarterly data of house price returns and price-income ratio over the period from 1974:01 to 2004:04 containing 124 observations in total was split into in-sample of fifteen years from 1974:01 to 1989:02, and out-of-sample of fifteen years from 1989:03 to 2004:04, consisting of 62 observations in each sample.

Forecasting accuracy tests

A number of forecasting accuracy tests were performed on the results obtained from the linear and STAR model forecasts. The chosen forecasting accuracy tests are comparable and thus allow the identification of the superior forecast for each set of data. The statistical loss tests included standard functions such as ME, MAE and RMSE, as well as the Diebold-Mariano test of equal forecasting accuracy, forecast encompassing test, and combination forecast tests. In addition, a simple trade rule procedure was used as an economic loss function test of accuracy of forecast.

ME, MAE, and RMSE

The results of initial tests for the random walk and linear regression forecasts are in the table below (Table 5.4). As indicated in the table, region 8 of Outer Metropolitan London seems to produce the most accurate statistics overall, followed by region 9 of London. While the linear regression forecast for region 2 of Yorkshire and Humberside produces the highest value for the trade rule test suggesting the highest speculative profit, followed closely by region 13 of Northern Ireland also generated by a linear regression forecast. The highest trade rule values produced by the random walk model belong to regions 3 and 11 of the North West and Wales respectively. Overall, the forecasting tests statistics differ only marginally, mostly displaying very similar outcomes for all the regions with no clear preference for either of the linear models.

Table 5.4. Random walk and linear regression models forecast statistics.

	Region	**ME**		**MAE**		**RMSE**		**Trade rule**	
		Random Walk	***Linear***	***Random Walk***	***Linear***	***Random Walk***	***Linear***	***Random Walk***	***Linear***
1	North	0.0274	-0.0040**	0.0396	0.0607	0.0481	0.0706	0.0056	0.0045
2	Yorkshire and Humberside	0.0271	0.0036	0.3722	0.0501	0.0465	0.0603	0.0026	0.0138*
3	North West	0.0253	0.0191	0.0310	0.0373	0.0387	0.0445	0.0093*	0.0037
4	East Midlands	0.0248	0.0206	0.0307	0.0336	0.0374	0.0405	0.0049	0.0053
5	West Midlands	0.0217	0.0813	0.0301	0.0340	0.0353	0.0400	0.0055	0.0033
6	East Anglia	0.0220	0.0205	0.0317	0.0324	0.0375	0.0381	0.0051	0.0059
7	Outer South East	0.0213	0.0215	0.0290	0.0300	0.0349	0.0360	0.0061	0.0070
8	Outer Metropolitan London	0.0175*	0.0177	0.0241*	0.0247*	0.0305*	0.0311*	0.0043	0.0053
9	London	0.0182	0.0195	0.0270**	0.0281**	0.0326**	0.0337**	0.0054	0.0078
10	South West	0.0222	0.0218	0.0315	0.0326	0.0360	0.0378	0.0051	0.0081
11	Wales	0.0295	0.0070	0.0364	0.0492	0.0452	0.0581	0.0068**	0.0037
12	Scotland	0.0178**	0.0011*	0.0291	0.0343	0.0369	0.0409	0.0054	0.0057
13	Northern Ireland	0.0219	0.0233	0.0293	0.0305	0.0354	0.0364	0.0064	0.0107**
14	UK	0.0230	0.0187	0.0286	0.0306	0.0334	0.0359	0.0062	0.0141

Note : * indicates the best statistic

** indicates the second best statistic

STAR models forecasts were assessed in the same fashion with results of accuracy tests provided in the tables below (Table 5.5 – 5.7). The statistics reveal a very similar pattern to linear results in terms of difficulty of determining clear preference for a specific model. However, the values of statistics in general seem to be fractionally better compared to those of linear forecasts, with the exception of the LSTAR forecast for region 11 of Wales which demonstrates the least favourable statistics results across all the linear and non-linear forecasts, yet producing a positive trade rule result.

Table 5.5. Asymmetric STAR model forecast statistics.

	Region	**ME**	**MAE**	**RMSE**	**Trade**
1	North	0.0002*	0.0251	0.0331	0.0140
2	Yorkshire and Humberside	0.0063	0.0262	0.0348	0.0125
3	North West	0.0084	0.0177*	0.0251**	0.0147
6	East Anglia	0.0091	0.0255	0.0308	0.0191**
7	Outer South East	0.0059	0.0183**	0.0245*	0.0220*
9	London	0.0056	0.0240	0.0306	0.0168
11	Wales	0.0115	0.0256	0.0326	0.0135
12	Scotland	0.0053**	0.0230	0.0310	0.0059
13	Northern Ireland	0.0132	0.0247	0.0309	0.0142

Note : * indicates the best statistic
** indicates the second best statistic

Table 5.6. ESTAR model forecast statistics.

	Region	**ME**	**MAE**	**RMSE**	**Trade**
1	North	0.0073	0.0460	0.0521	0.0097*
2	Yorkshire and Humberside	-0.0015*	0.0366*	0.0456*	0.0080
Note : * indicates the best statistic					

Table 5.7. LSTAR model forecast statistics.

	Region	**ME**	**MAE**	**RMSE**	**Trade**
1	North	-0.0276	0.0805	0.1420	0.0100**
2	Yorkshire and Humberside	0.0189**	0.0393	0.0528	0.0069
11	Wales	0.1264	0.1997	0.3045	0.0131*
12	Scotland	-0.0021*	0.0359**	0.0432**	0.0056
13	Northern Ireland	0.0219	0.0315*	0.0381*	0.0021
Note : * indicates the best statistic ** indicates the second best statistic					

Diebold-Mariano tests

The Diebold-Mariano test of equal forecasting accuracy (Diebold and Mariano, 1995) is designed to test whether the differences in MSEs of competing forecasts are statistically significant. Thus, assessing whether lower values of MSEs of one forecast are significant enough to validate the superiority of that forecast over competing alternatives. The null hypothesis of equal forecast accuracy is tested against an alternative hypothesis using the following test statistic:

$$S_1 = \left[\hat{V}(\bar{d})\right]^{-\frac{1}{2}}\bar{d} \tag{5.24}$$

where $\bar{d}$ is the mean of the coefficient d_t which is the difference between the sets of squared forecast errors from two competing models, $d_t = e_{1t}^2 - e_{2t}^2$, and $\hat{V}(\bar{d})$ is an estimate of the variance of $\bar{d}$.

The modified Diebold-Mariano statistic by Harvey et al. (1997) is more robust for two or more steps ahead and characterised with the ease of using the Student's *t*-test critical values as opposed to the standard distribution statistics (5.25).

$$S_1^* = \left[\frac{t + 1 - 2h + t^{-1}h(h-1)}{t}\right]^{-\frac{1}{2}} S_1 \tag{5.25}$$

where S_1 is the original Diebold-Mariano test statistic for h steps ahead forecast for time *t*.

The results of both Diebold-Mariano tests, the standard and modified, performed on the data revealed no significant statistics, thus rejecting H_0 of equal forecasting accuracy for all series suggesting that the differences in MSEs between competing forecasts are statistically insignificant (Table 5.8).

Table 5.8. Diebold-Mariano test statistics.

	Region		DM statistic	DM modified
1	North	RW/ linear	-0.0769	-0.0767
		RW/ AESTAR	0.0333	0.0332
		RW/ ESTAR	-0.1333	-0.1330
		RW/ LSTAR	-0.1692	-0.1688
		Linear/ AESTAR	0.9375	0.9353
		Linear/ ESTAR	-0.0364	-0.0363
		Linear/ LSTAR	-0.1692	-0.1688
2	Yorkshire and Humberside	RW/ linear	-0.0843	-0.0841
		RW/ AESTAR	0.3076	0.0306
		RW/ ESTAR	0.0869	0.0867
		RW/ LSTAR	0.0078	0.0078
		Linear/ AESTAR	0.5000	0.4988
		Linear/ ESTAR	0.1052	0.1050
		Linear/ LSTAR	0.0263	0.0262
3	North West	RW/ linear	-0.1876	-0.1871
		RW/ AESTAR	0.0217	0.0216
		Linear/ AESTAR	0.0688	0.0686
6	East Anglia	RW/ linear	-0.0400	-0.0399
		RW/ AESTAR	0.2272	0.2267
		Linear/ AESTAR	0.2272	0.2267
7	Outer South East	RW/ linear	-0.0765	-0.0763
		RW/ AESTAR	0.4666	0.2267
		Linear/ AESTAR	0.4666	0.2267
9	London	RW/ linear	-0.1135	-0.1132
		RW/ AESTAR	0.4000	0.3990
		Linear/ AESTAR	0.4000	0.3990
11	Wales	RW/ linear	-0.1000	-0.0997
		RW/ AESTAR	0.2857	0.2850
		RW/ LSTAR	-0.1397	-0.1394
		Linear/ AESTAR	0.4000	0.3990
		Linear/ LSTAR	-0.1353	-0.1350
12	Scotland	RW/ linear	-0.1192	-0.1189
		RW/ AESTAR	0.2000	0.1995
		RW/ LSTAR	-0.1168	-0.1166
		Linear/ AESTAR	0.1666	0.1666
		Linear/ LSTAR	-0.0916	-0.0913
13	Northern Ireland	RW/ linear	-0.0520	-0.0518
		RW/ AESTAR	0.1348	0.1345
		RW/ LSTAR	-0.1923	-0.1918
		Linear/ AESTAR	0.1111	0.1108
		Linear/ LSTAR	-0.1481	-0.1478
Note: RW – random walk				

Forecast encompassing tests

The forecast encompassing procedure tests whether two competing forecasts contain different additional information that is required to forecast the main variable. If one model does not contain such information it is said to be encompassed in the forecast produced by the other model. In the situation of both models contributing independent information towards forecasting of the variable, a combination of those forecasts might be considered. The simple version of the forecast encompassing test regresses the dependent variable Y_t on the forecasted values from both competing models $\hat{Y}_{t-s,t}^{(I)}$ and $\hat{Y}_{t-s,t}^{(II)}$ (5.26). The null hypothesis of the first model's forecast encompassing the forecast of the second, $\alpha_1 = 1$, is tested against the alternative of $\alpha_2 = 0$. The values of the coefficients will be reversed in the case of the first model's forecast being encompassed in the second, $\alpha_2 = 1$, $\alpha_1 = 0$. Any other outcome of the test will signify that neither of the models encompass the other. Non-zero values of coefficients will indicate that both models contain independent information and are required for the forecasting of the dependent variable, while zero coefficients will suggest that neither of the models contain information relevant to forecasting the variable.

$$Y_t = \alpha_1 \hat{Y}_{t-s,t}^{(I)} + \alpha_2 \hat{Y}_{t-s,t}^{(II)} + u_t \tag{5.26}$$

In addition, this paper will run the encompassing test of forecasting errors whereby regression is performed on the forecasting errors of the competing models, based on the approach adopted by Shiller (1990).

$$Y_t - Y_{t-s} = \alpha + \beta_1\left(\hat{Y}_{t-s}^{(I)} - Y_{t-s}\right) + \beta_2\left(\hat{Y}_{t-s}^{(II)} - Y_{t-s}\right) + u_t \tag{5.27}$$

where $\hat{Y}_{t-s}^{(I)}$ is the forecast of the dependent variable Y_t made from one of the competing forecasting models, and $\hat{Y}_{t-s}^{(II)}$ is the forecast of Y_t from the alternative forecasting model. The null hypothesis of the forecasts made by the first model contain no relevant information for forecasting the, Y_t, $(H_0: \beta_1 = 0)$, which is tested against the hypothesis that the alternative model contains no relevant information $(H_1: \beta_2 > 0)$. This study will not exercise the restriction of the sum of the coefficients being equal to unity, since, according to Fair and Shiller (1990), both models generating noise will result in both coefficient estimates to be zero, while both models containing independent information will generate the sum of coefficients equal to two. Table 5.9 demonstrates the results of forecast encompassing tests of STAR models against linear alternatives, random walk and linear regression forecasts.

Table 5.9. Forecast encompassing tests.

		Forecasting encompassing		**Forecasting errors encompassing**	
		t-statistic for β_1	t-statistic for β_2	t-statistic for β_1	t-statistic for β_2
1	**North**				
	RW/ AESTAR	1.0426	4.8563*	-144.7705*	0.0014
	RW/ ESTAR	1.5387	0.6945	-65.4633*	-4.7573
	RW/ LSTAR	1.4295	-0.5407	-168.3195*	3.5231*
	RW/ Linear	2.1478*	1.5688	-85.4498*	-11.3113
	Linear/ AESTAR	-0.3738	4.9105*	-21.6589*	-0.6897
	Linear/ ESTAR	-0.0976	0.0463	-5.0849*	-1.9252
	Linear/ LSTAR	-0.0672	0.1017	-23.4243*	-1.1878
2	**Yorkshire and Humberside**				
	RW/ AESTAR	-1.0135	1.5819	-134.9165*	-5.9216
	RW/ ESTAR	0.5995	3.3491*	-85.9780*	-1.4965
	RW/ LSTAR	0.9693	2.9024*	-94.5831*	-2.6104
	RW/ Linear	0.1307	-0.1235	-36.1964*	-7.5644
	Linear/ AESTAR	-0.4636	3.7033*	-13.7166*	1.9211
	Linear/ ESTAR	-0.9814	3.4550*	-23.2782*	-0.0797
	Linear/ LSTAR	-0.5456	2.7842*	-24.5957*	-0.0637
3	**North West**				
	RW/ AESTAR	-0.8272	0.7666	-82.2837*	-6.4145
	Linear/ AESTAR	1.3791	0.2746	-2.7083*	-1.8338
	RW/ Linear	4.8400*	4.1662*	-101.2296*	-15.2907
6	**East Anglia**				
	RW/ AESTAR	-0.1541	5.0364*	-85.0743*	-1.1247
	Linear/ AESTAR	-0.1544	5.0964*	-76.7021*	-0.4764
	RW/ Linear	-0.6154	-1.0360	-20.8630*	-21.4799

7	**Outer South East**				
	RW/ AESTAR	0.3688	8.0841*	-78.9998*	-0.3978
	Linear/ AESTAR	-0.1569	7.9797*	-120.8161*	-1.0119
	RW/ Linear	0.2036	-0.8351	-9.9626*	-19.4126
9	**London**				
	RW/ AESTAR	0.0707	6.4999*	-59.6645*	-0.0500
	Linear/ AESTAR	-0.7560	6.4542*	-119.1018*	-0.5699
	RW/ Linear	1.1888	-1.6545	-0.8286	-12.8206
11	**Wales**				
	RW/ AESTAR	1.2857	4.8765*	-68.6186*	-0.3374
	RW/ LSTAR	1.8069	0.3767	-123.8822*	-0.1672
	Linear/ AESTAR	-1.4094	5.2555*	-14.3058*	-2.1701
	Linear/ LSTAR	-0.9137	0.3474	-27.4761*	-1.4671
	RW/ Linear	1.8912	1.0751	-73.89599*	-14.9462
12	**Scotland**				
	RW/ AESTAR	0.4136	2.6716*	-65.3268*	1.2485
	RW/ LSTAR	1.2209	-0.0722	-64.4207*	-0.8648
	Linear/ AESTAR	-0.1514	2.4750*	-28.2226*	-6.7647
	Linear/ LSTAR	-1.5103	0.1072	-20.7665*	-1.6093
	RW/ Linear	0.4185	-0.9702	-32.9223*	-8.4447
13	**Northern Ireland**				
	RW/ AESTAR	1.4668	-0.2465	-13.9527*	-5.0907
	RW/ LSTAR	1.9974*	-2.1008	-30.9479*	0.4612
	Linear/ AESTAR	1.0767	0.1668	-13.4565*	-11.9912
	Linear/ LSTAR	1.3695	-1.8060	-18.8365*	-1.5611
	RW/ Linear	0.9821	0.1582	-9.4777*	0.5966

Note: * significant at 5%

RW – random walk

The results of forecast encompassing tests (Table 5.9) reveal that STAR models for most of the series encompass linear alternatives, with the exception of Northern Ireland (Region 13) where random walk encompasses the LSTAR model. According to the results of the forecast encompassing test for the North (Region 1) AESTAR is not encompassed by neither random walk nor linear regression models, while linear regression is encompassed by the random walk; for Yorkshire and Humberside (Region 2) AESTAR, ESTAR and LSTAR models encompass both linear alternatives; both random walk and linear regression forecasts seemed to contain independent information for the North West (Region 3); AESTAR model forecast was not encompassed by neither linear alternative for series of East Anglia (Region 6), Outer South East (Region 7), London (Region 9), Wales (Region 11), and Scotland (Region 12); only the random walk seems to contain an independent information toward the forecasting of Northern Ireland (Region 13) series.

The forecasting errors encompassing test, on the other hand, suggests that none of the STAR models encompass linear alternatives. In addition, the tests indicate that for most series random walk and linear models forecasts errors explain the forecasting variable, with the exception of London (Region 9) where neither the random walk model nor linear regression forecasting errors seem to contain independent information for forecasting the dependent variable.

By illuminating the models that contain information which is encompassed in other forecasts, assuming it can be utilised by replacing such forecasts with the superior alternatives, the combination of a random walk and STAR models can be confirmed as the optimal forecast encompassing the information contained in the linear regression forecasts for all series.

Combined forecast

Subsequent to the results of the forecast encompassing tests and as an additional mode of testing forecasting accuracy, all series for house prices were forecasted using combined forecast methodology. The forecasts are combined using the simple method of arithmetic average of combining weights which has been proven to be robust and reasonably accurate (5.28).

$$f_c(y) = \frac{1}{k}\sum_{i=1}^{k} f_i \tag{5.28}$$

where $f_c(y)$ is the equal weighting combined forecast of $f_1, f_2, \dots, f_k$ which are the forecasts for the dependent variable, y. The methodology is applied to generate forecasts in combinations of the random walk, linear regression and STAR models; random walk and STAR model; and random walk and linear regression.

Table 5.10. Combination forecasts statistics.

	Region	ME	MAE	RMSE	Trade
1	**North**				
	RW/ Linear/ AESTAR	0.0094*	0.0417	0.0506	0.0120**
	RW/ Linear/ ESTAR	0.0268	0.0431	0.0520	0.0004
	RW/Linear/ LSTAR	0.0805	0.0970	0.1707	-0.0011
	RW/ AESTAR	0.0179**	0.0267*	0.0330*	0.0131*
	RW/ ESTAR	0.0303	0.0372**	0.0455**	0.0077
	RW/ LSTAR	0.0391	0.0620	0.1057	-5.02E-05
	RW/Linear	0.0229	0.0406	0.0487	0.0056
2	**Yorkshire and Humberside**				
	RW/ Linear/ AESTAR	0.0066**	0.0320**	0.0415**	0.0134
	RW/ Linear/ ESTAR	0.0124	0.0352	0.0445	0.0082*
	RW/Linear/ LSTAR	0.0149	0.0368	0.0471	0.0088**
	RW/ AESTAR	0.0062*	0.0289*	0.0377*	0.0141
	RW/ESTAR	0.0129	0.0329	0.0418	0.0095
	RW/ LSTAR	0.0142	0.0343	0.0435	0.0124
	RW/Linear	0.0196	0.0392	0.0484	0.0129
3	**North West**				
	RW/Linear/ AESTAR	0.0163**	0.0260**	0.0334**	0.0140**
	RW/ AESTAR	0.0154*	0.0225*	0.0295*	0.0132
	RW/Linear	0.0226	0.0309	0.0389	0.0154*
6	**East Anglia**				
	RW/Linear/ AESTAR	0.0148	0.0331**	0.0395**	0.0129*
	RW/ AESTAR	0.0089**	0.0240*	0.0296*	0.0183
	RW/Linear	0.0088*	0.0469	0.0537	0.0130**
7	**Outer South East**				
	RW/Linear/ AESTAR	0.0050**	0.0210*	0.0272**	0.0192**
	RW/ AESTAR	0.0045*	0.0173*	0.0236*	0.0216*
	RW/Linear	0.0095	0.0455	0.0525	0.0120
9	**London**				
	RW/Linear/ AESTAR	0.0021*	0.0232**	0.0291**	0.0155
	RW/ AESTAR	0.0030**	0.0215*	0.0274*	0.0172**
	RW/Linear	0.0154	0.0346	0.0410	0.0076*
11	**Wales**				
	RW/Linear/ AESTAR	0.0092*	0.0287**	0.0364**	0.0098
	RW/ Linear/ LSTAR	0.0248	0.0420	0.0600	0.0092
	RW/ AESTAR	0.0114**	0.0261*	0.0332*	0.0142**
	RW/ LSTAR	0.0266	0.0379	0.0525	0.0118*
	RW/Linear	0.0233	0.0370	0.0462	0.0094
12	**Scotland**				
	RW/Linear/ AESTAR	0.0063**	0.0382	0.0483	0.0034
	RW/ Linear/ LSTAR	0.2327	0.2439	0.2706	0.0037**
	RW/ AESTAR	0.0080	0.0350**	0.0461**	0.0070*
	RW/ LSTAR	0.0217	0.0333*	0.0420*	0.0032
	RW/Linear	-0.0019*	0.0492	0.0582	0.0004

13	**Northern Ireland**				
	RW/Linear/ AESTAR	0.0265	0.0345	0.0431	0.0118*
	RW/ Linear/ LSTAR	0.0186**	0.0331	0.0435	0.0093
	RW/ AESTAR	0.0156*	0.0256*	0.0323*	0.0143**
	RW/ LSTAR	0.0186**	0.0271**	0.0342**	0.0087
	RW/Linear	0.0201	0.0282	0.0347	0.0039
Note : * signifies the value of the best statistic					

Following the results of combined forecasts statistical tests (Table 5.10), on the whole, combination forecasts containing the asymmetric ESTAR (AESTAR) model seem to produce the best performance in terms of forecasting accuracy as well as generating the highest trade rule results. Combinations of linear and random walk models do not generate even a marginal preference in comparison to other combinations of forecasts. Instead these are over performed by combinations of linear, random walk and STAR-type models. Moreover, overall, taking into account test statistics for all the forecasts, including linear and non-linear individual forecasts, asymmetric ESTAR (AESTAR) model forecasts and a combination of random walk and AESTAR models seem to produce most favourable statistics including ME, MAE, RMSE and trade rule. For nine individual non-linear stationary house prices AESTAR and combinations including AESTAR model generated the best statistics for forecasting accuracy test.

To conclude, based on the results obtained in this chapter, the preferred forecasting model for linearly stationary series is the random walk model (East Midlands, Region 4; West Midlands, Region 5; Outer Metropolitan London, Region 8; South West, Region 10; UK, Region 14), whereas the preferred forecasting model for non-linear stationary series is the AESTAR model (North, Region 1; Yorkshire and Humberside, Region 2; North West, Region 3; East Anglia, Region 6; Outer South East, Region 7; London, Region 9; Wales, Region 11; Scotland, Region 12; Northern Ireland, Region 13). The

AESTAR model generates the best forecasting tests statistics and highest trade rule results across the regions characterised by non-linear dynamics, followed closely by the combination of random walk and AESTAR models.

It is evident from the results presented in this chapter that house price returns are forecastable using a price-income ratio as a measure of affordability in order to establish the level of fundamental prices. In addition, the majority of the UK housing regions considered in this study exhibit non-linear adjustments to the equilibrium, which can be successfully forecasted with an application of STAR-type models, in particular asymmetric ESTAR (AESTAR).

5.5. Conclusion

This chapter applied the present value model and stock market approach to UK housing market data with the intent to carry out an econometric forecast of house price returns using non-linear modelling, in particular STAR-type models. Error-correction methodology was used to forecast house price returns using price-income ratio. The methodology was based on a procedure proposed by Black et al. (2005) which involved testing the stationarity of the price-income ratio as a measure of affordability in order to determine fundamental levels of house prices. Results reported by Black et al. (2005) found the price-income ratio to be non-stationary on the basis of the standard augmented Dickey-Fuller (ADF) test, which in turn suggested non-predictability of the house prices. However, it can be argued that such results imply that log house prices

and log income are simply not cointegrated with a vector $[1, -1]$, rather than complete absence of cointegration between these variables. These results could be due to presence of the non-fundamental components within the price-income ratio such that the transversality condition used to derive the ratio might not hold. Hence, these factors might be the reason for linear models to fail. In addition, many researchers have argued in the context of the stock market that the log dividend-price ratio remains stable in the long-run, despite temporary deviations from the equilibrium relationship as a result of the presence of non-fundamental components in the linear present value model. Such behaviour implies the presence of non-linear dynamics within the dividend-price relationship. Similar dynamics are observed in the housing market in the relationship between house prices and real income. Furthermore, researchers attribute a non-fundamental component to the presence of bubbles in both financial and housing markets, hence attempting to incorporate bubble dynamics into various non-linear modelling approaches (financial markets: Van Norden and Vigfusson, 1998; Bohl and Siklos, 2004; housing markets: Hall et al., 1997; Black et al., 2006; Goodman and Thibideau, 2008; Coleman et al., 2008). It must be pointed out that in most cases studies that intend to explain the presence of bubbles in stock prices mainly focus on so called 'rational' bubbles, or as described by Evans (1991), speculative, periodically collapsing bubbles. However, it appears that the presence of bubbles presents researchers with main problems of empirical identification of bubbles and theoretical doubts of the existence of rational bubbles.

Attempts to understand these concepts have led to the development of theories of behavioural finance, which examine the notion of market sentiment. The basis of behavioural finance is the argument that some financial phenomena can be understood

and explained assuming that some financial market participants are not fully rational (Barberis and Thaler, 2003). These theories particularly concentrate on the discussion of results of interaction between rational and irrational, or noise, traders. Behavioural finance theories argue that noise traders trade on the basis of predictive expectations, or in other words, trend-chasing. For instance, a price momentum trading strategy is based on the notion that the current trend in the financial market will continue. This can lead to investors' underreaction to the arrival of new information (Reilly and Brown, 2003). However, positive news in momentum trading can lead to overreaction, so that the change in price will exceed the actual price required by the news.

There is a possibility that the behavioural finance theories can be applied to the housing market participants, in the sense that the housing prices might also be driven by trend-chasing activities. Moreover, Case and Shiller (2004) found that, according to their survey, a majority of housing market participants hold the view that the market is not driven by psychology, while at the same time revealing to form their investment decisions based on expectations and word of mouth information. However, applications of behavioural finance theories to housing markets are yet to be addressed in the literature. In contrast, housing market literature comprises of ample research into the subject of presence of bubbles in house prices. Parallel to the financial market, the housing market is characterised by specific booms and bursts dynamics, which differently to the stock market, are exemplified by sluggish responses and slow mean reversion.

This chapter has provided an overview of academic literature on a general discussion and the subject of modelling of housing markets concentrating on performing a forecasting exercise of house price returns using non-linear models. The error-

correction methodology was applied to real house prices and price-income ratio as a determinant variable. The current study employed real income as a measure of affordability and individual wealth as a main determinant of house price suggested by previous studies on housing markets (Case and Shiller, 2004; Fraser et al., 2008; Fraser et al., 2009). The price-income ratio used in this study was based on the ratio proposed by Black et al. (2005) as a measure of affordability to determine fundamental price levels to use in conjunction specifically with housing prices. However, distinct from the approach employed by Black et al. (2005) of using conventional unit root tests to assess stationarity of the price-income ratio, this study applies non-linear stationarity tests specifically designed to test non-linear STAR-type stationarity. This study presumes that this main distinction from the Black's et al. (2005) approach in terms of using non-linear unit root tests and incorporating non-linear error-correction methodology enabled to reveal non-linear asymmetric adjustment of the housing prices. Thus, the error-correction methodology combined with STAR-type models was applied to quarterly house prices and real income data in order to generate a recursive out-of-sample one-step ahead forecast. The results of non-linear forecasts were compared to linear benchmarks in the form of a random walk model and a linear regression using a number of tests of forecasting accuracy. The empirical results revealed that while linear models performed well, STAR models seemed to perform marginally better. The random walk model was found to produce the best forecasting statistics for five linearly stationary series, with the asymmetric ESTAR (AESTAR) model generating the best forecast for all nine non-linear stationary series.

The objective of this study was to offer an extension of empirical evidence into non-linear out-of-sample forecasting of house prices which appears to be lacking such type

of investigation. The results of this study imply that the asymmetric ESTAR (AESTAR) model has generated the best overall performance in out-of-sample forecasting of the UK house prices, thus confirming the presence of asymmetric adjustment suggested in previous studies (Crawford and Fratantoni, 2003; Black et al., 2006; Miles, 2008; Gao et al., 2009). According to the results obtained here, the asymmetry in the quarterly housing price time-series is apparent to the extent of being able to utilise non-linear AESTAR modelling. This asymmetry is consistent with the housing market being characterised by a slow speed of adjustment due to considerable transaction costs and borrowing constraints. Moreover, it is possible that the slow asymmetric adjustment is due to the slow reaction to market changes which can be explained by the lack of professional arbitragers in the housing market who would otherwise ensure the correction of profitable deviations from the fundamental price levels.

Chapter 6
Summary and conclusion

The objective of this thesis was to examine time-series forecasting methodology and to extend analysis of predictability of financial assets using a non-linear approach. Chapter 2 offered an extensive literature review into time-series forecasting emphasising the importance of econometric modelling and forecasting techniques in a wide range of disciplines for varied market participants and policy-makers.

Following a non-linear forecasting methodology described in Section 2.3, this study presented three empirical chapters, where Chapter 3 assessed predictability of daily stock returns and forecasting abilities of non-linear models; Chapter 4 focused on detecting and forecasting non-linear dynamics within the price-dividend relationship of monthly stock returns by applying non-linear error-correction framework, further developing the research into long-horizon predictability; and Chapter 5 extended the research of financial assets predictability to the housing market and applied non-linear equilibrium methodology to monthly house prices for a number of UK regions.

Chapter 3 focused on examining forecasting abilities of non-linear smooth transition autoregressive (STAR) models using daily stock returns. Consistent with the previous literature (Abhyankar et al., 1995; Clements and Smith, 1999, 2001; McMillan, 2001; Lekkos and Milas, 2004; McMillan, 2004; Teräsvirta et al., 2005), the results confirmed the presence of predictability and STAR-type non-linearity within daily returns for FTSE, S&P, DAX and Nikkei indices. Moreover, the presence of STAR-type non-

linearity also seems to be consistent with the notion of the presence of market frictions (Martens et al., 1998; Kapetanios et al., 2003; McMillan, 2005). However, while STAR models performed well in forecasting exercises compared to the linear benchmark in terms of forecasting accuracy, the random walk model seemed to outperform non-linear alternatives in terms of simplicity of application whilst not compromising the accuracy of the forecast at the same time. Thus, following the results of the Chapter 3, the random walk can be recommended as the preferred model for forecasting stock returns on the daily level in terms of the combination of model's forecasting accuracy and its use in practical applications.

Building on the present value model approach, Chapter 4 (Section 4.4) applied non-linear error-correction methodology to forecasting the monthly price returns for FTSE, S&P, DAX and Nikkei indices using dividend yield and price-earnings ratio. Similar to the studies by McMillan and Speight (2006) and McMillan (2007), the empirical results confirmed non-linear predictability of monthly returns using STAR-type models. Determinants of the stock price, namely the dividend yield and price-earnings ratio, performed equally well in forecasting the returns with no clear preference for either of the variables. However, the best forecasting performance was achieved by a combined forecast of the random walk and STAR models. These results suggest that while the STAR models are able to capture asymmetric cyclical behaviour of returns series, random walk model compensates for the absence of non-linear adjustments in the periods of calm financial markets. Thus, for different states of the financial market characterised by either asymmetric cycles or periods of tranquillity, the combination of a random walk and non-linear STAR models forecast seems to be most preferable.

In addition, following research suggesting that stock predictability increases with the horizon (Fair and Shiller, 1990; Montgomery et al., 1998, Kim et al., 2005), Chapter 4 (Section 4.5) included an investigation of the predictability of long-horizon stock returns. The study applied a buy-and-hold strategy where the stock is assumed to be held for periods of three, six and twelve months before selling. The approach is different to the previous studies in terms of applying an out-of-sample forecast, as opposed to an in-sample prediction. The findings confirm the suggestion of improved long-horizon predictability, as the forecasts generate stronger results in terms of forecasting accuracy compared to monthly forecasts. Moreover, similarly to the results of monthly data forecasts, the combination of a random walk model and non-linear STAR model appears to be favoured.

Chapter 5 applied the financial market approach of the present value model to the housing market by utilising a non-linear error-correction methodology. The study approached the topic of house price predictability somewhat differently to the previous studies (Black et al., 2005; Black et al., 2006; Goodman and Thibideau, 2008) in terms of employing a non-linear framework to test for stationarity and application of STAR-type models to an out-of-sample forecasting exercise. The results confirmed the presence of non-linearity within house prices. Moreover, the findings revealed the asymmetric ESTAR (AESTAR) model as the most successful model in terms of forecasting performance. The results demonstrated a clear preference for the AESTAR model for forecasting house price returns using the price-income ratio, hence confirming the assumption of slow asymmetric mean reversion of the house prices suggested by previous studies (Holly and Jones, 1997; Crawford and Fratantoni, 2003; Black et al., 2006; Miles, 2008; Gao et al., 2009).

Clements et al. (2004) pointed out that the ability of a model to generate a good in-sample fit does not always translate into a good out-of-sample forecasting performance. In addition, Kanas and Yannopoulos (2001) highlighted the importance of including non-linear terms in out-of-sample forecasting in order to improve forecasting accuracy. Hence, this investigation focused on out-of-sample forecasting performances of non-linear STAR-type models. It is evident that financial as well as housing markets are characterised with non-linear predictability with STAR-type models providing adequate forecasting of these dynamics. The study confirmed an increase of predictability with the horizon and successful application of STAR models to generate accurate forecasts. These findings should be of interest to policy-makers and market participants concerned with long-horizon economic and financial forecasts, which could assist in examining and predicting possible cyclical trends or the state of economy. The results also suggest a strong presence of asymmetry in housing markets. Following the research by Koetter and Poghosyan (2009), who suggested property prices as an indicator of the overall state of financial and banking sector stability, and Das et al. (2009) and Gao et al. (2009) attributing the recent credit crunch to the burst of the housing bubble, the importance of investigating housing market dynamics cannot be over exaggerated. Therefore, this study emphasises the results of house prices predictability and possible uses of asymmetric non-linear modelling in conjunction with financial market forecasts, as it is apparent that the house prices can be treated as indicators of wealth, and thus can be successfully utilised by policy-makers.

The presence of non-linear dynamics within financial and housing markets is evident throughout the study. Despite the evidence in certain cases of linear models to generate equally accurate forecasts combined with the simplicity of implementation, it is

essential that the presence of non-linearity is not ignored. However, as pointed out by Chatfield (1997), there is a danger of overfitting or fitting an incorrect model due to a wide availability of specialised computer software. Thus, while the superiority of non-linear models is greatly attractive, practitioners must take care when utilising complex models in the context of forecasting.

In conclusion, it is worth to take into account that while econometric forecasting is an invaluable tool for market practitioners and policy makers, it is also, in fact, only an estimation of the future. Moreover, Hendry and Clements (2003) described economic forecasting as a mixture of science and art, while Armstrong and Fildes (2006) recommended an expansion of econometric forecasting techniques to other fields of sciences in order to improve existing forecasts by combining knowledge and developments in other disciplines, including medicine, geology, politics, weather and many others.

References

Abraham, J.M., Hendershott, P.H., 1996. Bubbles in metropolitan housing markets, *Journal of housing research*, 7 (2), pp. 191-207

Abhyankar, A., Copeland, L.S., Wong, W., 1995. Nonlinear dynamics in real-time equity market indices: Evidence from the United Kingdom, *The economic journal*, 105 (431), pp. 864-880

Abhyankar, A., Copeland, L.S., Wong, W., 1997. Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE 100, *Journal of business & economic statistics*, January 1997, 15 (1), pp. 1-14

Aihara, S., Bagchi, A., 2000. Estimation of stochastic volatility in the Hull-White model, *Applied mathematical finance*, 7, pp. 153-181

Alexander, C., 1999. Optimal hedging using cointegration, *Philosophical transactions: Mathematical, physical and engineering sciences*, Mathematics of Finance August 15th 1999, 357 (1758), pp. 2039-2058

Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002. Range-based estimation of stochastic volatility models, *Journal of finance*, 57 (3), pp. 1047-1091

Allen, F., Karjalainen, R., 1999. Using genetic algorithms to find technical trading rules, *Journal of financial economics*, 51, pp. 245-271

Andersen, A., 1977. An application of post-sample theory to time-series forecasting, *Operational research quarterly (1970-1977)*, 28 (1), pp. 93-98

Andersen, T.G., Bolleslev, T., Lange, S., 1999. Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon, *Journal of empirical finance*, 6, pp. 457-477

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2002. Modelling and forecasting realized volatility, *Working paper*

Andersen, T.G., Bolleslev, T., Meddahi, N., 2004. Analytical evaluation of volatility forecasts, *International economic review*, 45 (4), pp. 1079-1110

Ang, A., Bekaert, G., 2007. Stock return predictability: Is it there?, *Review of financial studies*, 20 (3), pp. 651-707

Armstrong, J.S., Fildes, R., 2006. Making progress in forecasting, *International journal of forecasting*, 22, pp. 433-441

Angelidis, T., Benos, A., Degiannakis, S., 2004. The use of GARCH models in VaR estimation, *Statistical methodology*, 1, pp. 105-128

Awartani, B.M.A., Corradi, V., 2005. Predicting the volatility of the S&P 500 stock index via GARCH models: the role of asymmetries, *International journal of forecasting*, 21, pp. 167-183

Baillie, R.T., Bollerslev, T., 1991. Intra-day and intra-market volatility in foreign exchange rates, *The review of economic studies*, Special issue: The econometrics of financial markets, 58 (3), pp. 565-585

Baillie, R.T., Bollerslev, T., 1992. Prediction in dynamic models with time-dependent conditional variances. *Journal of econometrics*, vol.31, pp.307-327

Baillie, R.T., Bollerslev, T., 1994. Cointegration, fractional cointegration, and exchange rate dynamics, *Journal of finance*, 49 (2), pp. 737-745

Baker, M., Wurgler, J., 2002. Market timing and capital structure, *The journal of finance*, February 2002, 57 (1), pp. 1-32

Bali, T.G., Demirtas, K.O., Levy, H., 2008. Nonlinear mean reversion in stock prices, *Journal of banking & finance*, 32, pp. 767-782

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration, *International economic review*, August 1997, 38 (3), pp. 627-645

Bandi, F.M., Russel, J.R., 2003. Microstructure noise, realized volatility, and optimal sampling, *CIRANO Conference Realized Volatility,* 7-8th November, 2003

Barberis, N., Thaler, R., 2003. A survey of behavioural finance Chapter 18, *Handbook of the economics of finance*, The Netherlands: Elsevier B.V., ISBN 0444513620

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the royal statistical society: series B*, 64, pp. 253-280

Barucci, E., Renò, R., 2002. On measuring volatility and the GARCH forecasting performance, *Journal of international financial markets, institutions and money*, 12, pp. 183-200

Batten, J.A., Ellis, C.A., 2005. Parameter estimation bias and volatility scaling in Black-Scholes option prices, *International review of financial analysis*, 14, pp. 165-176

Baum, C.M., Barkoulas, J.T., Caglayan, M., 2001. Nonlinear adjustment to purchasing power parity in the post-Bretton Woods era, *Journal of international money and finance*, 20, pp. 379-399

Bauwens, L., Lubrano, M., 2002. Bayesian option pricing using asymmetric GARCH models, *Journal of empirical finance*, 9, pp. 321-342

Bellini, F., Figà-Talamanca, G., 2005. Runs tests for assessing volatility forecastability in financial time series, *European journal of operational research*, 163, pp. 102-114

Beltratti, A., Morana, C., 2010. International house prices and macroeconomic fluctuations, *Journal of banking & finance*, 34, pp. 533-545

Berkowitz, J., Giorgianni, L., 1996. Long-horizon exchange rate predictability?, *Working Paper*

Berkowitz, J., 2001. Testing density forecasts, with applications to risk management. *Journal of business & economics statistics*, vol.19, no.4, pp.465-474

Bewley, R., Rees, D., Berg, P., 2004. The impact of stock market volatility on corporate bond credits spreads, *Mathematics and computers in simulation*, 64, pp. 363-372

Black, A., Fraser, P., Hoesli, M., 2005. House prices, fundamentals and inflation, *Research Paper N°129*, Université de Genevè

Black, A., Fraser, P., Hoesli, M., 2006. House prices, fundamentals and bubbles, *Journal of business finance & accounting*, 33 (9-10), pp. 1535-1555

Blanchard, O.J., Watson, M.W., 1982. Bubbles, rational expectations and financial markets, *In: Wachtel P. (ed), Crisis in the economic and financial structure,* Lexington: Lexington books, pp. 295-315

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns, *The journal of finance*, December 1992, 47 (5), pp. 1731-1764

Bohl, M.T., 2003. Periodically collapsing bubbles in the US stock market?, *International review of economics and finance*, 12, pp. 385-397

Bohl, M.T., Siklos, P.L., 2004. The present value model of U.S. stock prices redux: a new testing strategy and some evidence, *The quarterly review of economics and finance*, 44, pp. 208-223

Box, G.E.B., Jenkins, J.M., 1970. Time series analysis: forecasting and control*, Holden-Day, San Francisco*

Box, G.E.B., Jenkins, J.M., 1976. Time series analysis: forecasting and control*, Holden-Day, San Francisco*

Brooks, C., 2002. *Introductory Econometrics for Finance*. United Kingdom: Cambridge University Press, ISBN 0 521 79367 X

Brooks, C., Katsaris, A., 2003. Rational speculative bubbles: an empirical investigation of the London stock exchange, *Bulletin of economic research*, 55 (4), pp. 0307-3378

Brown, J.P., Song, H., McGillivray, A., 1997. Forecasting UK house prices: a time varying coefficient approach, *Economic Modelling*, 14, pp. 529-548

Bulkley, G., Harris, R.D., 1997. Irrational analysts expectations as a cause of excess volatility in stock prices, *The Economic Journal*, 107 (441), pp. 359-371

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. *The econometrics of financial markets*, Princeton University Press, ISBN 0691043019

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models, *The journal of political economy*, October 1987, 95 (5), pp.1062-1088

Campbell, J.Y., Shiller, R.J., 1988a. Interpreting cointegrating models, *Journal of economic dynamics and control*, 12, pp. 505-522

Campbell, J.Y., Shiller, R.J., 1988b. The dividend-price ratio and expectations of future dividends and discount factors, *The review of financial studies*, Autumn 1988, 1 (3), pp. 195-228

Campbell, J.Y., Shiller, R.J., 1998. Valuation ratios and the long-run stock market outlook, *Journal of portfolio management*, Winter 1998, 24, pp. 11-26

Campbell, J.Y., Shiller, R.J., 2001. Valuation ratios and the long-run stock market outlook: an update, *Cowles foundation discussion paper*, Yale University, 1295

Campbell, J.Y., Yogo, M., 2006. Efficient tests of stock return predictability, *Journal of financial economics*, 81, pp. 27-60

Caner, M., Hansen, B.U., 2001. Threshold autoregression with a unit root, *Econometrica*, November 2001, 69 (6), pp. 1555-1596

Caporale, G.M., Gil-alana, L.A., 2004. Fractional cointegration and tests of present value models, *Review of financial economics*, 13, pp. 245-258

Capozza, D.R., Hendershott, P.H., Mack, C., 2004. An anatomy of price dynamics in illiquid markets: analysis and evidence from local housing markets, *Real estate economics*, 32 (1), pp. 1-21

Case, K.E., Mayer C.J., 1996. Housing price dynamics within a metropolitan area, *Regional science and urban economics*, 26, pp. 387-407

Case, K.E., Shiller, R.J., 1989. The efficiency of the market for single-family homes, *The American economic review*, March 1989, 79 (1), pp. 25-137

Case, K.E., Shiller, R.J., 2004. Is there a bubble in the housing market?, *Brookings papers on economic activity*, 2, Cowles foundation for research in economics, Yale University, Cowles foundation paper N° 1089

Case, K.E., Quigley, J.M., Shiller, R.J., 2001. Comparing wealth effects: The stock market versus the housing market, *Cowles foundation discussion paper N°1335*, Cowles foundation for research in economics, Yale University

Chan, K.S., 1990. Testing for threshold autoregression, *The Annals of Statistics*, 18 (4), pp. 1886-1894

Chackraborty, S., 2009. The boom and the bust of the Japanese economy: A quantitative look at the period 1980-2000, *Japan and the world economy*, 21, pp. 116-131

Chan, K.S., 1990. Testing for threshold autoregression, *The annals of statistics*, December 1990, 18 (4), pp. 1886-1894

Chan, K.S., Tong, H., 1986. On tests for non-linearity in time series analysis, *Journal of forecasting*, 5 (4), pp. 217-228

Charles, A., Darné, O., 2005. Outliers and GARCH models in financial data, *Economics letters*, 86, pp. 347-352

Chatfield, C., 1977. Some recent developments in time-series analysis, *Journal of the royal statistical society*, Series A (General), 140 (4), pp. 492-510

Chatfield, C., 1997. Forecasting in the 1990s, *Journal of the royal statistical society. Series D* (*The statistician)*, 46 (4), pp. 461-473

Chen, L., 2009. On the reversal of return and dividend growth predictability: A tail of two periods, *Journal of financial economics*, 92, pp. 128-151

Chen, A.S., Leung, M.T., 2005. Modelling time series information into option prices: An empirical evaluation of statistical projection and GARCH option pricing model, *Journal of banking and finance*, Article in press

Cheung, Y.W., Lai, K.S., 1993. A fractional cointegration analysis of purchasing power parity, *Journal of business and economic statistics*, 11, pp. 103-112

Cho, V., 2002. A comparison of three different approaches to tourist arrival forecasting, *Tourism management*, 24 (3), pp. 323-330

Chong, Y.Y., Hendry, D.F., 1986. Econometric evaluation of linear macro-economic economic models, *The review of economic studies*, August 1986, 53 (4), pp. 671-690

Chortareas, G.E., Kapetanios, G., Chin, Y., 2002. Nonlinear mean reversion in real exchange rates, *Economics letters*, 77, pp. 411-417

Chou, R.Y., 2005. Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model, *Journal of money, credit and banking*, 37, pp. 561-582

Christoffersen, P.F., 1998. Evaluating interval forecasts. *International economic review*, vol.39, no.4, pp.841-862

Clark, T.E., McCracken, M.W., 2005. Evaluating direct multistep forecasts, *Econometric reviews*, October 2005, 24 (4), pp. 369-404

Clemen, R.T., Winkler, R.L., 1986. Combining economic forecasts, *Journal of business & economic statistics*, January 1986, 4 (1), pp. 39-46

Clements, M.P., Franses, P.H., Swanson, N.R., 2004. Forecasting economic and financial time-series with non-linear models, *International journal of forecasting*, 20, pp. 169-184

Clements, M.P., Harvey, D.I., 2007. Forecast encompassing tests and probability forecasts, April 2007, *Working paper*

Clements, M.P., Hendry, D.F., 1998. Forecasting economic processes, *International Journal of forecasting*, 14, pp. 111-131

Clements, M.P., Smith, J., 1997. The performance of alternative forecasting methods for SETAR models, *International journal of forecasting*, 13, pp. 463-475

Clements, M.P., Smith, J., 1999. A Monte Carlo study of the forecasting performance of empirical SETAR models, *Journal of applied econometrics*, March-April 1999, 14 (2), pp. 123-141

Clements, M.P., Smith, J., 2001. Evaluating forecasts from SETAR models of exchange rates, *Journal of international money and finance*, 20, pp. 133-148

Clements, M.P., Taylor, N., 2003. Evaluating interval forecasts of high-frequency financial data. *Journal of applied econometrics*, vol.18, pp.445-456

Cochrane, J.H., 2008. The dog that did not bark: a defence of return predictability, *Review of financial studies*, 21, pp. 1533-1575

Coleman IV, M., (Major Coleman IV), LaCour-Little, M., Vandell, K.D., 2008. Subprime lending and the housing bubble: Tail wags dog?, *Journal of housing economics*, 17, pp. 272-290

Crawford, G.W., Fratantoni, M.C., 2003. Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices, *Real estate economics*, 31 (2), pp. 223-243

Cuthbertson, K., Nitzsche, D., 2004. *Quantitative Financial Economics. Stocks, Bonds & Foreign Exchange*. Second Edition, John Wiley & Sons, Ltd., ISBN 10 0470 09171 1 (P/B)

Dacco, R., Satchell, S., 1999. Why do regime-switching models forecast so badly?, *Journal of forecasting*, 18, pp.1-16

Das, S., Gupta, R., Kabundi, A., 2009. Could we have predicted the recent downturn in the South African housing market?, *Journal of Housing Economics*, 18, pp. 325-335

Davies, N., Pemberton, J., Petruccelli, J.D., 1988. An automatic procedure for identification, estimation and forecasting univariate self exciting threshold autoregressive models, *The statistician*, Special issue: Statistical forecasting and decision-making, 37 (2), pp. 199-204

Davies, N., Petruccelli, J.D., 1986. Detecting non-linearity in time series, *The statistician*, Special issue: statistical modelling, 35 (2), pp. 271-280

De Gooijer, J.G., Hyndman, R.J., 2006. 25 Years of time series forecasting, *International journal of forecasting*, 22, pp. 443-473

De Gooijer, J.G., Kumar, K., 1992. Some recent developments in non-linear time series modelling, testing, and forecasting, *International journal of forecasting*, 8, pp. 135-156

De Gooijer, J.G., Molenaar, P., Schmitz, B., 1992. Dynamic factor analysis of nonstationary multivariate time series, *Psychometrika*, Springer, 57 (3), pp. 333-349

Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. *International economic review*, vol.39, no.4, pp.863-883

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy, *Journal of business & economic statistics*, July 1995, 13 (3), pp. 253-263

Diebold, F.X., Nason, J.A., 1990. Nonparametric exchange rate prediction?, *Journal of international economics*, May 1990, 28 (3-4), pp. 315-332

Doran, C.F., 1999. Why forecasts fail: The limits and potential of forecasting in international relations and economics, *International studies review*, Summer 1999, 1 (2), pp. 11-41

Duan, J.C., Simonato, J.G., 2001. American option pricing under GARCH by a Markov chain approximation, *Journal of economic dynamics and control*, 25, pp. 1689-1718

Edmonds, R.G. Jr., So, J.Y.C., 2004. Is exchange rate volatility excessive? An ARCH and AR approach, *The quarterly review of economics and finance*, 44, pp. 122-154

Edison, H.J., Klovland, J.T., 1987. A quantitative reassessment of the purchasing power parity hypothesis: evidence from Norway and the United Kingdom, *Journal of applied econometrics*, 2, pp. 309-333

Enders, W., Granger, C.W.J., 1998. Unit-root tests and asymmetric adjustment with an ample using of the term structure of interest rates, *Journal of business & economic statistics*, July 1998, 16 (3), pp. 304-311

Enders, W., Siklos, P.L., 2001. Cointegration and threshold adjustment, *Journal of business & economics statistics*, April 2001, 19 (2), pp. 166-176

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50 (4), pp. 987-1008

Engle, R.F., Granger, C.W.J., 1987. Core-integration and error corrections representation, estimation and testing, *Econometrica*, 55, pp. 251-276

Engle, R.F., Granger, C.W.J., Kraft, D., 1984. Combining competing forecasts of inflation using bivariate ARCH model, *Journal of economic dynamics and control*, 8, pp. 151-165

Englund, P., Hwang, M., Quigley, J.M., 2002. Hedging housing risk, *Journal of real estate and economics*, 24 (1-2), pp. 167-200

Ericsson, N.R., 1992. Parameter constancy, mean square forecast errors, and measuring forecast performance: an exposition, extensions, and illustration, *Journal of policy modelling*, 14 (4), pp. 465-495

Ericsson, N.R., Marquez, J., 1993. Encompassing the forecasts of U.S. trade balance models, *The review of economics and statistics*, February 1993, 75 (1), pp. 19-31

Evans, J.W., 1991. Pitfalls in testing for explosive bubbles in asset prices, *The American economic review*, September 1991, 81 (4), pp. 922-930

Fair, R.C., Shiller, R.J., 1990. Comparing information in forecasts from econometric models, *The American economic review*, June 1990, 80 (3), pp. 375-389

Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns, *Journal of financial economics*, 22, pp. 3-25

Fang, Y., Xu, D., 2003. The predictability of asset returns: an approach combining technical analysis and time series forecasts, *International journal of forecasting*, 19, pp. 369-385

Feige, E.L., Pearce, D.K., 1976. Economically rational expectations: Are innovations in the rate of inflation independent of innovations in measures of monetary and fiscal policy?, *Journal of political economy*, June 1976, 84, pp. 499-522

Figlewski, S., 1997. Forecasting volatility, *Financial markets, institutions & investments*, 6 (1), pp.1-88

Franses, P.H., van Dijk, D., 2000. Nonlinear time series models in empirical finance, Cambridge University Press, ISBN 0521779650

Fraser, P., Hoesli, M., McAlevey, L., 2008. House prices and bubbles in New Zealand, *Journal of real estate, finance and economics*, 37, pp. 71-91

Fraser, P., Hoesli, M., McAlevey, L., 2009. House prices, disposable income, and permanent and temporary shocks, *Working Paper*, Curtin University of Technology, November 2009, ISSN: 1835-9450

Froot, K.A., Obstfeld, M., 1991. Intrinsic bubbles: the case of stock prices, *The American economic review*, 81, pp. 1189-1214

Gallant, A.R., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics, *Journal of econometrics*, 81, pp. 159-192

Gao, A., Lin, Z., Fangzhou Na, C., 2009. Housing market dynamics: Evidence of mean reversion and downward rigidity, *Journal of housing economics*, 18, pp. 256-266

Garino, G., Sarno, L., 2004. Speculative bubbles in U.K. house prices: Some new evidence, *Southern economic journal*, 70 (4), pp. 777-795

Ghysels, E., Santa-Clara, P., Valkanov, R., 2003. Predicting volatility: Getting the most out of return data samples at different frequencies, *Working Paper*, National bureau of economic research, NBER working papers N° 10914

Gillen, K., Thibodeau, T., Wachter, S., 2001. Anisotropic autocorrelation in house prices, *Journal of real estate, finance and economics*, 23 (1), pp. 5-30

Goetzmann, W.N, Jorion, P., 1993. Testing the predictive power of dividend yields, *The Journal of finance*, June 1993, 48 (2), pp. 663-679

Goodman, A.C., Thibodeau, T.G., 2008. Where are the speculative bubbles in US housing markets?, *Journal of housing economics*, 17, pp. 117-137

González-Rivera, G., Lee, T.W., Mishra, S., 2004. Forecasting volatility: a reality check based on option pricing, utility function, value-at-risk, and predictive likelihood, *International journal of forecasting*, 20, pp. 629-645

Granato, J., Suzuki, M., 1996. Electoral inquiry. The use of the encompassing principle to restore empirical controversies in voting behaviour: An application to voter sophistication in congressional elections, *Electoral Studies*, 15 (3), pp. 383-398

Granger, C.W.J., 1980. Long memory relationships and aggregation of dynamic models, *Journal of econometrics*, October 1980, 14 (2), pp. 227-238

Granger, C.W.J., Newbold, P., 1977. *Forecasting Economic Time Series*. Academic Press, Inc., ISBN 0 12 295150 6

Granger, C.W.J., Newbold, P., 1986. *Forecasting Economic Time Series*, 2nd edition, Academic Press, ISBN 0-12-295184-0

Granger, C.W.J., Swanson, N.R., 1996. Future developments in the study of cointegrated variables, *Oxford bulletin of economics and statistics*, 58, pp. 537-553

Granger, C.W.J., Swanson, N.R., 1997. An introduction to stochastic unit-root processes, *Journal of econometrics*, 80, pp. 35-62

Granger, C.W.J., White, H., Kamstra, M., 1989. Interval forecasting. *journal of econometrics*, vol.40, pp.87-96

Gu, A.Y., 2002. The predictability of house prices, *Journal of real estate research*, 24 (3), pp. 213-233

Guerra, M.L., Sorini, L., 2005. Testing robustness in calibration of stochastic volatility models, *European journal of operational research*, 163, pp. 145-153

Guidolin, M., Hyde, S., McMillan, D., Ono, S., 2008. Non-linear predictability in stock and bond returns: When and where is it exploitable?, *Working paper*, Federal reserve bank of St. Louis.

Guidolin, M., Timmermann, A., 2003. Option prices under Bayesian learning: implied volatility dynamics and predictive densities, *Journal of economic dynamics & control*, 27, pp. 717-769

Guidolin, M., Timmermann, A., 2006. An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns, *Journal of applied econometrics*, 21, pp. 1-22

Guidolin, M., Timmermann, A., 2009. Forecasts of US short-term interest rates: A flexible forecast combination approach, *Journal of econometrics*, 150, pp. 297-311

Gujarati, D.N., 2003. *Basic Econometrics*, 4th edition, McGraw-Hill Companies Inc: Gary Burke, ISBN 0-07-112342-3

Gupta, S., Wilton, P.C., 1987. Combination of forecasts: an extension, *Management science*, March 1987, 33 (3), pp. 356-372

Hall, S.G., Mitchell, J., 2007. Combining density forecasts. *International journal of forecasting*, vol.23, pp.1-13

Hall, S., Psaradakis, Z., Sola, M., 1997. Switching error-correction models of house prices in the United Kingdom, *Econometric modelling*, 14, pp. 517-527

Harris, R., Sollis, R., 2003. *Applied time series modelling and forecasting*, Great Britain: John Wiley & Sons Ltd, ISBN 0-470-84443-4

Hartmann, D., Kempa, B., Pierdzioch, C., 2008. Economic and financial crisis and the predictability of U.S. stock returns, *Journal of empirical finance*, 15, pp. 468-480

Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors, *International journal of forecasting*, 13, pp. 281-291

Harvey, D., Leybourne, S., Newbold, P., 1998. Testing for forecast encompassing, *Journal of business and economic statistics*, April 1998, 16 (2), pp. 254-259

Harvey, D., Newbold, P., 2000. Tests for multiple forecast encompassing, *Journal of applied econometrics*, September-October 2000, 15 (5), pp. 471-482

He, H., Modest, D.M., 1995. Market frictions and consumption-based asset pricing, *The journal of political economy*, February 1995, 103 (1), pp. 94-117

Hendry, D.F., 1997. The econometrics of macro economic forecasting, *The economic journal*, September 1997, 107 (444), pp. 1330-1357

Hendry, D.F., 2000. *Econometrics: alchemy or science? Essays in econometric methodology*, Oxford: Oxford University Press

Hendry, D.F., Clements, M.P., 2003. Economic forecasting: some lessons from recent research, *Economic modelling*, 20, pp. 301-329

Henry, D.F., 2006. Robustifying forecasts from equilibrium-correction systems, *Journal of econometrics*, 135, pp. 399-426

Herwartz, H., Weber, H., 2005. Exchange rate uncertainty and trade growth – a comparison of linear and non-linear (forecasting) models, *Applied stochastic models in business and industry*, 21, pp. 1-26

Holden, K., Peel, D.A., Thompson, J.L., 1990. *Economic forecasting: An introduction*, Cambridge University Press, ISBN 0 521 35692 X

Holly, S., Jones, N., 1997. House prices since the 1940s: cointegration, demography and asymmetries, *Economic modelling*, 14, pp. 549-565

Hull, J.C., 2003. *Options, Futures, and Other Derivatives*, 5th edition, Prentice Hall, ISBN 0-13-046592-5

Himmelberg, C., Mayer, C., Sinai, T., 2005. Assessing high house prices: Bubbles, fundamentals and misperceptions, *Journal of economic perspectives*, Fall 2005, 19 (4), pp. 67-92

Jansen, E.S., Teräsvirta, T., 1996. Testing parameter constancy and super exogeneity in econometric equations, *Oxford bulletin of economics and statistics*, 58, pp. 735-763

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, November 1991, 59 (6), pp. 1551-1580

Kanas, A., 2005. Nonlinearity in the stock price-dividend relation, *Journal of international money and finance*, 24, pp. 583-606

Kanas, A., Yannopoulos, A., 2001. Comparing linear and nonlinear forecasts for stock returns, *International review of economics and finance*, 10, pp. 383-398

Kapetanios, G., Shin, Y., Snell, A., 2001.

Kapetanios, G., Shin, Y., Snell, A., 2003. Testing for a unit root in the nonlinear STAR framework, *Journal of econometrics*, 112, pp. 359-379

Kapetanios, G., Shin, Y., Snell, A., 2006. Testing for cointegration in nonlinear STAR error-correction models, *Econometric theory*, 22, pp. 279-303

Keenan, D.M., 1985. A Tukey nonadditivity-type test for time series nonlinearity, *Biometrika*, 72, pp. 39-44

Kilian, L., 1999. Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions?, *Journal of applied econometrics*, September–October 1999, 14 (5), pp. 491-510

Kilian, L., Taylor, M.P., 2003. Why is it so difficult to beat the random walk forecast of exchange rates?, *Journal of international economics*, 60, pp. 85-107

Kim, I.J., Kim, S., 2004. Empirical comparison of alternative stochastic volatility option pricing models: Evidence from Korean KOSPI 200 index options market, *Pacific-basin finance journal*, 12, pp. 117-142

Kim, S.W., Mollick, A.B., Nam, K., 2005. Another look at long-horizon stock returns: evidence from the G-7 stock markets, August 2005, *Working paper*, University of Texas pan American

Kirman, A.P., 1991. Epidemics of option and speculative bubbles in financial markets in M.P. Taylor (ed), *Money and financial markets*, Cambridge: Basil Blackwell, pp 354-368

Kirman, A.P., 1993. Ants, rationality, and recruitment, *Quarterly journal of economics*, 108, pp. 137-156

Koetter, M., Poghosyan, T., 2009. Real estate prices and bank stability, *Journal of Banking & Finance*, Article in press, doi:10.1016/j.jbankfin, 2009.11.010

Kräger, H., Kugler, P., 1993. Non-linearities in foreign exchange markets: a different perspective, *Journal of international money and finance*, 12, pp. 195-208

Krasker, W.S., 1990. Outliers, *Time series and statistics*, edited by Eatwell, J., Milgate, M., Newman, P., 1990. W. W. Norton & Company Inc, ISBN 0-393-95862-0

Lanne, M., 2002. Testing the predictability of stock returns, *The review of economics and statistics*, August 2002, 84 (3), pp. 407-415

Leitch, G., Tanner, J.E., 1991. Economic forecast evaluation: Profits versus the conventional error measures, *The American economic review*, June 1991, 81 (3), pp. 580-590

Lekkos, I., Milas, C., 2004. Time-varying excess returns on UK government bonds: A non-linear approach, *Journal of banking and finance*, 28, pp. 45-62

Lewellen, J., 2004. Predicting returns with financial ratios, *Journal of financial economics*, 74, pp. 209-235

Lim, C., McAller, M., 2001. Time series forecasts of international travel demand for Australia, *Tourism management*, 23 (4), pp. 389-396

Liu, J., Li, W.K., Li, C.W., 1997. On a threshold autoregression with conditional heteroscedastic variables, *Journal of statistical planning and inference*, 62, pp. 279-300

Lo, M., Wong, M.C.S., 2006. What explains the deviations of purchasing parity across countries? International evidence from macro data, *Economics letters*, 91, pp. 229-235

Locke, P.R., Sayers, C.L., 1993. Intra-day futures price volatility: Information effects and variance persistence, *Journal of applied econometrics*, 8 (1), pp. 15-30

Luukkonen, R., Saikkonen, P., Teräsvirta, T., 1988. Testing linearity against smooth transition autoregressive models, *Biometrika*, 75 (3), pp. 491-499

MacDonald, R., 1996. Panel unit root tests and real exchange rates, *Economics letters*, 50, pp. 7-11

Makridakis, S., Hibon, M., Moser, C., 1979. Accuracy of forecasting: an empirical investigation, *Journal of the royal statistical society*, Series A (general), 142 (2), pp. 97-145

Marmol, F., Velasco, C., 2004. Consistent testing of cointegrating relationships, *Econometrica*, November 2004, 72 (6), pp. 1809-1844

Martens, m., Kofman, P., Vorst, T.C.F., 1998 A threshold error-correction model for intraday futures and index returns, *Journal of applied econometrics*, May-June 1998, 13 (3), pp. 245-263

McLeod, A.I., Li, W.K., 1983. Diagnostic checking ARMA time series models using squared-residual autocorrelations, *Journal of time series analysis*, 4 (4), pp. 269-273

McMillan, D.G., 2001. Nonlinear predictability of stock market returns: Evidence from nonparametric and threshold models, *International review of economics and finance*, 10, pp. 353-368

McMillan, D.G., 2002. Non-linear predictability of UK stock market returns, *Working paper*, Department of accountancy and finance, University of Aberdeen

McMillan, D.G., 2003. Non-linear predictability of UK stock market returns, *Oxford bulletin of economics and statistics*, 65 (5), pp. 0305-9049

McMillan, D.G., 2004. Non-linear error correction: Evidence for UK interest rates, *The Manchester school*, 72 (5), pp. 626-640

McMillan, D.G., 2005a. A re-examination of the price-dividend ratio: evidence of non-linear behaviour with a break, *Working paper*, University of Durham

McMillan, D.G., 2005b. Non-linear dynamics in international stock market returns, *Review of financial economics*, 14, pp. 81-91

McMillan, D.G., 2006. Non-linear cointegration and adjustment: an asymmetric exponential smooth-transition model for U.S. interest rates, *Working paper*, February 2006, University of St. Andrews

McMillan, D.G., 2007. Bubbles in the dividend-price ratio? Evidence from an asymmetric exponential smooth-transition model, *Journal of banking & finance*, 31, pp. 787-804

McMillan, D.G., 2009. Are share prices still too high?, *Research in international business and finance*, 23, pp. 223-232

McMillan, D.G., 2009. Revisiting dividend yield dynamics and returns predictability: evidence from a time-varying ESTR model, *The quarterly review of economics and finance*, 49, pp. 870-883

McMillan, D.G., Speight, A.E.H., 2006. Nonlinear dynamics and competing behavioural interpretations: evidence from intra-day FTSE-100 index and futures data, *Journal of futures markets*, 26 (4), pp. 343-368

McNees, S.K., 1995. Forecast uncertainty: Can it be measured?, *Federal Reserve Bank of New York discussion paper*

McQuinn, K., O'Reilly, G., 2008. Assessing the role of income and interest rates in determining house prices, *Economic modelling*, 25, pp. 377-390

Meese, R., Wallace, N., 1994. Testing the present value relation for housing prices: Should I leave my house in San Francisco?, *Journal of urban economics*, 35, pp. 245-266

Michael, P., Nobay, A.R., Peel, D.A., 1997. Transactions costs and nonlinear adjustment in real exchange rates: An empirical investigation, *Journal of political economy*, 105, pp. 862-879

Mikhed, V., Zemčík, P., 2009. Do house prices reflect fundamentals? Aggregate and panel data evidence, *Journal of housing economics*, 18, pp. 140-149

Miles, W., 2008. Boom-bust cycles and the forecasting performance of linear and non-linear models of house prices, *Journal of real estate finance and economics*, 36, pp. 249-264

Mills, T.C., 1990. *Time Series techniques for Economists*, Cambridge University Press, ISBN 0 521 40574 2

Mitchell, J., Hall, S.G., 2005. Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR 'fan' charts of inflation. *Oxford bulletin of economics and statistics*, vol.67, supplement, pp.0305-9049

Montgomery, A.L., Zarnowitz, V., Tsay, R.S., Tiao, G.C., 1998. Forecasting the U.S. unemployment rate, *Journal of the American statistical association*, June 1998, 93 (442), pp. 478-493

Muellbauer, J., Murphy, A., 1997. Booms and bursts in the UK housing market, *The economic journal*, November 1997, 107 (445), pp. 1701-1727

McQuinn, K., OReilly, G., 2008. Assessing the role of income and interest rates in determining house prices, *Economic modelling*, 25, pp. 377-390

Nasseh, A., Strauss, J., 2004. Stock prices and the dividend discount model: did their reaction break down in the 1990s?, *The quarterly review of economics and finance*, 44, pp. 191-207

Newbold, C., Granger, C.W.J., 1974. Experience with forecasting univariate time series and the combination of forecasts, *Journal of the royal statistical society: series A*, 137, pp. 131-165

Ng, H.G., McAleer, M., 2001. The effects of outliers and model misspecification on recursive modelling of volatility, *Working paper*

Ng, H.G., McAleer, M., 2002. Recursive modelling of symmetric and asymmetric volatility in the presence of extreme observations, *Working paper*

Ortalo-Magné, F., Rady, S., 2006. Housing market dynamics: On the contribution of income shocks and credit constraints, *The review of economic studies*, April 2006, 73 (2), pp. 459-485

Pain, N., Westaway, P., 1997. Modelling structural change in the UK housing market: a comparison of alternative house price models, *Economic modelling*, 14, pp. 587-610

Pandey, A., 2005. Volatility models and their performance in Indian capital markets, *Vikalpa: The journal for decision makers*, 30 (2), pp. 27-46

Pesaran, M.H., Timmermann, A., 1995. Predictability of stock returns: Robustness and economic significance, *The journal of finance*, September 1995, 50 (4), pp. 1201-1228

Pesaran, M.H., Timmermann, A., 2000. A recursive modelling approach to predicting UK stock returns, *The economic journal*, January 2000, 110 (460), pp. 159-191

Petruccelli, J.D., Davies, N., 1986. A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series, *Biometrica*, 73 (3), pp. 687-694

Pong, S., Shackleton, M.B., Taylor, S.J., Xu, X., 2004. Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models, *Journal of banking and finance*, 28, pp. 2541-2563

Poon, S.H., Granger, C.W.J., 2003. Forecasting volatility in financial markets: A review, *Journal of economic literature*, 41, pp. 478-539

Poon, S.H., Granger, C., 2005. Practical issues in forecasting volatility, *Financial analysts journal*, 61 (1), pp. 45-56

Pascalau, R., 2007. Testing for a unit route in the asymmetric nonlinear smooth transition framework, August 2007, *Working paper*, University of Alabama

Peel, D., Davidson, J., 1998. A non-linear error correction mechanism based on the bilinear model, *Economics letters*, 58, pp. 165-170

Phelps, E.S., 1994. Structural slums: the modern equilibrium theory of unemployment, interest and assets, *Cambridge, MA: Harvard University press*

Phelps, E.S., Zoega, G., 2001. Structural booms, *Economic policy*, 32, pp. 83-126

Pindyck, R.S., Rubinfeld, D.L., 1998, *Econometric models and economic forecasts*, 4th edition, Singapore: Irwin McGraw-Hill, ISBN 0-07-118831-2

Psaradakis, Z., Sola, M., Spagnolo, F., 2004. On Markov error-correction models with an application to stock prices and dividends, *Journal of applied econometrics*, January-February 2004, 19 (1), pp. 69-88

Rapach, D.E., Wohar, M.E., Rangvid, J., 2005. Macro variables and international stock return predictability, *International journal of forecasting*, 21, pp. 137-166

Ready, M.J., 2002. Profits from technical trading rules, *Financial management*, Autumn 2002, 31 (3), pp. 43-61

Reilly, F.K., Brown, K.C., 2003. *Investment analysis and portfolio management*, 7th edition, United States of America: Thomson south-western, ISBN 0-324-17173-0

Rapach, D.E., Wohar, M.E., Rangvid, J., 2005. Macro variables and international stock return predictability, *International journal of forecasting*, 21, pp.137-166

Roche, M.J., 2001. The rise in house prices in Dublin: bubble, fad or just fundamentals, *Economic modelling*, 18, pp. 281-295.

Rosen, H.S., Rosen, K.T., Holtz-Eakin, D., 1984. Housing tenure, uncertainty, and taxation, *The review of economics and statistics*, August 1984, 66 (3), pp. 405-416

Sadorsky, P., 2005 Stochastic volatility forecasting and risk management, *Applied financial economics*, 15, pp. 121-135

Sandman, G., Koopman, S.J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood, *Journal of econometrics*, 87, pp. 271-301

Sarantis, N., 1999. Modelling nonlinearities in effective exchange rates, *Journal of international money and finance*, 18, pp. 27-45

Sarantis, N., 2001. Nonlinearities, cyclical behaviour and predictability in stock markets: international evidence, *International journal of forecasting*, 17, pp. 459-482

Sargan, J.D., 1964. Wages and prices in the United Kingdom: A study in econometric methodology, *Econometric analysis for national economic planning ed by P.E. Hart, G. Mills, J.N. Whitaker*, London: Butterworth

Schmidt, P., Phillips, P.C.B., 1992. LM Tests for a unit root in the presence of deterministic trends, *Oxford bulletin of economics and statistics*, 54, pp. 257-287

Schwarz, G., 1978. Estimating the dimension of a model, *Anals of statistics*, 6, pp. 461-464

Shiller, R.J., 1990. A scott-type regression test of the dividend ratio model, *The review of economics and statistics*, 72 (2), pp. 356-371

Shively, P.A., 2007. Asymmetric temporary and permanent stock-price innovations, *Journal of empirical finance*, 14, pp. 120-130

Shleifer, A., 2000. Inefficient markets. An introduction to behavioural finance, *Clarendon lectures in economics,* Oxford: Oxford University press

Siklos, P.L., 2002. Asymmetric adjustment from structural booms and slumps, *Economics letters*, 77, pp. 329-333

Siklos, P.L, Granger, C.W.J., 1997. Regime sensitive cointegration with an application to interest rate parity, *Macroeconomic dynamics,* 3, pp. 640-657

Sollis, R., 2009. A simple unit root test against asymmetric STAR nonlinearity with an application to real exchange rates in Nordic countries, *Economic modelling*, 26, pp. 118-125

Sollis, R., Leybourne, S.J., Newbold, P., 2002. Tests for symmetric and asymmetric non-linear mean reversion in real exchange rates, *Journal of money, credit and banking*, 34, pp. 686-700

Stambaugh, R.F., 1999. Predictive regressions, *Journal of financial economics*, 54, pp. 375-421

Stiglitz, J.E., 1990. Symposium on bubbles, *The journal of economic perspectives*, Spring 1990, 4 (2), pp. 13-18

Taylor, J.W., 2004. Volatility forecasting with smooth transition exponential smoothing, *International journal of forecasting*, 20, pp. 273-286

Teräsvirta, T., Anderson, H.M., 1992. Characterizing nonlinearities in business cycles using smooth transition autoregressive models, *Journal of applied econometrics*, 7, pp. S119-S136

Teräsvirta, T., van Dijk, D., Medeiros, M.C., 2005. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination, *International journal of forecasting*, 21, pp. 755-774

Timmermann, A.G., 1993. How learning in financial markets generates excess volatility and predictability in stock prices, *The quarterly journal of economics*, 108 (4), pp. 1135-1145

Timmermann, A.G., 1994. Can agents learn to rational expectations? Some results on convergence and stability of learning in the UK stock market, *The economic journal*, 104 (425), pp. 777-797

Timmermann, A.G., 1995. Cointegration tests of present value models with a time-varying discount factor, *Journal of applied econometrics*, 10 (1), pp. 17-31

Timmermann, A.G., 1996. Excess volatility and predictability of stock prices in autoregressive dividend models with learning, *The review of economic studies*, 63 (4), pp. 523-557

Tong, H., 1978. On a threshold model, *In C.H. Chen (ed): Pattern recognition and signal processing*, Cijhoff and Noordoff, Amsterdam

Tong, H., 1983. Threshold model is non-linear time series analysis, *New York: Springer-Verlag*

Tong, H., 1990. Non-linear time series: A dynamic system approach, *Oxford: Clarendon Press*

Tong, H., 1990. *Non-linear Time Series*. Clarendon Press Oxford, ISBN 0 190852224 X

Tong, H., Lim, K.S., 1980. Threshold autoregression, limit cycles and cyclical data, *Journal of the royal statistical society: series B*, 42, pp. 245-292

Tourous, W., Valkanov, R., Yan, S., 2004. On predicting stock returns with nearly integrated explanatory variables, *Journal of business*, 77 (4), pp. 937-966

Tsay, R.S., 1989. Testing and modelling threshold autoregressive processes, *Journal of the American statistical association*, 84 (405), pp. 231-240

Turner, T.M., 2003. Does investment risk affect the housing decisions of families?, *Economic enquiry*, 41 (4), pp. 675-691

Valkanov, R., 2003. Long-horizon regressions: theoretical results and applications, *Journal of financial economics*, 68, pp. 201-232

Van Dijk, D., Franses, P.H., 2003. Selecting a nonlinear time series model using weighted tests of equal forecast accuracy, *Working paper*, Econometric Institute Report, EI 2003-10

Van Norden, S., Vigfusson, R., 1998. Avoiding the pitfalls: can regime-switching tests reliably detect bubbles?, *Studies in non-linear dynamics and econometrics*, 3 (1), pp. 1-22

Wagle, B.V., 1965. A review of two statistical aids in forecasting, *The statistician*, 15 (2), pp. 191-196.

Wallis, K.F., 2003. Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts. *International journal of forecasting*, vol.19, pp.165-175

Wallis, K.F., 2005. Combining density and interval forecasts: a modest proposal. *Oxford bulletin of economics and statistics*, vol.67, supplement, pp.0305-9049

West, K.D., 1988. Bubbles, fads and stock price volatility tests: A partial evaluation, *The journal of finance*, 43 (3), pp. 639-656

Wheaton, W.C., Nechayev, G., 2008. The 1998-2005 housing bubble and the current correction: What's different this time?, *Journal of real estate research*, 30 (1), pp. 1-26

White, E.N., 1990. The stock market boom and crash of 1929 revisited, *The journal of economic perspectives*, Spring 1990, 4 (2), pp. 67-83

Winkler, R.L., Clemen, R.T., 1992. Sensitivity of weights in combining forecasts, *Operations research*, May-June 1992, 40 (3), pp. 609-614

Wolf, M., 2000. Stock returns and dividend yields revisited: a new way to look at an old problem, *Journal of business & economic statistics*, January 2000, 18 (1), pp. 18-30

Zakoian, J.M., 1994. Threshold heteroscedastic models, *Journal of economic dynamics and control,* 18 (5), pp. 931-955

Zandi, M., Chen, C., 2006. Is the price right?, *Regional financial review*, March 2006, pp. 11-25