A re-examination of the relationship between FTSE100 index and futures prices

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: https://dspace.lboro.ac.uk/2134/8071

Publisher: © Juan Tao

Please cite the published version.
This item is held in Loughborough University’s Institutional Repository (https://dspace.lboro.ac.uk/) and was harvested from the British Library’s EThOS service (http://www.ethos.bl.uk/). It is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
A Re-examination of the Relationship between FTSE100 Index and Futures Prices

by

Juan Tao

A Doctoral Thesis
Submitted in Partial Fulfilment of the Requirements for the Award of Doctor of Philosophy

Department of Economics
Loughborough University

© Juan Tao, 2008
Acknowledgements

I am greatly indebted to my supervisor, Professor Christopher Green, for his wise guidance, useful advice and constant encouragement and support. Without his help, this work would not be possible.

I would like to express my sincere and deep gratitude to Dr Lawrence Leger for all the valuable discussions and generous and sociable help.

I would also like to thank Professor Terence Mills for his great patience with my questions and smart suggestions on my research method and data problem.

I also wish to express my gratitude to the Department of Economics, Loughborough University, for their generous financial support.

My deepest gratitude goes to my family. I am grateful to my parents for their generous love, considerate thoughts and continuous encouragement. I am also deeply grateful to my parents-in-law for their generous support and care of my son. I owe my special loving thanks to my dear husband, Zhihua Lu and my dear son, Xingjian Lu. They have lost a lot due to my research abroad in the UK. Without their understanding and support, it would have been simply impossible to finish this work.
Abstract

This thesis examines the validity of the cost of carry model for pricing FTSE100 futures contracts and the relationship between FTSE100 spot and futures markets during two sub-periods characterised by different market trading systems employed by the LSE and LIFFE. The empirical work is carried out using three approaches to econometric modeling: a basic VECM for spot and futures prices, a VECM extended with a DCC-TGARCH framework to account for the conditional variance-covariance structure for spot and futures prices and a threshold VECM to capture regime-dependent spot-futures price dynamics.

Overall, both the basic VECM and the DCC-TGARCH analysis suggest that there are deviations from the cost of carry relationship in the first sub-sample when transactions costs in both markets are relatively high but that the cost of carry relationship tends to be valid in the second sub-sample when transactions costs are lower. This is further confirmed by the evidence of higher conditional correlations between the two markets in the second sub-sample as compared with the first, using the DCC-TGARCH analysis. This implies that the no-arbitrage cost of carry relationship between spot and futures markets is more effectively maintained by index arbitrageurs in the second period when market conditions are closer to perfect market assumptions, and hence the cost of carry model could be more reasonably used as a benchmark for pricing stock index futures.

The threshold VECM analysis depicts regime-dependent price dynamics between FTSE100 spot and futures markets and leads to some interesting and important findings: arbitrage may not be practicable under some market conditions, either because it is difficult to find counterparties for the arbitrage transactions, or because there is significant risk associated with arbitrage; as a result, the cost of carry model may not always be suitable for pricing stock index futures. Furthermore, the threshold values yielded from estimating the threshold VECM reflect the average transaction costs for most arbitrageurs that are more reliable and fair than subjective estimations.

Keywords: cost of carry; trend-corrected basis; artificial price jumps; vector error correction model (VECM); DCC-TGARCH; CCF test; non-linear cointegration; mispricing; threshold VECM
Contents

Chapter 1 Introduction ...................................................................................................... 1
  1.1 Background and Motivations .................................................................................. 1
  1.2 Thesis Overview and Aims ..................................................................................... 3
    1.2.1 Sample selection .............................................................................................. 4
    1.2.2 Methodological choices ................................................................................... 6
  1.3 Contributions to Knowledge ................................................................................... 7
  1.4 Structure of the Thesis .......................................................................................... 10

Chapter 2 An Introduction to Futures and Stock Index Futures ..................................... 12
  2.1 Futures markets and stock index futures .............................................................. 12
    2.1.1 The mechanics of futures markets ................................................................. 12
    2.1.2 Some advantages of trading in futures markets ............................................. 13
  2.2 The economic functions of futures and stock index futures .................................. 15
    2.2.1 Risk transfer ................................................................................................... 15
    2.2.2 Price discovery .............................................................................................. 16
    2.2.3 Other functions .............................................................................................. 17
  2.3 The price relationship between index and its futures price .................................. 17
  2.4 Index Arbitrage in the UK .................................................................................... 21
    2.4.1 Hurdles to arbitrage ...................................................................................... 21
    2.4.2 The effect of index-tracking funds on index arbitrage .................................. 24

Chapter 3 Futures Mispricing, the Implications for Arbitrage, and Causal Relationships
between Stock Index and Futures Prices Literature Review .......................................... 26
  3.1 Introduction ........................................................................................................... 26
  3.2 Futures contract mispricing and the implications for arbitrage ............................ 27
    3.2.1 Empirical investigations of futures mispricing .............................................. 28
    3.2.2 Interpretation of mispricing: does it always represent arbitrage opportunity? ........................................................................................................... 29
    Marking to market and stochastic interest rates ...................................................... 29
    Seasonal dividends ................................................................................................. 30
    The influence of tax and the tax timing option ..................................................... 31
    Transaction costs and arbitrage risks .................................................................... 33
    Short sale restrictions ......................................................................................... 35
    The impact of index composition ........................................................................ 36
Alternatives to the cost of carry model for pricing stock index futures

3.3 Causal relationships

3.3.1 Why do causal relationships exist between spot and futures prices?

Non-synchronous trading
Transaction costs
Leverage effect
Different information arrival

‘Economy-wide’ versus ‘firm-specific’ information
‘Good’ or ‘bad’ news

3.3.2 Cointegration and causality-in-mean

Linear cointegration
Threshold cointegration

3.3.3 Conditional heteroskedasticity

The relation between price volatility and futures contract maturity
The relation between price volatility and trading volume

3.3.4 Causality-in-variance

3.4 Summary

3.5 Conclusions

Chapter 4 Unified Tests of Causality and Cost of Carry – A VECM Analysis of FTSE100 Spot and Futures Prices

4.1 Introduction

4.2 Data and Preliminary Analysis

4.2.1 Data and sample descriptions
4.2.2 Normality test
4.2.3 Unit root test
4.2.4 Autocorrelation Test

4.3 Methodology

4.3.1 Cointegration Test
4.3.2 An Error Correction Model
4.3.3 Joint Test of Causality and Cost of Carry

4.4 Empirical results

4.4.1 Causality and cointegration
4.4.2 Wald test of the cost of carry model
4.4.3 Causality and cointegration under accepted cost of carry restrictions
List of Figures

Figure 4.1 Time plot of FTSE100 index price, 28/10/1986 - 30/12/2005 ..........................64
Figure 4.2 Time plot of FTSE100 futures price, 28/10/1986 - 30/12/2005 ......................65
Figure 4.3 Time plot of the cost of carry for FTSE100 index and futures prices, 28/10/1986 - 17/10/1997 .........................................................................................66
Figure 4.4 Time plot of the cost of carry for FTSE100 index and futures prices, 27/10/1997 - 30/12/2005 .........................................................................................67
Figure 4.5 Time plot of the basis for FTSE100 futures and index prices, 28/10/1986 - 17/10/1997 .........................................................................................68
Figure 4.6 Time plot of the basis for FTSE100 futures and index prices, 27/10/1997 - 30/12/2005 .........................................................................................69
Figure 4.7 Serial and cross-correlations of the residuals estimated from the VAR of equations (4.8a) and (4.8b) (4 lags), 28/10/1986 - 30/12/2005 .................................92
Figure 4.8 Serial and cross-correlation in the residuals estimated from the VAR of equations (4.8a) and (4.8b) (4 lags), 28/10/1986 - 17/10/1997 .....................................................96
Figure 4.9 Serial and cross-correlations of the residuals estimated from the VAR of equations (4.8a) and (4.8b) (5 lags), 27/10/1997 - 30/12/2005 .................................98
Figure 4.10 The plots of CUSUM against 5% critical bounds for residuals of equation (4.8a) and equation (4.8b), 28/10/1986 - 17/10/1997 .........................................................99
Figure 4.11 The plots of CUSUM against 5% critical bounds for residuals of equation (4.8a) and equation (4.8b), 27/10/1997 - 30/12/2005 .........................................................100
Figure 4.12 The plot of residuals estimated from equation (4.8a) for spot returns, 28/10/1986 - 17/10/1997 .........................................................................................102
Figure 4.13 The plot of residuals estimated from equation (4.8b) for futures returns, 28/10/1986 - 17/10/1997 .........................................................................................103
Figure 4.14 The plot of residuals estimated from equation (4.8a) for spot returns, 27/10/1997 - 30/12/2005 .........................................................................................104
Figure 4.15 The plot of residuals estimated from equation (4.8b) for futures returns, 27/10/1997 - 30/12/2005 .........................................................................................105
Figure 4.16 The plot of residuals estimated from equation (4.11a) for spot returns, 28/10/1986 - 17/10/1997 .........................................................................................107
Figure 4.17 The plot of residuals estimated from equation (4.11b) for futures returns, 28/10/1986 - 17/10/1997 .........................................................................................108
Figure 5.1 Time plot of daily FTSE 100 spot returns (in logs), 28/10/1986 – 17/10/1997 ................................................................. 137
Figure 5.2 Time plot of daily FTSE 100 spot returns (in logs), 28/10/1986 – 17/10/1997 (19/10/1989 – 23/10/1987 not included) ................................................................. 138
Figure 5.3 Time plot of daily FTSE 100 spot returns (in logs), 27/10/1997 – 30/12/2005 ................................................................. 139
Figure 5.4 Time plot of daily FTSE 100 futures returns (in logs), 28/10/1986 – 17/10/1997 ................................................................. 140
Figure 5.5 Time plot of daily FTSE 100 futures returns (in logs), 28/10/1986 – 17/10/1997 (19/10/1989 – 23/10/1987 not included) ................................................................. 141
Figure 5.6 Time plot of daily FTSE 100 futures returns (in logs), 27/10/1997 – 30/12/2005 ................................................................. 142
Figure 5.7 FTSE100 spot conditional variance versus spot price, 03/11/1986 – 17/10/1997 ................................................................. 164
Figure 5.8 FTSE100 futures conditional variance versus future price, 03/11/1986 – 17/10/1997 ................................................................. 165
Figure 5.9 FTSE100 spot conditional variance versus spot price, 03/11/1997 – 30/12/2005 ................................................................. 166
Figure 5.10 FTSE100 futures conditional variance versus future price, 03/11/1997 – 30/12/2005 ................................................................. 167
Figure 5.11 Futures conditional variance versus spot conditional variance, 03/11/1986 – 17/10/1997 ................................................................. 168
Figure 5.12 Futures conditional variance versus spot conditional variance, 03/11/1997 – 30/12/2005 ................................................................. 169
Figure 5.13 Correlation between spot and futures markets, 03/11/1986 – 17/10/1997170
Figure 5.14 Correlation between spot and futures markets, 03/11/1997 – 30/12/2005171
Figure 6.1 The no-arbitrage window determined by transaction costs................. 177
List of Tables

Table 3.1 The relationship between futures and cash index prices with taxes .................32
Table 4.1 Summary statistics for FTSE100 spot and futures returns ..............................71
Table 4.2 Unit Root Tests, 28/10/1986 – 30/12/2005 ..................................................73
Table 4.3 ACF and PACF of FTSE100 spot and futures returns ....................................75
Table 4.4 Cointegration Tests .......................................................................................78
Table 4.5 Testable restrictions imposed by the continuous stochastic cost of carry
relationship (equation 4.4) on selected VECM equations (4.8a) and (4.8b) ...............88
Table 4.6 VAR lag order selected by different information criteria ..............................91
Table 4.7 Chow tests for stability of estimated coefficients of equations (4.8a) and (4.8b)
(4 lags), 28/10/1986 – 30/12/2005 ...........................................................................94
Table 4.8 VAR lag order selected by different information criteria ..............................95
Table 4.9 VAR lag order selected by different information criteria ..............................97
Table 4.10 SUR estimation of equations (4.31a) and (4.31b) ......................................110
Table 4.11 SUR estimation of equations (4.8a) and (4.8b) ........................................113
Table 4.12 Wald test of the restrictions imposed by the cost of carry model on the
selected VECM .............................................................................................................116
Table 4.13 SUR estimation of the restricted equations (4.31a) and (4.31b) .................122
Table 4.14 SUR estimation of the restricted equations (4.8a) and (4.8b) .................123
Table 5.1 ACF and PACF of the returns and squared returns ......................................144
Table 5.2 ARCH-LM test on OLS residuals of equations (5.1a) and (5.1b) ................145
Table 5.3 Ljung-Box Q test on standardized residuals (e1) and their squared values (e2)
estimated from the fitted DCC-TGARCH (1,1) model ................................................153
Table 5.4 Results of DCC-TGARCH(1, 1), 28/10/1986 – 17/10/1997 .........................157
Table 5.5 Results of DCC-TGARCH(1, 1), 27/10/1997 – 30/12/2005 .........................161
Table 5.6 CCF test for causality-in-variance .................................................................173
Table 6.1 Tsay test for threshold nonlinearity in the basis ...........................................184
Table 6.2 Ljung-Box test on residuals estimated from the selected threshold VECM for
spot and futures returns .................................................................................................190
Table 6.3 Least squares estimates and their t-ratios of the threshold VECM, 28/10/1986
- 17/10/1997 ..............................................................................................................193
Table 6.4 Least squares estimates and their t-ratios of the threshold VECM, 27/10/1997
- 30/12/2005 ..............................................................................................................197
Chapter 1 Introduction

1.1 Background and Motivations

"By far the most significant event in finance during the past decade has been the extraordinary development and expansion of financial derivatives. ... These instruments enhance the ability to differentiate risk and allocate it to those investors most able and willing to take it. This unbundling improves the ability of the market to engender a set of product and asset prices far more calibrated to the value preferences of consumers than was possible before derivatives markets were developed. The product and asset price signals enable entrepreneurs to finely allocate real capital facilities to produce those goods and services most valued by consumers, a process that has undoubtedly improved national productivity growth and standards of living."

— Alan Greenspan, 19 March 1999

Every day, millions of individuals and corporations make decisions about where to invest their savings and how to manage the potential risks associated with owning (or potentially owning) an asset in volatile and unpredictable markets. In response to the increasing demand for investment and risk management, financial futures have evolved quickly from the original futures trading on agricultural products. The late seventies and early eighties first saw the introduction of financial futures contracts on U.S. Treasury bills and bonds, stock indices, and Eurodollar time deposits in response to radical changes in the international currency system (the collapse of the Bretton Woods system in 1971) and in the way the Federal Reserve managed the nation's money supply. Active financial futures markets were also quickly introduced in countries outside the U.S., notably England and Japan. Nowadays, futures trading plays a noticeable role in the global financial system, accounting for trillions of dollars in trades every day. The ‘price discovery’ function of futures markets helps people to make more efficient investment decisions, and hence capital is allowed to flow to its most highly valued use. The ‘risk transfer’ function of futures markets allows the transfer of risk from risk-averse investors to those who are most able and willing to take it, and hence helps to re-distribute risk efficiently across market participants. In this way, futures markets contribute substantially to the effective functioning and

---

1 Alan Greenspan was the chairman of the Federal Reserve of the United States from 1987 to 2006. These remarks are from his speech on financial derivatives given in Boca Raton, Florida on 19 March 1999.
sustainable development of capital markets, which is important to individual economic units, the financial system and the economy as a whole.

Stock index futures are one of the most important financial futures contracts. Historically, stock index futures have supplemented, and often replaced, the secondary stock market as a stock price discovery mechanism. They allow significant improvements in market timing, and an investment strategy based on the outlook for an aggregate market rather than for a particular financial asset. Furthermore and most importantly, they provide investors with an efficient and cost-effective means of hedging against market risks. Institutional investors such as mutual funds and pension funds are the main beneficiaries in this regard. Stock index futures are therefore an important tool that can be used to optimize the functioning of capital markets and improve the risk-avoidance capability and competitive strength for each economic unit.

Stock index futures have a history of more than 20 years in western developed countries. Since the Kansas City Board of Trade introduced the first stock index futures on the Value Line Index on 24 February 1982, followed immediately by the Chicago Mercantile Exchange’s introduction of futures contracts on the S&P500 index just about two months later, there has been a host of stock index futures introduced by other developed countries. The FTSE100 futures were introduced in the U.K. on 3 January 1984. More recently, stock index futures have been gradually introduced by some developing countries. For example, stock index futures were launched by India on 12 June 2000, by Mexico on 2 January 2003, and by Turkey on 4 February 2005. China is currently also in the preparation stage of launching stock index futures, expected to materialize in 2008, based on the Shanghai and Shenzhen 300 Index.

While the U.S. financial futures market has been thoroughly investigated, studies of the U.K. financial futures market are comparatively few, despite its importance in the financial world. The development of the U.K. financial futures market in the past two decades has not been smooth and has experienced a range of different shocks. Some representative events include: the deregulation of the financial market on 27 October 1986; the stock market crash in October 1987; the dot-com bubble of the late 1990s and its subsequent bursting in the early 2000s; the microstructure transformations of both the London Stock Exchange (LSE) and London International Financial Futures Exchange (LIFFE). The performance of the U.K. financial futures market during this
period is therefore worth investigation. Furthermore, the development of financial futures markets in western developed countries may have important practical as well as regulatory lessons for developing countries such as China that are still in the early stage of promoting financial futures. All these considerations have formed the initial motivation for this research.

1.2 Thesis Overview and Aims

Given that stock index futures are often believed to provide 'price discovery' services to the underlying market, it is important to understand the mechanism for pricing stock index futures and to determine the efficiency with which these services are performed. Index arbitrage is believed to be necessary for an efficient and thriving index futures market. The cost of carry model based on a simple no-arbitrage argument between futures and the underlying index is the most widely used model for pricing stock index futures. Key issues concerning stock index futures include the following. Is the index futures market performance efficient? Do index futures qualify as an efficient stock price discovery tool for the underlying market? Is the widely cited cost of carry model valid for pricing stock index futures? How can stock index futures be used for effective hedging against market risks? Being guided by these questions, this research has an empirical examination of daily FTSE100 futures price dynamics and their relation to the underlying index.

This research has three main objectives. The first aim is to test lead-lag (or causal) relationships between FTSE100 spot and futures markets. Since the U.K. indices are quote-driven, they reflect more up-to-date information than transaction-price-based indices such as the US indices, which are more subject to infrequent trading problems (Yadav and Pope, 1990, 1994). Empirical evidence of lead-lag relationships between FTSE100 index and futures prices is therefore less subject to bias introduced by infrequent trading. It also provides an alternative and fair context in which to test the hypothesis that stock index futures usually provide a 'price discovery' vehicle for the underlying index. The second aim is to have a more formal test of the validity of the cost of carry model for pricing stock index futures (this deserves high attention given that the cost of carry model is very widely used for pricing stock index futures). Third, arbitrage with zero initial investment that yields risk-free profit may be the most attractive trading strategy one can have. However, since the real world can never be
perfect and frictionless, deviations from the arbitrage-induced cost of carry relationship, or 'mispricing' of futures contracts, may not always represent arbitrage opportunities. Another aim of this research is to identify the degree of 'mispricing' that could really yield an arbitrage profit.

1.2.1 Sample selection

This research is based on nearly 20 years of daily data for FTSE100 spot and futures prices, from 28/10/1986 to 30/12/2005, split into two sub-samples for empirical analysis for the following reasons. An important linkage between stock index futures and the underlying spot markets should be maintained by arbitrage. This involves simultaneously taking a short position in one market and an opposite long position in the other market if the prices for the same underlying asset are misaligned in the two markets. Due to quickly changing prices, immediate implementation of both legs of an arbitrage strategy is important. Delayed execution in either market is potentially risky and could result in an arbitrage loss. Therefore, trading systems and transaction costs in both spot and futures markets are important factors that could affect arbitrage. Given that arbitrage affects spot-futures price dynamics, it is then important to analyse empirically price relationships between spot and futures markets over a time period with a relatively stable trading system and transactions cost structure in each market, which is used as the principle for selecting the sample period and sub-samples in this research.

Arbitrage transactions costs in London market have been most affected by two discrete microstructure changes at the LSE and one microstructure change at LIFFE. In 1986, London became one of the first financial centers to see trading move out of the pits, off the market floor and on to the telephone. This was achieved by using a screen-based electronic bulletin board SEAQ (Stock Exchange Automated Quotations System), historically known as 'Big Bang' on 27 October 1986. Fixed commission charges were abolished since then. 28 October 1986 was therefore set as the start date of our sample period. The trading system employed by the LSE was further transformed on 20 October 1997 from the dealership system (SEAQ) to a fully electronic order-driven trading system (SETS). The implementation of the cash leg of an index arbitrage through SETS becomes much quicker and cheaper than through SEAQ. On 30 November 1998, LIFFE also replaced the old 'open outcry' system with a new
electronic platform, LIFFE CONNECT. The implementation of the futures leg of an index arbitrage through LIFFE CONNECT should also be much quicker and cheaper than before. The arbitrage linkage between FTSE100 spot and futures markets, and hence the dynamics between them, might therefore be different before and after 20 October 1997, and before and after 30 November 1998. Ideally therefore, we should consider and compare three sub-sample periods separated by 20 October 1997 and 30 November 1998, that is, 28/10/1986–17/10/1997, 20/10/1997–27/11/1998 and 30/11/1998–30/12/2005.

However, the time period 20/10/1997–27/11/1998 is short and there is an insufficient number of observations for efficient estimation of the models tested in this thesis. As will be seen in later empirical chapters, over the sub-sample period 30/11/1998–30/12/2005, the estimation of the DCC-TGARCH model (Chapter 5) has a problem with convergence and the results of estimating the threshold VECM (Chapter 6) suffer from bias created by a very small sample for the upper regime. To avoid these problems, we have finally chosen 28/10/1986–17/10/1997 as the first sub-sample when FTSE100 securities were traded through a dealership system and FTSE100 futures were traded through ‘open outcry’ and 27/10/1997–30/12/2005 as the second sub-sample, when FTSE100 shares were traded through an order-driven system and FTSE100 futures were traded on an electronic platform (only after 30/11/1998). Statistical tests have also been used to check for parameter stability in the two sub-samples. Furthermore, as the aggregate market displayed an overall rising trend in the first sub-sample but mainly a falling pattern in the second sub-sample, this provides a natural experiment of the impact of market sentiment on stock index arbitrage.

Small changes in arbitrage transactions costs should have also occurred at other times, which may have affected index arbitrage and hence spot-futures price dynamics, but this impact is expected to be minor in comparison with the effects resulting from market microstructure changes in the two markets, namely the ‘Big Bang’ and the moves to SETS on LSE and the employment of LIFFE CONNECT by LIFFE.

---

2 In fact the second sub-sample starts from 27/10/1997 rather than 20/10/1997 (omitting 5 observations) to allow for the necessary lags in the error-correction model.
1.2.2 Data and methodological choices

Daily closing price for FTSE100 index and daily settlement price for FTSE100 futures were used in this study. As daily closing time in the spot market has usually been different from daily settlement time in the futures market, our study was based on the assumption that daily closing price for FTSE100 index was synchronous with daily settlement price for FTSE100 futures. Given that each individual futures contract has a limited life span, the single time series of FTSE100 futures price used in this study was created artificially, as what is normally done in the literature. The price series of successive futures contracts were spliced together in which observations shift to the next near contract on each expiration day. The final three months’ futures prices before expiration of each futures contract therefore form a single time series of futures price.

To calculate the cost of carry for FTSE100 futures price, the LDMON (London Discount Market Overnight Rate) was used as a proxy for the risk free interest rate and the realized ex post dividend yield was used as a proxy for the forecasted dividend yield for FTSE100 index.

Given the arbitrage-induced cointegrating relationship between futures and the underlying spot markets, it has become routine in the literature that the empirical examination of dynamics between a stock index and its futures price should be conducted within a vector error correction model (VECM). Following Green and Joujon (2000), we have included three other variables, $k_t$, $z_t$, and $T-t+k_t$, in addition to lagged own- and cross-market returns and a lagged error correction term in a traditional standard VECM for spot and futures returns. Here, $k_t$ is the time interval between two consecutive observations, added to capture the effect of the passage of time on returns; $z_t$ is a dummy variable used to account for 'artificial' futures price jumps at each contract roll-over when prices from consecutive futures contracts are spliced into a single time series; $T-t+k_t$ is the time to maturity, and is included to induce stationarity in the basis for each individual futures contract.

Short-run lead-lag relationships (or causality-in-mean) and long-run cointegration between FTSE100 spot and futures markets are first examined within the basic linear VECM framework (Chapter 4). Restrictions imposed by the cost of carry model on the linear VECM are derived, allowing a formal test of the cost of carry model within the
basic VECM framework. Given the widely documented evidence of ‘volatility clustering’ and ‘leverage effects’ in financial time series, causality-in-mean is further examined within the basic VECM but extended by DCC-TGARCH (threshold GARCH) to account for time-varying volatilities and possible asymmetric effects in volatilities (Chapter 5). With standardized residuals estimated from the DCC-TGARCH model, a two-step CCF (cross correlation function) test is also used to test for causality-in-variance, which, together with the evidence on causality-in-mean, provides a more complete and accurate picture of the dynamics between FTSE100 spot and futures markets.

The basic linear VECM analysis and the DCC-TGARCH analysis are both based on the assumption that the arbitrage-induced adjustment toward equilibrium is continuous and at constant speed, regardless of the size of the deviation from equilibrium. Provided that arbitrage in the real world is neither risk-free nor free of transaction costs, the adjustment cannot be continuous and cointegration between spot and futures markets could be nonlinear depending on the presence or absence of arbitrage. Therefore the basic VECM analysis is further extended to a threshold VECM framework, in which the assumption of continuous adjustment was relaxed and an examination of threshold cointegration and regime-dependent price dynamics is performed (Chapter 6).

1.3 Contributions to Knowledge

A VECM framework first used by Green and Joujon (2000) for French data is used as the basic model structure in this research. Based on this, the thesis has led to at least five contributions to the literature. First, by examining FTSE100 spot-futures price dynamics over a sample period of nearly 20 years, this thesis provides some new and important evidence about the behaviour of index spot and futures prices in London and more generally. In particular, since the FTSE100 index is a quote-based index, it provides a more reasonable environment for testing lead-lag (or causal) relationships between a stock index and its futures price that is less subject to the infrequent trading problem. Furthermore, the empirical findings are compared over two sub-samples. These subsamples are differentiated by (i) different trading systems employed by both the LSE and LIFFE and (ii) different performance in the aggregate market. During the first sub-sample a dealership system and ‘open outcry’ were used at the LSE and LIFFE respectively, but this changed to an order-driven trading system at the LSE and an
electronic platform at LIFFE in the second sub-sample. Performance was characterized by an overall bull market in the first sub-sample but a rather lengthy early 2000s recession in the second sub-sample. These changes provide a natural experiment of the impact of market frictions and market expectations on arbitrage and hence on the futures price process.

Second, following Green and Joujon (2000), restrictions implied by the cost of carry model on the VECM are carefully derived. Empirical coefficients of the VECM for FTSE100 spot and futures returns therefore allow a formal test of the validity of the cost of carry model for pricing stock index futures. The findings suggest that during the first sub-sample, when transaction costs in both markets are high, the arbitrage-induced cost of carry relationship between FTSE100 spot and futures markets is rejected by the data. During the second sub-sample when transaction costs are comparatively low, the cost of carry relationship tends to be valid. This implies that the performance of the cost of carry model as a benchmark model for pricing stock index futures is more reliable when market conditions are closer to the perfect markets assumption.

Third, the basic VECM of Green and Joujon (2000) is extended to include DCC-TGARCH to model ‘volatility clustering’ and ‘leverage effects’ in volatilities and dynamic conditional correlation (DCC) between FTSE100 spot and futures returns. Because of the similarity of the two-step estimation approach for estimating the DCC-GARCH models of Engle (2002) and the two-step CCF test for causality-in-variance proposed by Cheung and Ng (1996), a CCF test for causality-in-variance can be performed within the DCC-TGARCH model for FTSE100 spot and futures returns. With this combination, both volatility processes for FTSE100 spot and futures prices and causality-in-variance between them can be examined within an easy-to-implement multivariate GARCH – the DCC-TGARCH framework. The evidence of time-varying conditional correlation between FTSE100 spot and futures markets has important implications for those using stock index futures in their daily risk management (e.g. fund managers): to have an efficient hedge against market risks, hedge ratios should be monitored carefully and updated frequently to follow the changing correlations between spot and futures markets.

The fourth contribution comes from the findings yielded by estimating a three-regime threshold VECM for FTSE100 spot and futures returns. The three regimes reflect the
consideration that arbitrageurs tend to react to a large enough negative mispricing or a large enough positive mispricing in a previous period, while no arbitrage occurs in the inner regime when the mispricing is small and the marginal cost of arbitrage exceeds the marginal benefit. I find that for the first sub-sample, the arbitrage-induced error correction term is insignificant in the upper regime when futures contracts are overpriced; for the second sub-sample, the arbitrage-induced error correction term is insignificant in the lower regime when futures contracts are underpriced. These seemingly puzzling results have important implications. One is that arbitrage may not be practicable under some market conditions, either because it is difficult to find counterparties for the arbitrage transactions, or because there is significant risk associated with arbitrage. Another implication is that since some other factors might be relevant and important in determining futures prices, especially under certain market conditions, the cost of carry model may not always be suitable for pricing stock index futures, and hence the observed 'mispricing' of futures contracts may not always represent an arbitrage opportunity.

Last but not least, this thesis contributes to the literature regarding a reasonable selection of the threshold variable in the threshold cointegration analysis. Arbitrage is to a great extent affected by transaction costs involved in trading in the two markets. Accordingly, in an empirical study using threshold cointegration analysis, the location of the thresholds should rely primarily on transaction costs associated with arbitrage because the thresholds separate different regimes depending on the presence or absence of arbitrage. The comparability of the threshold variable with transactions costs is therefore a promising idea. ‘Percentage mispricing’, as proposed by MacKinlay and Ramaswamy (1988), was first used in this research as the threshold variable. Percentage mispricing can be directly compared to the actual transaction costs of arbitrage because both are expressed as a proportion of the underlying index value. The threshold values yielded from estimating the threshold VECM therefore reflect the average transaction costs for most arbitrageurs. These are more reliable and fair than subjective estimation because they are based on the observed market pattern that the price dynamics switch between regimes depending on the presence or absence of arbitrage activities.
1.4 Structure of the Thesis

The rest of this thesis consists of 6 chapters organised as follows.

Chapter 2 introduces background knowledge about futures markets, with particular attention to stock index futures, including descriptions of the mechanics of the futures market, the advantages of trading in futures markets and some important economic functions of futures markets. This is followed by an explanation of the cost of carry model for pricing stock index futures and various testable hypotheses implied by the model. Challenges to the cost of carry model are also briefly analysed. Finally the chapter discusses market conditions for index arbitrage in the U.K. in particular, which are useful for understanding the empirical findings on price dynamics between FTSE100 spot and futures markets.

Chapter 3 is a review of relevant papers on stock index futures, organized according to two main lines of research – namely the ‘mispricing’ of stock index futures, and its implication for index arbitrage, and ‘lead-lag’ (or causal) relationships between spot and futures prices. The review forms a basis for the research and raises several questions for this research to explore.

Chapter 4 first presents the data used in this research. The dynamics between FTSE100 spot and futures markets are then examined within a linear VECM framework. The restrictions implied by the cost of carry model are also derived. With the empirical evidence on the behaviour of FTSE100 index and futures prices, a formal test of the cost of carry model for pricing FTSE100 futures is performed.

Chapter 5 re-examines the spot-futures price dynamics of the FTSE100 by extending the basic linear VECM developed in Chapter 4 to include DCC-TGARCH. This models ‘volatility clustering’ and ‘leverage effects’ in the volatilities of FTSE100 spot and futures prices. Furthermore, based on empirical results estimated from the DCC-TGARCH model, the two-step CCF test for causality-in-variance is applied to examine ‘volatility spillover’ effects between the two markets.

Chapter 6 further extends the basic VECM to a threshold VECM, which is used to investigate regime-dependent price dynamics between FTSE100 spot and futures
markets. The average transactions costs faced by those arbitrageurs most active in the market are also estimated. Furthermore, based on the empirical results, the chapter includes a discussion of the difficulties associated with index arbitrage under unusual market conditions. If index arbitrage is not always practicable, cost of carry models based on simple no-arbitrage arguments are almost certainly not always suitable for pricing stock index futures.

Chapter 7 provides an overall conclusion to the thesis and discusses further possible research directions implied by its findings.
Chapter 2 An Introduction to Futures and Stock Index Futures

This chapter provides a basic introduction to futures markets. Some features relevant to stock index futures, in particular to stock index futures in the UK, which is the focus of this research, will also be addressed. The introduction is organized in the following order. First, we explain briefly the mechanics of futures markets and some advantages of trading in futures. Second, we summarize the important economic functions of futures markets. Third, the price relationship between a stock index itself and a stock index futures contract is analyzed and the cost of carry model, so far the most widely used model for pricing stock index futures, is explained. Two alternatives to the cost of carry model for pricing stock index futures are also briefly presented. Fourth, we have an analysis of real world transaction costs and potential risks associated with index arbitrage and the introduction of index-tracking funds on index arbitrage.

2.1 Futures markets and stock index futures

2.1.1 The mechanics of futures markets

Futures contracts are standardized agreements to buy or sell some underlying item on a specified future date, the settlement date, at a price negotiated at the time of the futures transaction. Futures contracts trade on organized exchanges. The basic purpose of an exchange is to provide an organized marketplace, with uniform trading rules and standardized contracts. Futures contract standardization reduces transaction costs since it obviates the need to negotiate all the terms of a contract with every transaction, like forward contracts. Each futures exchange has an affiliated clearinghouse\(^3\). The basic function of a clearinghouse is to clear futures contracts, with its operations as follows. It matches and records all trades and guarantees contract performance by interposing itself between buyers and sellers, assuming the role of counterparty to the contract for both parties. The clearinghouse guarantee relieves traders of the default risk of the other party to the contract. To protect itself against counterparty credit risk, the clearinghouse requires a margin deposit as collateral against default on each futures contract to ensure that traders honor their contractual obligations.

\(^3\) The clearinghouse can either be a separately incorporated membership association or organized as a division within the exchange corporation. In order to insulate the exchange from the legal liability of the clearing corporation, however, the latter practice is often avoided (Duffie, 1989).
The clearinghouse also requires all buyers and sellers to realize any gains or losses on their outstanding futures positions at the end of each trading day through a daily settlement procedure known as 'marking to market'. Specifically, the clearinghouse collects payments, called variation margin, from all traders incurring a loss and transfers the proceeds to those traders who have earned a profit on the current trading day. If a trader’s margin account falls below a specified minimum, called the maintenance margin, he faces a margin call requiring the deposit of additional margin money. This has the equivalent effect as if every futures position were liquidated at the current trading day’s closing price and the trader begins the next trading day with a commitment to buy or sell the underlying item at the previous day’s closing price. The daily settlement mechanism therefore strictly limits counterparty credit risks and prevents accumulated losses on each futures position. In addition, the clearing house monitors the financial integrity of its members and also maintains a guarantee fund to cover default in case it does occur. By all these measures — record keeping, margin requirements, financial oversight of members, and a guarantee fund — a clearinghouse supports the financial integrity of a futures market. Futures contracts, especially financial futures, are usually based on cash delivery rather than physical delivery. This allows the design of some futures contracts for which either physical delivery is difficult, such as a stock index or no underlying deliverable asset even exists, such as an inflation index (Duffie, 1989; Kuprianov, 1992).

2.1.2 Some advantages of trading in futures markets

Trading in futures markets has some advantages over trading in spot markets. Stock index futures, in particular, have a variety of attractive features for a trader who wishes to trade a portfolio of shares corresponding to the index. First, short selling shares is restricted in many countries. For example, the “uptick” rule in the US prevents shares being sold short unless the last price movement was up. In the UK only market makers (including equity options market makers) can borrow UK shares, and they must do so via money brokers and are required to pay a fee of roughly 1% per annum of the value of the shares. On the other hand, a short position is easy for futures. Indeed, it is necessarily the case that exactly half the trades in futures involve taking a short position.

Second, transaction costs of trading in futures markets, including commission, the bid-ask spread, adverse selection costs, the opportunity cost of funds used in paying the
initial margin and which are set aside to meet variation margin, are lower than in spot markets. The bid-ask spread for shares is often markedly larger than for index futures because a market maker in particular shares, who is exposed to both systematic and unsystematic risk, will require a larger bid-ask spread than a market maker in index futures, who is only subject to systematic risk, as unsystematic risk is diversified away by holding a position in a widely diversified portfolio. Adverse selection costs due to private information are greater for shares than index futures since diversification of the index portfolio of securities reduces the effect of information asymmetry. The opportunity cost of funds tied up in a futures position is also significantly lower than that in a spot position, because a futures position requires futures traders to deposit only a small fraction of the contract value as the initial margin (leverage effect). Moreover, the initial margin can be posted in the form of Treasury Bills, which allows the investor to earn the risk-free interest rate on his/her capital in the initial margin account.

Third, futures markets are much more liquid than spot markets. By deliberate design, the futures contracts of the same maturity are perfect substitutes for each other and the validity of a futures contract is independent of the identity of the buyer and the seller. This means that the buyer does not need to be careful about the identity of the seller, nor need the seller be concerned about the identity of the buyer (Telser, 1981, 1986). Contract standardization, along with the clearinghouse guarantee, allows futures contracts to be a transferable agreement. These characteristics contribute to high liquidity in futures markets.

Fourth, trading futures and trading shares are subject to different tax regulations. For example, in the UK stamp duty of 0.5% (1% prior to Big Bang) is payable on share transactions (market makers and charities can avoid paying stamp duty on purchases of shares that are sold within seven days), but no stamp duty is charged on futures transactions. Furthermore, earnings from futures trading by pension funds and authorized unit trusts in the UK are exempt from taxation (prior to July 1990, authorized unit trusts, pension funds, and investment trusts were exempt from capital gains tax on futures transactions for investment purposes, but were liable to pay tax on any trading income) (Sutcliffe, 2006).
2.2 The economic functions of futures and stock index futures

2.2.1 Risk transfer

Risk transfer has been viewed as the primary economic function of futures markets. This view suggests that hedgers and speculators, among others, are two main types of participants in futures markets. Futures contracts are essentially insurance contracts for hedgers, providing insurance against uncertain terms of trade on spot markets at the delivery date of relevant futures contract (Duffle, 1989). By trading in futures, hedgers can effectively transfer the price risk associated with owning (or a potential purchase of) an item to someone else by selling (or buying) a futures contract on the same item. Futures market speculators, on the other hand, are risk takers. They buy or sell futures contracts solely in an attempt to profit from price changes. The potential risk is that the futures price may move in the opposite direction to their expectation, or their bet. The existence of futures speculators makes it possible and convenient for hedgers to transfer risks to someone who is willing to take them. Just as Adam Smith explained, speculators – in pursuing their own interests – are making the markets more liquid and stable.

High liquidity and low transaction costs in futures markets are important reasons why futures contracts are so popularly used by some market participants in their risk management activities. A futures contract is a transferable agreement, i.e. it can be bought or sold through the clearinghouse at any time before maturity to liquidate an open futures position. It is cheap and easy to trade in futures markets. These characteristics of futures contracts are especially attractive to market makers (or dealers) and other intermediaries whose cash positions change continually, along with their exposure to price risk. For example, securities dealers must stand ready to buy and sell securities in response to customer orders and therefore constantly change the composition of their holdings. Thus they hedge using futures contracts because the great liquidity and low transaction costs in futures markets mean that a futures hedge can be readjusted frequently with relatively little difficulty and at minimal cost. But on the other hand, contract standardization, while contributing to futures market liquidity, practically implies that futures contracts will not be perfectly suited to the needs of each hedger, whose planned transaction dates rarely coincide with standardized futures delivery dates. Thus most hedgers using futures contracts must unwind their futures
positions before the contracts mature and therefore face 'basis risk' — risk that the futures price is not directly tied to the underlying index, except for the final settlement price on the expiration date (Figlewski, 1984), which is much easier to avoid with customized forward contracts\textsuperscript{4} (Telser 1981, 1986; Kuprianov, 1992).

Stock index futures provide investors an efficient and cost-effective means of hedging a portfolio of shares or equity index options against systematic risk; meanwhile, they also provide a convenient and cheap means for market timing — the strategy of making buy or sell decisions of financial assets (often stocks) by attempting to predict future market price movements based on the outlook for the aggregate market, rather than for a particular financial asset. Fund managers have been found to actively use stock index futures in order to achieve an optimal combination of risk and expected return for the entire trust or investment fund. For example, on the one hand, they use stock index futures in their market timing strategy to benefit from expected market changes; on the other hand, they use stock index futures in hedging the funds against market risk (Stein, 1986).

2.2.2 Price discovery

The futures market also provides an important price discovery vehicle for the underlying assets. This is fulfilled by all market participants involved in futures trading. For example, when speculators and hedgers trade futures contracts, they are acting on their own market analysis of supply and demand conditions in the underlying assets. Their trading, in turn, affects futures prices. The discovery of the underlying asset's price is completed as individuals' information and analysis become visible to the wider market through their trading in futures markets (Patel and Tkac, 2007). The introduction of futures contracts also provides a platform for arbitrageurs, who exploit pricing anomalies between spot and futures markets to produce a riskless profit. Arbitrage activities should therefore ensure that any information first impounded in futures markets should immediately be reflected in the underlying spot markets, and vice versa. This has the effect of stopping prices in both markets from diverging too far

\textsuperscript{4} In a remark on 'financial derivatives' given on March 19, 1999, Alan Greenspan pointed out that of the $33 trillion notional value of outstanding derivatives contracts at year-end, only $4 trillion were exchange-traded derivatives; the remainder were off-exchange or over-the-counter (OTC) derivatives. The greater use of OTC derivatives reflects the attractiveness of customized over standardized products.
away from their equilibrium values and plays an important role in mitigating irrationality in the stock market.

2.2.3 Other functions

Futures trading itself represents a new means of investment and has led to the creation of trading strategies based on futures. Because of the institutional arrangements, a futures market is in many cases more convenient for trading than the corresponding spot market, so futures contracts are sometimes used as substitutes for spot transactions. Since the price of a stock index future is likely to move in tandem with the prices of the underlying stocks, it should in theory give the same return as owning the stocks. The stock index future conveys no rights to the dividends but is cheaper and easier to buy and may be exempt from certain taxes and charges to which stock ownership is subject. Trading using futures could involve, for instance, spread trading (simultaneous purchase of one future and the sale of another in anticipation of exploiting expected changes in the relative prices of two futures) and trading of futures options. Investing via the use of stock index futures would allow stock portfolio managers to use the leverage effect of futures trading in increasing their exposure to movements in a particular index or market sector without having to actually purchase shares directly. Futures contracts are useful even to those who do not trade them, since futures prices provide some public indication of future demand and supply conditions of the underlying stocks. (Duffie, 1989; Sutcliffe, 2006).

In summary, the introduction of futures contracts has greatly extended the range of investment and risk management strategies available to market participants and plays an important role in maintaining stock market efficiency. Stock index futures, in particular, play a non-negligible role in establishing a healthy and stable financial system, which is an essential task for each economic unit.

2.3 The price relationship between index and its futures price

The best-known model for pricing stock index futures is undoubtedly the cost of carry model, developed by Cornell and French (1983a). The derivation of this model is based on a simple no-arbitrage argument that two different assets, or combinations of assets, that yield the same return should sell for the same price. Otherwise, arbitrage profit is
available. This no-arbitrage argument therefore predicts that the futures price of an asset should just equal its spot market price on the maturity date of the futures contract because buying futures on the maturity date of a futures contract is equivalent to buying the underlying asset in the spot market. Buying a futures contract before the contract maturity date fixes the cost of acquiring the underlying asset in the future. But the cost of future availability of the asset can also be fixed in advance by buying and holding that asset. Buying and holding the asset entails opportunity costs in the form of interest foregone on the funds used to purchase the asset and, in some instances, storage costs (negligible for financial assets). On the other hand, buying and holding assets may bring benefits such as convenience yield and dividend payments. The benefits yielded by physical holding of the asset offset a fraction of any financing costs. The difference between the cost of financing the purchase of an asset in the cash market and the benefit arising from holding the asset is known as the net cost of carry. Since a futures position can be replicated by spot positions in the stock and T-bill markets, the no-arbitrage argument predicts that the net cost of carry should determine the relationship between futures and spot prices. This leads to the derivation of the cost of carry model, which expresses the futures price in terms of the underlying stock index value, the risk-free interest rate and the dividend yield for the index (Kuprianov, 1992).

Using the simple no-arbitrage argument, the derivation of the cost of carry model also relies on some simplifying assumptions as follows:

1. Capital markets are perfect and frictionless (i.e. no taxes or transactions costs; all transactions are by cash; no restrictions on short sales; perfectly divisible assets).
2. The interest rate is deterministic.
3. Each investor faces a single, known, risk-free borrowing and lending rate.
4. Dividends are paid continuously at a constant rate.

If borrowing and lending take place at a constant, continuously compounded interest rate $r$, and the basket of stocks representing the index pays dividends continuously at a constant dividend yield $d$ (the dividend yield is defined to be the ratio of the present value of dividends received to the cash index price at time $t$), the no-arbitrage
The relationship between the current price of a futures contract for delivery at time \( T \), \( F_t \), and the price of the underlying index, \( S_t \), should be given by

\[
F_t = S_t e^{(r-d)(T-t)}
\]  

(2.1)

where \( T-t \) represents time to maturity of the futures contract. Rearranging equation (2.1) by taking logarithms on both sides gives

\[
f_t = s_t + (r-d)(T-t)
\]  

(2.2)

or alternatively,

\[
\Delta f_t = \Delta s_t - (r-d)
\]  

(2.3)

where \( f_t = \ln F_t \), \( s_t = \ln S_t \), \( \Delta f_t = \ln F_t / F_{t-1} \), and \( \Delta s_t = \ln S_t / S_{t-1} \). The logarithm of the futures price exceeds the logarithm of the spot price by the 'net cost of carry' for purchasing the index portfolio and carrying it until the futures contract matures. The net cost of carry, \((r-d)(T-t)\), is equal to the interest forgone on the funds tied up in the spot position minus the dividend accumulated on the spot position from time \( t \) until the maturity of the futures contract, at time \( T \). Clearly, therefore, as the futures contract approaches maturity, the stock index futures price converges to the underlying index price. Hereafter, while using logarithms of the futures and spot prices, for convenience reference will be made simply to the 'futures price' and the 'spot price'.

If equation (2.1) did not hold, stock index arbitrage transactions would occur immediately in an efficient market. If index futures were under-priced, arbitrage would entail taking a long position in the futures contract and a simultaneous short position in the underlying index, investing the sale proceeds at the risk-free interest rate. If index futures were over-priced, opposite transactions would be carried out, financed by borrowing at the risk-free interest rate. In both cases, arbitrageurs would be able to lock in a risk-free profit with zero initial investment. This can be accomplished through a low cost and highly efficient computer-assisted trading technique on a large scale, known as 'programme trading'. Such arbitrage transactions would quickly bring spot and futures prices back to the no-arbitrage equilibrium defined by equation (2.1). An
implication of equation (2.3) is that if stock index and futures markets are frictionless and functioning efficiently, and if the interest rate and dividend yield are deterministic, spot and futures price changes will be perfectly contemporaneously correlated and no lead-lag relationship will exist. Furthermore, the variance of futures price changes will be perfectly contemporaneously correlated with the variance of the spot price changes (MacKinlay and Ramaswamy, 1988; Stoll and Whaley, 1990; Abhyankar, 1995, 1998).

Despite its popularity, there have been both theoretical and empirical challenges to the cost of carry model in the literature. Two representative theoretical challenges to the cost of carry model are Hemler and Longstaff (1991) and Hsu and Wang (2004). Hemler and Longstaff (1991) point out that the cost of carry model is only a partial equilibrium model because it assumes that the stock market is exogenous. This could fail to capture the dynamic interactions between spot and futures markets and therefore could result in bias in estimating the fair futures price. Furthermore, the cost of carry model is actually a forward (not futures) pricing model. To apply it to stock index futures, one must assume that forward and futures prices are equal. Since futures contracts are resettled daily, futures and forward prices need not be equal if interest rates are stochastic. Also, other factors such as market volatility might also have significant explanatory power for stock index futures prices. Motivated by these considerations, Hemler and Longstaff (1991) develop a closed-form general equilibrium model of stock index futures prices in a continuous-time production economy characterized by stochastic interest rates and fluctuating levels of market uncertainty. The general equilibrium model allows for interactions between futures, cash, and credit markets. The risk-free interest rate, the stock index futures price, and the level of the stock index itself are all determined endogenously as part of the equilibrium. Hsu and Wang (2004) argue that since capital markets are not perfect or frictionless, standard arbitrage is exposed to such large risk. As a result, arbitrage mechanisms cannot be complete, particularly for index arbitrage, and price expectations and risk aversion may also play an important role in determining index futures prices in real futures markets. Moreover, arbitrage mechanisms may not work in certain circumstances (for example during stock market crashes, where spot markets or derivative markets are illiquid, or where the relevant markets do not exist). Based on these concerns, Hsu and Wang (2004) develop a pricing model of stock index futures in imperfect markets that incorporates the factor of price expectations and partial (rather
than complete) arbitrage as a natural principle. Empirically, there have also been substantial challenges to the cost of carry model (the literature review in chapter 3 provides details). Nevertheless, the predominant position of the cost of carry model in providing a benchmark model for pricing stock index futures has been maintained from its introduction until now, thanks to the simple and yet solid theoretical premise it is based on.

2.4 Index Arbitrage in the UK

2.4.1 Hurdles to arbitrage

The cost of carry relationship is often used as a theoretical benchmark to evaluate stock index futures pricing efficiency. The real world, however, is not as perfect and frictionless as in theory. Since stock index arbitrage involves transactions in both spot and futures markets, marginal transaction costs associated with arbitrage include round-trip trading costs (including commissions and bid-ask spread\(^5\)) in both cash and futures markets, transaction taxes (or stamp duty\(^6\)) and the costs of borrowing fixed interest capital and index stocks\(^7\). Cash market marginal transaction costs can be close to zero for market makers and for institutional arbitrageurs who can negotiate index arbitrage trades at mid-market prices (Yadav and Pope, 1994). Market makers and brokers/dealers in the UK are exempt from stamp duty if they buy and sell shares within seven days, which would result in a lower transaction cost for this category of arbitrageurs compared to other arbitrageurs (who must pay stamp duty of 0.5% when buying shares). Borrowing costs are only faced by arbitrageurs who do not have capital in treasury bills (or equivalent fixed interest securities) and index component stocks. Real-world transaction costs would create a band (or an arbitrage window) in which the futures price is free to diverge from the single theoretical 'fair' value defined by the cost of carry model, in which no arbitrage is available. The width of the band is determined by the transaction costs of the most favourably situated arbitrageurs in the market, usually the market makers and/or brokers. The implication is that while futures contracts may often be observed to be 'mispriced', these do not necessarily represent profitable arbitrage opportunities.

\(^5\) Bid-ask spread is sometimes called market impact cost and the market impact cost is usually increased at times of volatility. There is often an expected discount on large sell orders and premium on large buy orders.

\(^6\) A stamp duty is paid on purchase of shares of any British companies on the London Stock Exchange. Round-trip transaction tax (or stamp duty) is 0.5% of stock value (1% before the 27/10/1986 'Big Bang').

\(^7\) Capital in Treasury bills is needed for arbitrage strategies involving long stock and short futures and capital in index stocks is needed for arbitrage strategies involving long futures and short stocks.
Arbitrage is further influenced by three other factors.

(1) Stockholders can defer the realization of capital gains, and hence the associated tax liability, while futures traders necessarily pay taxes in the year that capital gains are realized. The cash investors' tax timing option that is not available to futures traders should therefore result in the actual futures price being less than the theoretical 'fair' value (Cornel and French, 1983b).

(2) Institutional restrictions on the short sale of stocks can inhibit the arbitrage process. In the UK, only registered market makers have special stock borrowing privileges. In case of futures being underpriced, non-market makers can only undertake arbitrage transactions if their trading books are already long in stocks (Yadav and Pope, 1994). If factors (1) and (2) are important, futures contracts would tend to be underpriced and futures mispricing — the difference between the actual futures price and the theoretical futures price — would tend to be more negative.

(3) Early unwinding and rollover options might have the countervailing effect of reducing the arbitrage window. Arbitrageurs have the option to unwind (or close) their arbitrage positions prior to expiry once the no-arbitrage condition is met. At any time before delivery the initial mispricing may be reversed, and new mispricing occur in the opposite direction. Under this circumstance, early unwinding of the arbitrage position will lead to an extra riskless profit. Therefore, early unwinding avoids the risk of being unable to close out the share position at the futures delivery price and raises the possibility of making a profit from an arbitrage position, over and above the riskless arbitrage profit that was initially locked in. With the early unwinding option, therefore, arbitrageurs may enter the market even before the mispricing is sufficient to cover transaction costs, in the expectation that the mispricing will subsequently be reversed and can be unwound early at a net profit (Sutcliffe, 2006). But early unwinding involves additional market impact costs associated with closing the arbitrage position, so it is worth unwinding early only if the magnitude of reversed mispricing is sufficient to cover the additional market impact costs (Yadav and Pope, 1990). On the other hand, instead of liquidating the arbitrage position when the current futures contract matures, arbitrageurs have the option to roll forward their futures position into the next near contract if the next near contract is mispriced in the same direction as the initial arbitrage. Since there is no need to trade shares, there are no additional transaction costs
in the stock market and no additional stamp duty in the UK case. The only incremental transaction costs involved in this new arbitrage position are those for trading the futures. Thus, establishing an arbitrage position gives the additional embedded option to delay unwinding by rolling over the arbitrage maturity with very low additional transaction costs, which may also encourage some arbitrageurs to enter the market even before the mispricing is sufficient to cover transaction costs (Yadav and Pope, 1990; Sutcliffe, 2006).

In addition to transaction costs and the other factors listed above, there are other potential risks involved in index arbitrage.

(1) Risks associated with dividend uncertainty: since dividends are paid unevenly throughout the year, the cost of carry model should reflect only those dividends to be paid from the time of entry into the futures contract to the settlement date. This can be highly subjective, given different forecasts as to amount and timing.

(2) Risks of delayed execution: due to delayed execution, prices quoted when arbitrage orders are submitted are not necessarily the execution prices obtained for the orders and the ‘mispricing’ actually obtained by arbitrageurs is often less than when the trade is initiated (Neal, 1996). The delayed execution is especially serious with the cash leg of the arbitrage because market makers normally have excess stock when there is an arbitrage-related sell program and are short of stock when there is an arbitrage-related buy program (Yadav and Pope, 1994).

(3) The daily re-settlement cash flows of futures contracts are uncertain because of price uncertainty.

(4) Tracking error risks arise if only a subset of index stocks is used to track the index value. The tracking error risk is higher during highly volatile periods.

These risks and uncertainties will lead to a risk premium and an effective increase in the width of the arbitrage window. These risk factors are expected to be positively related to time to expiration. Therefore the width of the arbitrage window, hence the

---

8 But dividends are usually announced at least a few weeks before ex-dividend date, major problems are not expected if the analysis is restricted to the near contract.
magnitude of ‘mispricing’, is expected to be positively related to time to expiry (MacKinlay and Ramaswamy, 1988).

In terms of index arbitrage in the UK, some factors in addition to transaction costs and arbitrage risks have also restricted the involvement of managed funds in index arbitrage, especially in earlier years. For example, prior to July 1990, authorized unit trusts, pension funds, and investment trusts were exempt from capital gains tax on futures transactions for investment purposes, but were liable to pay tax on any income earned from trading futures. The fear of being deemed to be trading (rather than investing) in futures contracts, and so liable to taxation, may have deferred such institutions from using futures as part of their investment strategy. But the 1990 Finance Act made clear that all futures trades conducted after 26th July 1990 by pension funds and authorized unit trusts would be exempt from taxation, which has encouraged these institutions to participate in futures markets. Furthermore, the rules governing the investment policy of a fund may prohibit the use of futures contracts, so preventing the fund from engaging in index arbitrage. Even if it is legally allowed to trade futures, the approval of trustees is required to permit funds to trade in index futures. The trustees may be unfamiliar with index futures and impose limits on the proportion of the fund that can be invested in derivatives, which also limits the participation of funds in index arbitrage.

To summarise, index arbitrage in the UK is much cheaper and easier for market makers, who can avoid stamp duty on purchase of shares sold within seven days and who have special stock borrowing privileges (which permits taking a short position in shares). They also hold a position in almost all the stocks in the index basket, and therefore they can avoid ‘tracking error’ risk that may arise when using only a subset of the index stocks in implementing index arbitrage. The ‘in-house’ trading of market makers also reduces the risk associated with execution lag in index arbitrage. Market makers therefore have considerably lower transaction costs and lower arbitrage risks than other potential arbitrageurs and are the main participants in index arbitrage in the UK market (Sutcliffe, 2006).

2.4.2 The effect of index-tracking funds on index arbitrage

The introduction of exchange-traded index-tracking funds (ETFs) that attempt to track all kind of indices offers an alternative to futures as a way of trading the index basket of shares
with lower transaction costs. Since ETFs trade on the market, investors can carry out the same types of trades that they can with a stock. For instance, investors can sell short (subject to market regulations), use a limit order, use a stop-loss order, and invest as much or as little money as they wish (there is no minimum investment requirement). Also, an ETF is continuously priced throughout the day and allows the user to react to adverse or beneficial market conditions on an intraday basis. This stock-like liquidity allows an investor to trade the ETF for cash throughout regular trading hours, and often after hours through electronic communication networks (ECNs). Furthermore, trading in ETFs is not subject to stamp duty. This makes purchase and sale of the index basket of shares in spot market much quicker, easier and cheaper than before (Sutcliffe, 2006). The use of ETFs in index arbitrage will also avoid the tracking error risk of implementing index arbitrage strategies with only a subset of index stocks. Therefore, the introduction of ETFs is believed to have the effect of improving index arbitrage and reducing the magnitude of mispricing of index futures. This hypothesis is supported by practices in different markets. For example, Park and Switzer (1995) find that the introduction of TIPs (Toronto Index Participations) in 1990 has led to an increase in futures arbitrage using TIPs rather than shares, and that daily mispricing of TSE35 futures decreased in magnitude after the introduction of TIPs, supporting the view that TIPs facilitate arbitrage. Switzer et al. (2000) find that the introduction of SPDRs (the ETF on the S&P index) on 29 January 1993 mitigated the extent of pricing errors of S&P500 futures and suggest that market efficiency has been enhanced by SPDRs. Kurov and Lasser (2002) find that the introduction of Cubes (the Nasdaq-100 Index Tracking Stock) in March 1999 has led to a reduction in the size and frequency of violations of the no-arbitrage boundary. Furthermore, there appears to be an increase in the speed of the market response to observed violations. A bit later, in April 2000, Barclays Global Investors (BGI) also introduced an exchange traded fund (ETF) designed to track the FTSE 100 index. Among many other advantages, the ETF purchases made on the LSE are not directly subject to stamp duty and the ETF can also be sold short. This should make purchase and sale of FTSE100 index portfolio in the spot market (and hence index arbitrages between FTSE100 spot and futures markets) much quicker, easier and cheaper than before.
Chapter 3 Futures Mispricing, the Implications for Arbitrage, and Causal Relationships between Stock Index and Futures Prices

Literature Review

3.1 Introduction

The aim of this chapter is to review the literature to discover open questions regarding stock index futures that require further investigation. To summarize the papers dealing with stock index futures pricing, this chapter is organized according to two primary issues that have been addressed in the literature. The first issue is about futures pricing efficiency, or whether stock index futures are 'mispriced', and the second is about causal relationships between spot and future markets, both of which may generate signals for arbitrage. Examination of both 'mispricing' of futures contracts and causal relationships between spot and futures markets is motivated by identifying opportunities for arbitrage, since the 'no-arbitrage' condition provides the base for deriving the cost of carry model for pricing stock index futures.

The cost of carry model is so far the most widely used model for pricing stock index futures. It is suggested that if the cost of carry relationship between futures and spot markets is not met, stock index arbitrage transactions would occur immediately to exploit a riskless profit with zero initial investment. Such arbitrage transactions would quickly bring spot and futures prices back to no-arbitrage equilibrium. Earlier research also suggests that an implication of the cost of carry model is that if stock index and futures markets are frictionless and functioning efficiently, and if interest rate and dividend yield are deterministic, spot and futures price changes should be perfectly contemporaneously correlated and the variance of changes in the futures price will be perfectly contemporaneously correlated with the variance of changes in the spot price (no lead-lag, or causal, relationship should exist between the two markets). From an econometric point of view, the cost of carry model implies that the futures and spot prices are cointegrated in the long run and the spot-futures price dynamics should therefore be conducted within the framework of an error correction model. The literature regarding the price relationship between stock index and futures markets has
therefore focused either on the ‘mispricing’ of stock index futures contracts and the
implication for arbitrage (e.g. MacKinlay and Ramaswamy, 1988; Klemkosky and Lee,
1991; Yadav and Pope, 1990, 1994), or testing within a VECM the causal relationships
between spot and futures price changes (e.g. Ghosh, 1993; Wahab and Lashgari, 1993),
or testing within a basic VECM but extended with GARCH for the residuals for both
causality-in-mean and causality-in-variance between the two markets (e.g. Chan et al.,
1991; Tse, 1999b; Bhar, 2001; Zhong et al., 2004).

Empirical investigations have presented a large amount of evidence on substantial and
persistent ‘mispricing’ of futures contracts and ‘lead-lag relationships’ between futures
and spot prices. This casts doubt on the validity of the strict cost of carry model for
pricing stock index futures, because a profitable arbitrage opportunity should not last
long in the market if advances in computer trading can easily take advantages of such
an opportunity when it appears. This has led to further more recent research regarding
the pricing of stock index futures and the cost of carry model from different
perspectives. For example, some have analysed limitations of the cost of carry model
and developed alternatives for pricing stock index futures (Hemler and Longstaff, 1991;
Hsu and Wang, 2004). Balke and Fomby (1997) consider the limitations of the
traditional error correction model in analyzing cointegrating relationships induced by
arbitrage. They propose a threshold error correction model to analyze regime-dependent
price dynamics related to the presence or absence of arbitrage activities. Green and
Joujon (2000) derive the restrictions imposed by the cost of carry model on the VECM
for spot and futures returns, and point out that some causal relationships between spot
and futures markets are not necessarily inconsistent with the cost of carry model. They
seek to perform a more formal test of the cost of carry model within the same VECM
framework for examining the lead-lag relationships.

3.2 Futures contract mispricing and the implications for arbitrage

Based on a simple no-arbitrage argument, the cost of carry model was developed by
Cornell and French (1983a) for pricing stock index futures. To earn a risk free profit
from a zero investment arbitrage strategy is one of the best investments one can ever
make. Such a profitable arbitrage opportunity should not last long in the market because
advances in computer trading can immediately take advantages of this opportunity
when it appears. Then, if the cost of carry model is valid for pricing stock index futures

27
one should not see futures contracts being persistently mispriced when using the ‘fair’ price given by the cost of carry model as a benchmark. Otherwise, it means either that the cost of carry model may not be suitable for pricing stock index futures, or that the futures market is seriously inefficient and persistent arbitrage opportunities are present in the market. Persistent inefficiency of futures markets is unlikely to be present given the low transaction costs and the ease of trading in the futures market. These issues have generated a lot of research interest in futures contract ‘mispricing’ examinations of the cost of carry model and the implications for arbitrage.

3.2.1 Empirical investigations of futures mispricing

Many empirical studies have focused on futures contract mispricing and the availability of profitable arbitrage opportunities. Surprisingly, to both academics and practitioners, researchers have reported substantial and sustained deviation in futures prices from their theoretical values given by the cost of carry model. Early empirical studies focused on American market, particularly the S&P500 futures prices. For example, Cornell and French (1983a, b) report that during the first seven months of stock index futures trading, from March through September 1982, the S&P500 and the NYSE composite index futures were often below the spot price, contrary to the prediction of the cost of carry model with perfect market assumptions. MacKinlay and Ramaswamy (1988) report that the mispricing of S&P500 futures contracts from June 1983 through June 1987 had a tendency to persist above or below zero for substantial lengths of time and that the magnitude of the mispricing was positively related to time until maturity. Klemkosky and Lee (1991) find that for the March 1983 through December 1987 period, S&P500 futures contracts were generally more overpriced and that both the frequency and the degree of mispricing diminished as the expiration day approached. For the Australian market, Bowers and Twite (1985) provide evidence of the All Ordinaries Share Price Index Futures prices violating the no-arbitrage bounds in the 1980s. Heaney (1995) demonstrates poor descriptive ability of the simple cost of carry model for the All Ordinaries Share Price Index Futures contract for the period from March 1983 to March 1990. For FTSE100 futures contracts traded in the UK, Yadav and Pope (1990, 1994) find that over the period from April 1986 to March 1990, the far contract and the near contract tended to be mispriced in the same direction, the absolute magnitudes of mispricing for the far contract appeared to be considerably larger than for the near contract, and the mispricing tended to persist over long periods. They argue
that since FTSE100 index values are based on firm quotes on which respective market makers are obliged to trade up to very large sizes, the stock index values thus represent actually tradable values and identified arbitrage opportunities are therefore actually exploitable and economically significant. Sutcliffe (2006) gives a comprehensive summary of empirical evidence from different markets on index futures mispricing.

### 3.2.2 Interpretation of mispricing: does it always represent arbitrage opportunity?

Cornell and French (1983a, b) point out that the perfect market cost of carry model is based on several fairly strict assumptions, which may mask various important factors that could explain the observed futures prices and mispricing. It is known that arbitrage trading is not free of cost in the real world. Transaction costs, dividend uncertainty, taxes, tracking error risk, the risk of delayed execution, short sale restrictions in the stock market, the lack of arbitrage capital and exchange-imposed position limits are all important factors restricting arbitrage. The most commonly cited explanations for the observed mispricing of futures contracts in the literature are summarized below.

**Marking to market and stochastic interest rates**

Empirical studies on mispricing of index futures have centered around arbitrage pricing errors derived from the cost of carry relationship. The standard cost of carry model for pricing forward contracts given by equation (1) is applicable to futures contracts only if the interest rate is non-stochastic (Cox *et al.*, 1981). However, interest rates have usually been found to be stochastic. Since futures contracts are resettled daily, the futures price must reflect any unanticipated interest earnings or costs from financing the mark-to-market cash flow in the futures position (Yadav and Pope, 1990). To demonstrate the effect of stochastic interest rates on futures pricing, we first assume that changes in share prices are positively correlated with changes in interest rates. Due to daily resettlement, the gains on a long futures position arising from increasing share prices can be realized immediately and reinvested at the higher new interest rate, while the loss on a long futures position arising from decreasing share prices can be financed at a lower new interest rate. A long position in a forward contract, however, is not affected by interest rate movements in the same way. As a result, a positive correlation between changes in risk-free interest rate and changes in share prices suggests that the futures price is greater than the forward price. A similar argument suggests that if there
is a negative correlation between changes in risk-free interest rate and changes in share prices, the futures price will be less than the forward price (Cox et al., 1981; Hull, 2000).

**Seasonal dividends**

The standard cost of carry model assumes that the dividend flow from the underlying security is constant and continuous. Cornell and French (1983a, b) point out that seasonal dividends, the fact that most firms only pay dividends quarterly at most, may have a significant effect on the observed index futures prices. It is less serious a problem for an index than for individual stocks because the lumpiness of dividend payments is reduced by collecting stocks into portfolios. On the other hand, since many firms issue their quarterly dividends at about the same time, index portfolios may also display seasonal fluctuations in their dividend flow. Consequently, the simple cost of carry model, which assumes constant and continuous dividend payment over the full year, will tend to overprice a futures contract if there are higher than average dividend payments during the life of the futures contract, and *vice versa*.

Cornell and French (1983a, b) have extended the perfect markets model by introducing seasonal dividends and stochastic interest rates. Following Cornell and French (1983a, b), Yadav and Pope (1990, 1994) propose that with non-constant discrete dividends and stochastic interest rates, the arbitrage free ‘fair’ value at time $t$ of an index futures contract maturing at time $T$ is given by:

$$ F_t = S_t e^{r_{t,T}(T-t)} - \sum_{w=t+1}^{T} D_w e^{r_{w,T}(T-w)} $$  \hspace{1cm} (3.1)

where $r_{t,T}$ is the interest rate on a risk-free discount bond that is issued at time $t$ and maturing at time $T$, $r_{t,w,T}$ is the forward interest rate at time $t$ for a loan at time $w$ that matures at time $T$ and $D_w$ is the aggregate dividend paid by underlying stocks on day $w$ during the life of the futures contract ($t < w < T$). The second term in equation (3.1), $\sum_{w=t+1}^{T} D_w e^{r_{w,T}(T-w)}$, equals the time $T$ value of the dividends an index portfolio holder receives over the life of the futures contract. The model still assumes that the dividends
are known at time $t$. However, Yadav and Pope (1994) argue that because dividends are typically announced at least a few weeks before the ex-dividend date\(^9\) and those companies going ex-dividend before futures maturity should have declared their dividends, there is essentially no dividend uncertainty with index arbitrage transaction over the relatively short life of a futures contract. The dividend certainty assumption is especially innocuous if research analysis is restricted to the near contract, which is frequently the case. Cornell and French (1983a, b) point out that since both the long-term interest rate $r_{t,T}$ and the forward interest rate $r_{t,w,T}$ can be observed at time $t$, the model applies whether the term structure moves deterministically or stochastically.

**The influence of tax and the tax timing option**

Cornell and French (1983a, b) have further examined the effect of taxes on futures pricing, which are assumed to be zero in the standard cost of carry model. They assume a simple tax structure in which interest and dividend payments are taxed at the ordinary income tax rate, $i$, capital gains and losses on shares are taxed at capital gains rate, $g$, and profits and losses from futures trading are taxed at the futures rate, $f$. While the arbitrage transaction may have zero profit under the assumption of zero tax, the post-tax profit or loss might not be zero, or *vice versa*, suggesting the need to consider the tax effects in the no-arbitrage model. Cornell and French (1983a, b) point out that income tax will reduce the effective dividend yield and interest rate. Reduced dividend yield has the effect of raising the futures price, while a reduced interest rate has the effect of lowering the futures price. Since the interest rate is usually above the dividend yield, the interest rate effect dominates the dividend yield effect. As a result, the ordinary income tax will have the effect of lowering the actual futures prices. Cornell and French (1983a, b), and later Klemkosky and Lee (1991), have extended the simple cost of carry model to allow for taxation effects. Considering two investment strategies in Table 3.1, either (A) investing in $(1-g)/(1-f)$ index futures contracts or (B) borrowing money and investing in one unit of stock index, both strategies have the same cash flow at time

\(^9\) There are four major dates in the process of a company paying dividends: declaration date, ex-dividend date, date of record and date of payment. The record date is designated by a company such that dividends are paid to the list of shareholders who hold stock on the record date. In order to allow time for the ‘settlement’ of a stock purchase, stock exchanges set a date – generally two business days prior to the record date – the ex-dividend date. Someone who purchases the stock on (or after) the ex-dividend date will not receive the dividend, as the purchase will not ‘settle’ by the record date, and therefore the buyer will not be on the list of shareholders to which the company pays its dividends.
To avoid arbitrage, the two investments should have the same initial cash flows at time \( t \), that is, \([(1 - g)F_t + gS_t + D]e^{-\gamma (T - t)} - S_t = 0\). As a result, the no-arbitrage futures pricing model can be generalized to allow for the defined simple tax structure

\[ F_t = \{S_t[e^{(1+i)\tau (T-t)} - g] - \sum_{w=t+1}^{T}(1-i)D_we^{(1+i)\tau (T-w)}\}/(1-g) \] (3.2)

**Table 3.1 The relationship between futures and cash index prices with taxes**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Cash flow at time ( t )</th>
<th>Cash flow at time ( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Buy ( \frac{1-g}{1-f} ) futures contracts</td>
<td>0</td>
<td>((1-g)(S_T - F_t))</td>
</tr>
<tr>
<td>B. Buy one unit of index and borrow</td>
<td>(-S_t)</td>
<td>(S_T - g(S_T - S_t) + D)</td>
</tr>
<tr>
<td></td>
<td>([((1-g)F_t + gS_t + D]e^{-\gamma (T-t)})</td>
<td>(-(1-g)F_t - gS_t - D)</td>
</tr>
<tr>
<td></td>
<td>([(1-g)F_t + gS_t + D]e^{-\gamma (T-t)} - S_t)</td>
<td>((1-g)(S_T - F_t))</td>
</tr>
</tbody>
</table>

where \( D = \sum_{w=t+1}^{T}(1-i)D_we^{(1+i)\tau (T-w)} \).

Empirical evidence seems to be mixed. Cornell and French (1983a, b) report that, even after introducing stochastic interest rates, seasonally varying dividends and a simple tax structure, the richer cost of carry model still leads to predicted futures prices that are generally higher than actual prices of the S&P500 and NYSE composite index futures from March through September of 1982. On the other hand, Klemkosky and Lee (1991) find that, when taxes are considered, the frequencies of violations of no-arbitrage pricing conditions and excess returns notably decrease for S&P500 futures contracts for the period March 1983 through December 1987.

Constantinides (1983) demonstrates that stockholders in many countries have a valuable ‘tax timing option’ – the option to sell an asset and claim a loss for the purpose
of tax refunds or not to sell the asset and defer the payment of capital gains tax – which affects the equilibrium stock price. In response to this, Cornell and French (1983a, b) suggest that, because investors who hold futures contracts do not have this tax timing option, it must have the effect of further lowering the futures price. Yadav and Pope (1994) point out that for those tax-exempt institutions and those investors (such as arbitrageurs or floor traders) who cannot hold the cash index into the next tax period, the tax timing option is ‘not valuable’. *Ceteris paribus*, the tax timing option should be more valuable when the index is more volatile since this option increases the value of the spot index relative to the index futures. Hence an increase in spot volatility will lead to futures becoming underpriced relative to a theoretical futures price that ignores the tax timing option – that is, there should be a negative relationship between the ‘mispricing’ (the difference between actual futures prices and theoretical fair values) and spot volatility if the tax timing option is a relevant factor for index futures pricing (Yadav and Pope, 1994). However, Yadav and Pope (1994) document a positive relationship between ‘mispricing’ and spot volatility for the FTSE 100 from 1986 to 1990, suggesting that the tax timing option is of no value for UK indices.

**Transaction costs and arbitrage risks**

It is now well known that the presence of transaction costs in the real world has the effect of allowing the futures price to fluctuate within a band around its theoretical value without representing a profitable arbitrage opportunity. The width of this band (or arbitrage window) should be dictated by the transaction costs of the most favourably situated arbitrageurs. Since stock index arbitrage involves transactions in both markets, transaction costs include round-trip commissions and market impact costs (due to bid/ask spread) in the stock and futures markets, the costs of borrowing fixed interest capital and index stocks, and transaction taxes (or stamp duty) in the stock market in the UK (Yadav and Pope, 1990, 1994). If the arbitrage position is held until expiration, no market impact costs are incurred by unwinding the index arbitrage positions, since the stock position can be closed at the market-closing price, which is the same as the

---

10 The market impact cost is determined by the market maker’s bid/ask spread. A market maker’s quotes are firm for only a fixed transaction size. Large orders may move the bid (ask) quote downward (upward). In fast markets, due both to the backlog of orders and quotes that may not be current, there is a greater possibility that shares will not be available at the quote made when the order was initially placed, so large orders are often filled in segments at several different prices. Normally, there is an expected discount on large sale orders or a premium on large buy orders. The magnitude of the market-impact cost reflects, among other things, the liquidity and depth of a market (Fleming *et al.*, 1996).
terminal futures price (MacKinlay and Ramaswamy, 1988; Brennan and Schwartz, 1990).

Early unwinding and rollover options might have the countervailing effect of reducing the arbitrage window. Arbitrageurs have the option to reverse their positions prior to expiry if the mispricing changes sign. Since early unwinding involves additional market impact cost associated with closing the futures position (transaction costs in the cash market can be close to zero for market makers and institutional arbitrageurs), it is worth implementing an early unwinding if the magnitude of reversed mispricing is sufficient to cover or exceed the additional market impact cost in the futures market. Arbitrageurs also have the option to roll forward their futures position into the next near contract if the direction of mispricing of the next near contract is the same as the mispricing direction when the arbitrage position was initiated. Since there is no need to trade shares, the new arbitrage program initiated on the expiration date does not involve additional transaction costs or stamp duty in the stock market. The only incremental transaction costs involved in this new arbitrage position are those for trading the futures. Therefore, the rollover option would yield profit to arbitrageurs if the extent of mispricing on the expiration date exceeds the transaction costs associated with initiating the arbitrage plus the new incremental costs associated with rollover the arbitrage. With early unwinding and rollover options, arbitrageurs may enter the market even before the mispricing reaches the transaction cost boundaries (Yadav and Pope, 1990).

MacKinlay and Ramaswamy (1988) argue that the width of the arbitrage window should be constant over the life of the futures contract if it is decided by transaction costs alone, because transaction costs are independent of the remaining maturity of the contract. They find that the magnitude of ‘mispricing’ for S&P500 futures contracts is positively related to time until maturity and suggest that the arbitrage window is also affected by other risk factors influenced by time to expiration, such as dividend uncertainty, uncertain marking-to-market flows, and ‘tracking error’ risk of implementing index arbitrage strategies with only a subset of index stocks. These risks are expected to be larger with longer times to expiration.

Another risk involved in arbitrage transactions is the risk of delayed execution, that is, prices quoted when arbitrage orders are submitted are not the execution prices obtained for the orders. Execution lags may have an adverse impact on arbitrage profit - that is,
the mispricing actually obtained by the arbitrageur is likely to be less than when the trade is initiated (Neal, 1996). On the other hand, Yadav and Pope (1994) point out that the high degree of persistence in mispricing suggests that the possibility of delayed execution may not be a serious risk for index arbitrageurs. Overall, these risks and uncertainties will lead to a risk premium and an effective increase in the width of the arbitrage window.

**Short sale restrictions**

Short selling shares involves various difficulties, including the costs of borrowing shares, restrictions on short selling in some markets and even a ban on short selling in many countries. Arbitrage for the case when the index futures are underpriced is therefore difficult for those who are not net long in spot positions. The simple no-arbitrage condition might then become $F_{t,T} < S_t e^{(r-d)(T-t)}$. Short sale restrictions may lead to a widening of the arbitrage window defined above by permitting greater underpricing of index futures. Neal (1996) argues that short sale restrictions are unlikely to affect cash-futures mispricing because institutional traders are typically net long in stocks and can avoid short sale restrictions by selling the stock directly. Kurov and Lasser (2002) investigate whether the introduction of the Nasdaq-100 Index Tracking Stock (referred to as Cubes) has led to significant changes in Nasdaq-100 spot-futures pricing relationship. Possible reasons for improvements include smaller transaction costs, shorter execution lags, smaller tracking risk, and fewer short-sale restrictions in the cash index market when Cubes are used in arbitrage. They do find that futures price boundary violations become less frequent after the introduction of Cubes and the violations are eliminated faster. However, they also find that the elimination of short sale restrictions after introduction of Cubes (which can be shorted on the downtick) had little impact on the effective transaction costs of short arbitrage trades (defined as buying futures and short selling the spot index). They conclude that their findings support the view of Neal (1996) that institutional traders can avoid the short-sale restrictions by using direct sales, and that the significant reduction in the frequency of boundary violations is primarily a function of lower transaction costs, shorter execution lags, and the reduction of tracking risk.
The impact of index composition

A recent paper by Bortoli and Frino (2006) suggests that transaction costs, dividend uncertainty and 'tracking error' risk proposed by MacKinlay and Ramaswamy (1988) are all a function of the composition of equity indices. They argue that for futures contracts listed on broad versus narrow indices, arbitrage transactions will involve increased transaction costs, more dividend uncertainty, greater difficulty of constructing portfolios of stocks to track the underlying index and higher 'tracking error' risk. Consequently, compared with futures listed on narrow indices, broad based futures contracts will exhibit greater 'mispricing' and wider arbitrage windows around the fair value. Bortoli and Frino (2006) further argue that risks associated with dividend uncertainty and tracking error decrease as futures contracts approach expiry. However, with longer times to expiration, the risks associated with broad-based contracts are proportionally greater than those associated with narrow-based contracts. The difference in the magnitude of mispricing between futures contracts based on broad and narrow indices is therefore positively related to time until maturity, ceteris paribus.

Alternatives to the cost of carry model for pricing stock index futures

Given the limitations of the cost of carry model, as analyzed above, some researchers have tried to seek more accurate alternatives to the cost of carry model for pricing stock index futures. For example, Hemler and Longstaff (1991) develop a closed form general equilibrium model for pricing stock index futures that allows for stochastic interest rates as well as the influence of market volatility on futures prices. The general equilibrium model also allows both spot and futures prices to be endogenously determined by the model, rather than taking spot price as given. By testing the restrictions imposed by the cost of carry model and the general equilibrium model within the same framework they find some support for the general equilibrium model, whereas the cost of carry model is invalid for pricing stock index futures if interest rates are stochastic. They also find that market volatility has significant explanatory power for stock index futures prices, providing a possible explanation for the perceived failure of the cost of carry model in turbulent and volatile markets. Hsu and Wang (2004) construct a theoretical foundation to explain why the price expectation of the underlying asset should enter the pricing formula of stock index futures in imperfect markets. They develop a pricing model of stock index futures that accounts for the price
expectation and incompleteness of arbitrage. These research findings have important implications for practitioners and, while the cost of carry model is based on a seemingly reasonable no-arbitrage argument, one should be careful in interpreting it in the real world. Because capital markets can never be as perfect and frictionless as is normally assumed for theoretical purposes, the arbitrage mechanism cannot be complete, particularly for index arbitrage. Since the cost of carry model ignores other factors (such as volatility and price expectation) it may fail to define true fair values for stock index futures contracts.

3.3 Causal relationships

A traditional inference drawn from the strict cost of carry model is that spot and futures price changes should be perfectly contemporaneously correlated (there should not be any lead-lag (or causal) relationship between them). This hypothesis on short-run price dynamics has been proposed in earlier research to test the validity of the cost of carry model and/or the price efficiency of stock index futures (Stoll and Whaley, 1990). On the other hand, some argue that in the presence of market frictions and trading costs, new information will tend to be incorporated with greater speed in one market relative to the other (Abhyankar, 1995; Fleming et al., 1996). More recent research also suggests that lead-lag relationships between spot and futures prices conform to the cointegrating relationship between the two price series rather than provide evidence against the cost of carry model (Engle and Granger, 1987; Green and Joujon, 2000). There has been considerable research interest in price dynamics between spot and futures markets. With theoretical developments and the aid of advanced computer technology, there have been substantial improvements in econometric modeling for empirical studies of economic and financial issues. Such improvements include the creation of ARCH models by Engle (1982) and their generalization to GARCH models by Bollerslev (1986), the initial introduction of cointegration and error correction model (ECM) by Engle and Granger (1987), and the development of threshold cointegration and threshold error correction model (ECM) by Balke and Fomby (1997). Accordingly, empirical testing of the relationships between spot and futures markets has become more sophisticated: VECM modeling has taken over from the traditional (spurious) VAR study, threshold cointegration analysis has addressed limitations associated with
linear cointegration analysis and GARCH models have been widely used to provide more complete insight into both causality-in-mean and causality-in-variance.

3.3.1 Why do causal relationships exist between spot and futures prices?

Non-synchronous trading

'Non-synchronous trading' of the component shares of a stock index is one of the most widely quoted explanations of why the cash index may lag the futures price. The 'nonsynchronous trading' problem refers to the fact that stock indices are recorded at the end of trading using the last transaction prices of component stocks. If those stocks did not trade at the same time (some stocks are traded less frequently than others due to transactions costs concern) and did not trade exactly at the time when the stock index was recorded, then descriptions of the characteristics of the index would be subject to nontrading-induced biases. The best known characteristic is the spurious positive autocorrelation of index returns. The non-synchronous trading problem might also be due to pure technical reasons, for example, time delays in the computation and reporting of the stock index value (Stoll and Whaley, 1990). As a result of non-synchronous trading, the reported value of the cash index may not be able to update information quickly and would be a stale indicator of the actual index value. Reported index futures, on the other hand, can continuously reflect all currently available information, because investors can trade the whole index in one transaction at lower transaction costs in the futures market (MacKinlay and Ramaswamy, 1988; Stoll and Whaley, 1990). Chan (1992) suggests that spurious conclusions caused by nonsynchronous trading could be avoided by using narrowly based indices when investigating lead-lag relationships, because the infrequent trading problem is less serious. Green and Jouflon (2000) point out that the non-synchronous trading problems that result from either infrequent trading or purely technical reasons are likely to be most serious with short time intervals. This argument would suggest that futures prices tend to lead the underlying index. The time span of the lead is likely to be relatively short if the index is being updated – probably at most an intra-day period. Yadav and Pope (1990, 1994) argue that studies of the lead-lag relationship between stock index and futures price based on US data potentially suffer from measurement errors induced by the use of last transaction prices in index computation. On this argument, non-synchronicity should be less of a problem for the UK because London Stock Exchange
indices are computed from the available mid-market quotes. These should reflect more up-to-date information than indices calculated from the last transaction prices of the component shares, since the quotes represent market makers’ (or market participants’) estimates of the current tradable value of component shares.

Empirical evidence on the non-synchronous trading hypothesis is mixed. MacKinlay and Ramaswamy (1988) find that the S&P500 index series is positively autocorrelated at the first lag but that the autocorrelation disappears as the interval length is increased. In contrast, the autocorrelations of the futures series are close to zero at all lags. Shyy et al. (1996) find that CAC-40 futures lead the cash price when transaction price data are used. These are most likely to be subject to non-synchronous trading and stale price effects. However, cash leads futures when bid/ask quote-mid-point data are used (reflecting more current information). Their findings provide support for the argument of Yadav and Pope (1990, 1994) and suggest that previous results showing that futures lead cash may be primarily due to the use of transactions price data. But Alphonse (2000) points out that the finding of Shyy et al. (1996), that cash leads futures, may be due to their use of the second nearest futures contract. Alphonse (2000) reports opposite results when using the first nearest futures contract. Theobald and Yallup (1998) examine how the actual (observed) price partially adjusts to the true (intrinsic) value across stock index futures and cash markets in the UK. They report that price adjustments are fuller in futures markets – while the futures price adjustments are not significantly different from 100%, the cash market adjustments can be as low as 80%. When non-synchronicities in the cash market are adjusted for, price adjustments are increased and tend not to be significantly different from one. Their findings lend support to the theory that non-synchronicity will cause futures price to lead the underlying cash index.

On the other hand, Stoll and Whaley (1990) report that S&P500 and MMI index futures lead cash prices by about 5 minutes on average, even after stock index returns have been purged of infrequent trading effects using an ARMA process. But Chan (1992) points out that since Stoll and Whaley (1990) assume the parameters of the ARMA model to be constant, it may not adequately eliminate the infrequent trading components if the effects of infrequent trading are changing. Chan (1992) finds that the lead-lag relation between futures and component stocks of MMI (where infrequent
trading is not a problem) does not suggest that futures lead only less actively traded stocks. In fact, even for stocks like IBM, AT&T, and Exxon, which seem to be more actively traded than the MMI futures, the returns are led by futures returns. Fleming et al. (1996) report that S&P100 and S&P500 futures lead the cash index after infrequent trading effects are controlled. These findings suggest that the lead of futures price over the cash index cannot be completely explained away by the nonsynchronous trading problem.

**Transaction costs**

Given that transaction costs exist and differ across markets, price discovery will tend to occur first in the lowest-cost market, because information-based trades are executed where they produce the highest profit (Fleming et al. 1996). Lower transaction costs in futures markets than in spot markets is an important reason why price movement in futures markets should lead price movement in the underlying spot markets. Adverse selection costs, inventory costs, order processing cost, and liquidity are four important reasons why transaction costs in futures markets are lower than that in spot markets. The models of Subrahmanyam (1991) and Gorton and Pennacchi (1993) show that a factor favouring trading in index futures, especially for uninformed liquidity traders, is that the diversification of the portfolio of securities reduces the effects of private information. Therefore, adverse selection costs induced by information asymmetry will be much lower for trading stock index futures than for trading individual securities in the spot market. Market makers manage risk and control inventory levels by adjusting their bid and ask quotes\(^\text{11}\), which has a direct impact on transaction costs for market participants. Transaction costs due to market makers’ inventory costs are small in futures trading and traders find it easy to control inventory in futures markets (Manaster and Mann, 1996; Tse, 1999a). Order processing costs are lower in the futures market because stock index futures are equivalent to one security and transactions for all the component shares can be conducted with only one order, compared to separate cash orders for each security in the spot market. Futures contract standardization and the

---

\(^{11}\) Since dealers (or market makers) must stand ready to sell securities to customers at their ask prices and purchase from them at their bid prices, inventory management is the essence of their business. Dealers intend to offset the transactions relatively quickly rather than acquire a long-term investment portfolio. They are often unable to even out their positions quickly without suffering large loss. Consequently, they face risks in managing their temporary portfolio of inventory (Stein, 1986, p.190).
participation of the clearing house in each transaction also make futures contracts highly liquid and very cheap to trade (Kuprianov, 1992).

Indeed, trading S&P500 futures costs only about 3% of the cost of trading an equivalent portfolio of index stocks, according to Fleming et al. (1996). Berkman et al. (2005) also find that the effective half (bid/ask) spread in the futures market is small compared to stock market for FTSE 100 index. Abhyankar (1995) finds that the sizeable reduction in transactions costs in the London equity market resulting from the Big Bang of 1986 appears to have reduced both the size and the significance of the lead of the FTSE100 futures returns over the index returns series.

**Leverage effect**

Index futures are regarded by some investors as a better investment vehicle than the underlying indices because they require only a small capital outlay for the initial margin, equal to a fraction of the contract value, hence allowing an investor to obtain a desired leverage that is unavailable in stock markets. Moreover, if the margin is posted in the form of Treasury Bills\(^{12}\), which allows the investor to earn risk-free interest rate on his/her capital, the opportunity cost to the investor of taking a position in futures markets is actually zero (Antoniou and Garrett, 1989). This implies that the futures price should lead the underlying index. However, Fleming et al. (1996) provide evidence contrary to the argument that the futures price should lead the underlying asset price because of leverage privileges associated with derivatives.

**Different information arrival**

`Economy-wide' versus `firm-specific' information

When new information arrives, traders can choose whether to exploit this knowledge in the futures or spot markets\(^{13}\). On arrival of firm-specific information, informed traders will probably choose to buy or sell individual shares rather than index futures because the movement in the index will be much smaller than in the share prices of the affected companies. In response to firm-specific information, therefore, the index changes might lead the futures price changes if the affected shares represent large weights in the index.

---

\(^{12}\) Treasury bills can be used to satisfy margin requirements on futures contracts.

\(^{13}\) Some informed investors (e.g. pension funds) are prohibited from trading derivatives and can trade only on the stock market, which could result in lead of spot price over futures price.
and therefore have caused a change in the index. On the other hand, for economy-wide information that relates to the economy in general rather than to particular companies, the assessment of the information by informed traders may first be reflected in futures price changes because futures markets have many advantages over spot markets, including higher liquidity, lower transaction costs, easily available short positions, low margins and rapid execution (Sutcliff, 2006). Indeed, Chan (1992) finds that the futures market is the main source of market-wide information.

"Good" or "bad" news

Due to short selling restrictions in the spot market, futures price changes are expected to lead spot price changes more clearly in response to bad news than in response to good news. Empirical evidence on this is mixed. For example, Puttonen (1993) reports that changes in the Finnish index futures price predict changes in the cash price better under bad news than under good news, implying that the short selling constraint is a major factor causing the lead-lag relationship. However, Chan (1992) finds that for MMI the futures price changes do not seem to lead the index changes only under bad news, suggesting that short selling restrictions are not a main cause for the observed lead-lag relationship.

3.3.2 Cointegration and causality-in-mean

Early empirical studies of the lead-lag (or causal) relationship between futures and spot prices, concentrating on the US market, were based on a bivariate VAR in returns:

\[
\Delta s_t = \alpha_0 + \sum_{j=1}^{M} \alpha_{1j} \Delta s_{t-j} + \sum_{j=1}^{M} \alpha_{2j} \Delta f_{t-j} + e_{1t} \tag{3.3a}
\]

\[
\Delta f_t = \beta_0 + \sum_{j=1}^{N} \beta_{1j} \Delta s_{t-j} + \sum_{j=1}^{N} \beta_{2j} \Delta f_{t-j} + e_{2t} \tag{3.3b}
\]

Here \( \Delta s_t \) and \( \Delta f_t \) are the first differences of spot and futures prices respectively. Since almost all asset prices are I(1), whereas returns are I(0), to avoid spurious regression both spot and futures prices are first differenced to induce stationarity. \( e_{1t} \) and \( e_{2t} \) are white noise errors. Kawaller et al. (1987) and Stoll and Whaley (1990), among others,
have used VAR in studying lead-lag relationships between stock indices and their futures prices. They find that futures prices tend to lead the underlying spot prices in response to new information.

**Linear cointegration**

Granger's representation theorem states that if two I(1) series are cointegrated, there exists a linear combination of these series that is I(0) and an error correction representation will exist (Engle and Granger, 1987). Futures and spot prices have both been found to be I(1) in the literature and are linked by the cost of carry relationship, and therefore are expected to be cointegrated in the long run. Systems in which variables are cointegrated can be characterized by the error correction model (ECM). Wahab and Lashgari (1993), Ghosh (1993) and Brenner and Kroner (1995) point out that the traditional VAR of equations (3.3a) and (3.3b) is misspecified since it excludes the error correction term \((f_{t-1} - \mu s_{t-1})\), where \(\mu\) is the cointegrating coefficient defined so that \((f_{t-1} - \mu s_{t-1})\) is I(0). The error correction term is used to model the effect of divergence from equilibrium on short-run price movement. The economic intuition is that, while short-run dynamic components of the futures and spot price relationship may diverge from the long-run equilibrium defined by the cost of carry, the previous period's divergence from equilibrium may carry predictive information for future price movements in spot and/or futures markets. The VECM in the form of equations (3.4a) and (3.4b) was thereafter proposed as a proper framework in which to study causal relationship between futures and spot markets.

\[
\Delta s_t = \alpha_0 + \sum_{j=1}^{M_1} \alpha_{1j} \Delta s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j} \Delta f_{t-j} + \alpha_3 (f_{t-1} - \mu s_{t-1}) + e_{1t} \tag{3.4a}
\]

\[
\Delta f_t = \beta_0 + \sum_{j=1}^{N_1} \beta_{1j} \Delta s_{t-j} + \sum_{j=1}^{N_2} \beta_{2j} \Delta f_{t-j} + \beta_3 (f_{t-1} - \mu s_{t-1}) + e_{2t} \tag{3.4b}
\]

Empirical studies using vector error-correction models (VECM) have found evidence of one-way causality from futures to spot (e.g. Tse (1995) for the Japanese market), two-way causality between futures and spot (e.g. Ghosh (1993) for the US market and Wahab and Lashgari (1993) for both the US and UK markets), or even one-way
causality from spot to futures (e.g. Shyy et al. (1996) and Green and Joujon (2000), both for French market).

While the existence of causal relationships between futures and spot markets has traditionally been viewed as evidence against the cost of carry model, Granger's representation theorem implies that if spot and futures prices are cointegrated then there must exist some causal ordering between spot and futures, although the direction of this ordering is not known a priori. Green and Joujon (2000) also point out that once the restrictions imposed by cost of carry are set out, it transpires that it is not true that one-way causality from futures to spot (or vice versa) is necessarily inconsistent with cost of carry. On the contrary, under reasonable assumptions, cost of carry implies that there necessarily exists some causal ordering between spot and futures. Another possible limitation of virtually all the empirical studies is that they adopt a trading time model while ignoring the possible effects of different intervals between trades or otherwise-reported prices (Green and Joujon, 2000).

Given the important role of interest rates in the cost of carry relationship, Brenner and Kroner (1995) further argue that if interest rates and therefore the differential (i.e. the difference between the risk-free interest rate and the dividend yield for the index) have a stochastic trend and are nonstationary, then spot and forward (or futures) prices will not be cointegrated by themselves, and the differential must be included in the system to find the like cointegrating relationship. But their proposition assumes that the time to expiration of the futures contract, \((T-t)\), is fixed, and that the time of expiration, \(T\), is changing. If this were true and if the interest rate had a stochastic trend, the net cost of carry, \((r-d)(T-t)\), could be nonstationary, affecting the cointegrating relationship between spot and futures prices. In fact, the time to expiration of a futures contract is not fixed but declines as time passes and the contract approaches the fixed delivery date. Given that each individual futures contract has a limited life span, a common practice in testing academic hypotheses about futures markets involves linking the price series of individual futures contracts through time so that a longer artificial price history can be created. However, there does not appear to be a consensus regarding the procedure to construct these artificial series, especially regarding the point in time at which the current contract is rolled over to the next. For example, in an investigation of the daily effect in stock index futures markets, Junkus (1986) splices together the price series of
individual futures contracts, excluding prices in the delivery month. This avoids any nonstationarity in the return series due to increase in variance associated with decreasing time-to-maturity. When testing the weekend effect in the S&P500 futures index, Dyl and Maberly (1986) set the rollover date at the third Wednesday of the current contract month – two days prior to the actual delivery day of the S&P500 futures contracts. Herbst et al. (1989) create a ‘perpetual contract’ with the daily price computed as a weighted average of prices on outstanding contracts, where the weights represent each contract’s number of days from the present. Ma et al. (1992) provide a summary and analysis of various rollover methods. No matter which day is used as the rollover date, the ‘artificial’ time series of the cost of carry created by linking the price series of individual futures contracts through time is unlikely to be nonstationary because the cost of carry must be mean-reverting (enforced by the decreasing time-to-maturity of each individual futures contract). The mean-reverting cost of carry should therefore not affect the cointegrating relationship between spot and futures prices themselves.

**Threshold cointegration**

More recently, Balke and Fomby (1997) point out that implicit in much of the previous discussion of cointegration and its corresponding ECM is the assumption that such a tendency to move toward equilibrium (in expectation) is present every time period. Yet the presence of fixed costs of adjustment, along with other factors such as capital constraint, interest rate, and execution risk, may prevent economic agents from adjusting continuously. They propose that only when the deviation from equilibrium exceeds a critical threshold do the benefits of adjustment exceed the costs, and hence allow economic agents (in particular, arbitrageurs) to act to move the system back towards the equilibrium. They have therefore developed a threshold error correction model (TECM), which allows for nonlinear adjustment to long-run equilibrium. That is, the cointegrating relationship is inactive inside a given band, normally defined by transaction costs involved in arbitrage, and becomes active once the system goes beyond the transaction cost band. One could assume symmetric thresholds, in which short-term dynamics are assumed to be the same in the outer regimes, as well as asymmetric thresholds, which would allow different short-term dynamics in the lower and upper outer regimes. Asymmetric thresholds could be caused by the presence of
asymmetric transaction costs. For example, short-selling restrictions in the stock market may cause a short position in shares to be more expensive than a long position in shares, such that positive deviations from equilibrium (where the futures price is too high) could be expected to exhibit quicker reversion than negative deviations (involving the potential short-selling of the stock underlying the index) (McMillan and Speight, 2006).

With the advantage of allowing arbitrageurs to respond differently to different levels of pricing errors, TECM has been used by some researchers in examining the nonlinear adjustment to equilibrium in U.S. stock index-futures arbitrage. For example, Tsay (1998) applies a three-regime error correction model with asymmetric thresholds to study the pricing relationship between S&P500 index in May 1993 and its June futures contract. As anticipated, he finds that both spot and futures return series do not depend on the error correction term in the middle regime. On the other hand, for the two outer regimes, the error correction term is highly significant in the spot equation, so the past futures returns appear to be more informative in explaining the variations in both return series. Martens et al. (1998) select a five-regime threshold error correction model in examining the price relationship between the S&P500 index and its matching futures contract maturing in June and December 1993. They find that there is a clear lead of the futures market over the spot market and that the impact of the futures market on the spot market is larger when the mispricing error (defined as deviation of actual futures price from its theoretical value) is negative. They also find that the deviation from the no-arbitrage relation becomes more important for current returns the further the futures price moves from its theoretical value, and that mean reversion is weaker for mispricing that is close to zero (below the smallest positive threshold and above the smallest negative threshold, indicating the absence of arbitrage in this range).

One limitation of threshold models is that they assume that the adjustment mechanism is either entirely in one regime or the other with a sharp (or abrupt) transition at a common threshold. This implies homogeneity of traders, with identical transaction costs, who agree on the fair price of the futures contract and act simultaneously in a uniform manner (Tse, 2001; McMillan and Speight, 2006). To accommodate heterogeneous investors, Tse (2001) models the mispricing of DJIA futures as a smooth transition autoregressive (STAR) process with the speed of adjustment toward equilibrium varying directly with the degree of mispricing and demonstrates that the
observed mean reversion in mispricing changes is induced by heterogeneous arbitrageurs, instead of being simply a statistical illusion created by infrequent trading of index portfolio stocks, as has been suggested by Miller et al. (1994). Using model selection tests, McMillan and Speight (2006) find that where arbitrageurs face different transaction costs, perceive different fair prices and face different levels of capital constraint, a smooth transition model allowing for smooth transition between regimes of behavior is more appropriate. They conclude that the examination of heterogeneous trader types can provide a richer understanding of market dynamics. However, experience shows that the transition parameters of a STAR model are hard to estimate and interpretation of an estimated STAR model is often very complicated (Tsay, 2005).

3.3.3 Conditional heteroskedasticity

Analysis of causal relationships based on models that concentrate only on the first moment of the cash index and futures price series implicitly assumes constant variance in cash and futures markets. However, the volatility of financial time series is related to the rate of information flow to the market and changes in variance are commonly argued to reflect the arrival of information and the extent to which the market evaluates and assimilates new information (Clark, 1973; Ross, 1989). If information shocks are uneven over time we would expect that, over any given time period, a higher rate of information arrival implies more volatile price changes (Karpoff, 1987). In fact, financial time series usually exhibit ‘volatility clustering’, in which large (small) price changes tend to follow large (small) price changes. Volatility clustering could occur because the arrival of new information is serially correlated or because the market is not sufficiently liquid to absorb large trades, resulting in dependence in sequential price changes (Wong and Vlaar, 2003). In addition to properly modelling causality and correlation between two financial time series, volatility is important in many other areas of research, including the pricing of options, volatility estimation and forecasting, risk management, hedging performance, and so on. A correct understanding and proper measurement of volatility are therefore essential for both academic and practical purposes. Two important hypotheses regarding heteroskedasticity in stock and futures markets are analyzed below.
The relation between price volatility and futures contract maturity

The Samuelson Hypothesis, proposed by Samuelson (1965), states that as the delivery date of a futures contract approaches, the volatility of its price changes will increase. Duffie (1989) suggests that the notion underlying the Samuelson Hypothesis is that the current futures price reflects current information about the spot price at delivery time. If this information is received more quickly as the delivery date approaches, one might expect futures prices to show correspondingly higher volatility. Sutcliffe (2006) suggests that the intuition behind the Samuelson Hypothesis is that a long time to the delivery date means a lot of time for new information to affect the final delivery price, so that any single piece of information is relatively unimportant. However, when delivery is about to take place there is little time for further information to arrive, so the information that does appear is relatively important. However, Samuelson’s argument is not conclusive. Rutledge (1976) considers two of the most plausible specifications of the generation of spot prices, yet they yield conflicting results as to the variability of futures prices: one specification implies that futures price volatility increases as the delivery date approaches and provides support to the Samuelson Hypothesis; in contradiction to the Samuelson Hypothesis, the other specification implies that futures price volatility decreases as the delivery date approaches. While empirical studies of the volatility-maturity relationship for non-financial assets have found much support for the Samuelson Hypothesis, the available evidence does not provide clear support in the case of index futures (Sutcliffe, 2006).

In particular, markets are expected to be highly volatile on ‘Triple Witching Days’ (when contracts for stock index futures, stock index options and various individual stock options all expire) as traders rush to offset (or unwind) their positions before the closing bell. When the consequent cash trades happen to be predominantly on one side of the market, there may be substantial order imbalances. If the specialists handling the underlying stock cannot provide sufficient liquidity, these order imbalances can lead to sharp price movements, up or down. The result may be significant price volatility on expiration days (Edwards, 1988). Edwards (1988) also reports that for some American indices, stock price volatility is in fact greater on those days when index options and futures expire together than when either only options expire or when neither futures nor options expire.
The relation between price volatility and trading volume

Several theories predict a positive contemporaneous relation between price volatility and trading volume. The `Mixture-of-distributions hypothesis', developed by Clark (1973) and then extended by Epps and Epps (1976), Tauchen and Pitts (1983) and Harris (1986), assumes that the variance of the price change on a single transaction is conditional upon the volume of that transaction. Transaction price changes are then sampled from a mixture of distributions with volume per transaction acting as the mixing variable. Copeland (1976, 1977), Morse (1981), Jennings et al. (1981), and Jennings and Barry (1983) develop and extend 'sequential arrival of information' models in which information is disseminated sequentially to traders, which generates both trading volume and price movements, both of which increase during periods characterized by numerous information shocks. Cornell (1981) supports the view that uncertainty introduces two motives for futures trading - either a desire to transfer risk leads to hedge transaction or differential assessments of information regarding the future value of an asset causes belief trading (speculation). Therefore an increase in uncertainty should lead to an increase in both hedging and belief trading, implying a positive relation between trading volume and price variability in futures markets. Admati and Pfleiderer (1988) suggest that, in general, in order to use the activity generated by uninformed noise traders as camouflage, both discretionary liquidity traders (who buy or sell according to their activity in the underlying market, not due to information arrival) and informed traders prefer to time their dealing activity and trade when the market is active. The clustering of trades at certain (undefined) points causes more information to be released, yielding more changeable prices in periods of concentrated trading. As a result, the periods of higher trading volume also tend to be the periods of higher volatility of price changes.

A positive relation between volatility and volume has been widely documented in empirical studies. Karpoff (1987) reviews previous research on the relation between price changes (indicating volatility) and trading volume in a variety of financial markets and finds two empirical relations emerging as 'stylized facts' - volume is positively related to the magnitude of the price change in both equity and futures markets, and the correlation between volume and the price change *per se* is positive in equity markets. For the Major Market Index and its futures contracts, Chan and Chung (1993) find that
cash volume is significantly related to both lagged spot and futures price volatility, and spot volatility has a stronger impact than futures volatility on cash market trading volume. Bessembinder and Seguin (1993) find that for the eight physical and financial futures markets examined, futures price volatility is positively related to both the expected and unexpected components of volume; but that unexpected volume shocks have, on average, seven times the effect on price volatility as changes in expected volume. Abhyankar et al. (1999) find that both volume and volatility of FTSE100 futures trading exhibit a U-shaped pattern over the day. Using volume data categorized by type of trader (including market makers, clearing members trading for their own accounts, floor traders trading for other exchange members and the general public), Daigler and Wiley (1999) find that a positive volatility-volume relation is driven by the general public, a group of traders who do not have access to precise information on order flow. Clearing members and floor traders who observe order flow often decrease volatility.

3.3.4 Causality-in-variance

An important feature of financial data is that the variance changes over time. Given the studies by Clark (1973), Tauchen and Pitts (1983) and Ross (1989), it is the variance of price changes, not simple price changes, that is related to the rate of information flow to the market. Chan et al. (1991) argue that previous studies of causal relationships between futures and spot markets ignoring time-varying volatility may be subject to model misspecification and therefore offer only inconclusive evidence on causal relationships between the two markets. Time-varying volatility of price changes in the cash and futures markets provides another way in which information flows to those two markets can be measured. Causality-in-variance (or 'volatility spillover'), along with causality-in-mean, between spot and futures markets has since attracted extensive research interest.

Proper modeling of time-varying volatility is important to test for causality-in-variance. For financial data, time-varying volatility often displays a pattern of 'volatility clustering'. Another important feature of financial data is the 'leverage effect', in which a negative shock to the market is likely to cause volatility to rise by more than a positive shock of the same magnitude. The most common explanation for this effect is that a decline in a firm's stock price resulting from a negative shock raises the debt-
equity ratio of the firm. A larger debt-equity ratio increases the risk of owning the stock and results in an increase in the volatility of the stock return (Brooks, 2002). Another possible cause of asymmetric volatility is that short sales, which are normally believed to be undertaken by informed traders, signify bad news and cause markets to overreact compared to good news. Asymmetric responses to positive and negative shocks are therefore exacerbated by short selling (Henry and McKenzie, 2006). The two important characteristics of financial asset returns, ‘volatility clustering’ and the ‘leverage effect’, can be modeled within the same framework. GARCH class models specify how conditional variances change over time and have proven to be successful in capturing ‘volatility clustering’ of financial time series. The exponential GARCH (EGARCH) model proposed by Nelson (1991) and the GJR-GARCH model, also known as threshold GARCH model (or TGARCH), proposed by Glosten et al. (1993) incorporate asymmetry in return volatilities and can be used to handle leverage effects.

Univariate and multivariate GARCH models have their own pros and cons, and both have been widely used in empirical research on conditional variance. Univariate GARCH models are limited in that they assume the volatility generating processes of two time series are entirely independent and fail to capture correlation between them. Given their simplifying assumptions, univariate GARCH models have the advantage of being easy to implement and have been used by some researchers to examine volatility spillover effects between spot and futures markets. Two steps are involved in this methodology. Univariate GARCH models are first estimated for both spot and futures returns separately. Two series of conditional variance estimates for the spot and futures returns are generated from the first-step estimation. The test for causality-in-variance is then based on OLS regression of conditional variance of the spot (or futures) return on leads and lags of conditional variance of the futures (or spot) return. For example, Abhyankar (1995) uses a univariate EGARCH model to investigate hourly returns on the FTSE100 spot index and futures contract and finds no clear pattern in which one market systematically leads the other in terms of volatility. Cheung and Ng (1996) propose a two-step CCF test to examine causality-in-variance. This involves a first step of estimating univariate time series models that allow for variation in both conditional mean and variance, and a second step of computing the cross correlation function (CCF) of the squared standardized residuals, which allows inference to be made on causality-in-variance. According to the two-step CCF test, Cheung and Ng (1996) report
feedback in variances of the 15 minute S&P500 index and futures returns, while Cheung and Fung (1997) find evidence of volatility spillover between Eurodollar spot and futures markets. Since it does not involve simultaneous modeling and estimation of multiple price series, the two-step procedure is easy to implement compared with multivariate GARCH models. However, Hafner and Herwartz (2004) find that the two-step CCF procedure is characterized by a severe shortfall in terms of empirical power and that it may not be able to find the evidence in favor of volatility spillovers.

Several different multivariate GARCH models have been proposed in the literature, including the VECH, diagonal VECH, BEKK, CCC-GARCH and DCC-GARCH models. While multivariate GARCH models are able to make more efficient use of information in the variance-covariance matrix than are univariate GARCH models, and have found considerable empirical success, the estimation of large time-varying matrices of unrestricted multivariate GARCH models usually proves to be a formidable task. The exact interpretation and impact of the individual coefficients is also difficult to discern. These problems have restricted researchers to estimating multivariate GARCH models with limited scope or considerable restrictions (Sheppard, 2001). For example, Chan et al. (1991), Tse (1999b), and Bhar (2001) extend the univariate GARCH model to the multivariate setting, but impose a restriction of constant covariance (or correlation) between spot and futures returns. Using a bivariate GARCH model, Chan et al. (1991) uncover a strong intermarket dependence in the volatility of the S&P500 cash and futures returns, stronger than the intermarket dependence of price changes themselves. Based on a bivariate EGARCH model, Tse (1999b) finds a significant bidirectional information flow, with the futures market volatility-spillovers to the stock market being greater than vice versa for the DJIA index. By extending the bivariate EGARCH model to incorporate the possibility that the lagged error correction term might also have predictive power for the conditional volatility, Bhar (2001) reports that both markets display volatility spillover effects and asymmetric responses of volatility to past innovations, with cointegrating residuals having explanatory power for both conditional mean and conditional volatility. Recent papers by Tsui and Yu (1999) and Zhong et al. (2004), however, find that constant correlation can be easily rejected for price changes in both spot and futures and in some other assets. Ignoring this fact can result in serious problems for hedging and other relevant investing strategies such as diversification. The Dynamic Conditional Correlation (DCC)
GARCH model (DCC-GARCH) recently developed by Engle (2002) has the flexibility of univariate GARCH models coupled with parsimonious parametric models for the conditional correlations. The DCC-GARCH model is estimated in two steps: univariate GARCH models are estimated for each asset series in the first step; the standardized residuals resulting from the first step are then used in the second step to generate a time-varying correlation between two asset series. Wong and Vlaar (2003) report that while the two-step estimation procedure reduces the computational burden of traditional multivariate GARCH models, it may result in loss of efficiency. Being easy to implement, the DCC-GARCH model has been used in many empirical investigations of dynamic conditional correlations between financial time series (e.g. Wong and Vlaar, 2003; Lanza et al., 2006).

3.4 Summary

In summary, despite widely documented evidence of divergence from the cost of carry relationship between stock index futures price and the underlying index value, the cost of carry model as a primary approach for pricing stock index futures has never been changed so far, both theoretically and practically. On the other hand, given the real world frictions created by transaction costs in both spot and futures markets, dividend uncertainty associated with component shares, the influence of tax, arbitrage risks (e.g. tracking error risk, the risk of delayed execution) and short sale restrictions, people do realize that the futures price is allowed to fluctuate within a band around the theoretical value defined by the cost of carry model without triggering profitable arbitrage opportunities. The width of the band is acknowledged to be determined by the general transaction costs of the most favorably situated arbitrageurs in the market. Therefore, using the cost of carry model as the benchmark for evaluating stock index futures prices, the observed ‘mispricing’ is sometimes only a statistical illusion created by real-world frictions rather than by arbitrage opportunities. Then how far can futures contracts be freely ‘mispriced’ and when does this trigger profitable arbitrage opportunities? These questions are clearly interesting to both researchers and practitioners and therefore worth further investigation. Some researchers try to estimate the no-arbitrage band subjectively, using substantive information about transaction costs and other factors that may affect arbitrage (Yadav and Pope, 1990, 1994; Butterworth and Holmes, 2000). The introduction by Balke and Fomby (1997) of the concept of threshold cointegration
explores a new way to study regime-dependent price dynamics depending on the presence or absence of arbitrageurs. This alternative way of estimating the no-arbitrage band should provide a more objective and reliable reflection of the average transaction costs faced by most active arbitrageurs.

While real world frictions provide a reasonable explanation for the frequently observed 'mispricing' of stock index futures contracts, they can also explain why there may be lead-lag relationships between spot and futures prices. Since transaction costs and other market imperfections differ between the two markets, they will affect the 'price discovery' contribution of each market at new arrival of information, causing one market to react faster to the same information than the other. While this used to be regarded as evidence against the cost of carry model, it has more recently been realized that if spot and futures prices are cointegrated in the long run, there must exist some causal ordering between spot and futures price changes (Engle and Granger, 1987) and that this is implied by the cost of carry model. But similar to the concern on 'mispricing', a big concern here is the length of time that a lead-lag relationship can be sustained. Even in the presence of market imperfections, there should be no systematic lagged responses of spot or futures prices large enough or long enough to exploit profitably after taking transactions costs into account.

3.5 Conclusions

This review suggests that most studies of futures pricing issues have been carried out for the US market and the number of papers addressing the UK market is very small, despite its important position in the financial world. Testing for futures pricing efficiency suffers from the joint hypothesis problem associated with any empirical test of the efficient market hypothesis (EMH) – that is, market efficiency per se is not testable because the question whether the price of financial assets has fully reflected all relevant information always depends on the model of asset pricing that the researcher is using. Despite the joint hypothesis problem, tests of futures pricing efficiency using the cost of carry model improve our understanding of the behavior of returns across time and markets. They also help to improve existing models for pricing stock index futures and are valuable to financial-markets professionals.
Regarding tests of causal relationships between spot and futures markets, the literature suggests that the test strategies based on very simple modeling (such as a VAR) using no other factors than past returns would miss many other features of the market and that any inference that follows could therefore be unreliable. Given the complexities of stock market trading that are caused by real world market frictions, such as transaction costs, price impact, short sale constraints, changeable volatilities and other institutional limitations, a more refined analysis using multiple approaches could yield more complex and accurate pictures of the very same events. For example, time-varying volatility can be accounted for by GARCH models while the regime-dependent price dynamics induced by transaction costs and arbitrage can be described by threshold error correction models. Indeed, applying sophisticated models has become a trend in empirical studies of financial issues over the past decade. This is in parallel with the promotion and popularity in the investment industry of analysis involving numerical and quantitative techniques, including mathematical computational and statistical models.

Based on the review of the literature on stock index futures pricing, the main objective of this research is to observe arbitrage relationships between the FTSE100 spot and futures markets over the past two decades. Meanwhile, it is a joint test of the validity of the cost of carry model for pricing the FTSE100 futures contracts. Different modeling techniques for FTSE100 spot and futures returns should provide explanations of price dynamics from different perspectives and therefore provide alternative insights into the same issue. Furthermore, examining pricing efficiency for FTSE100 futures using different empirical approaches should also yield results that allow more robust inference to be drawn.
Chapter 4 Unified Tests of Causality and Cost of Carry – A VECM Analysis of FTSE100 Spot and Futures Prices

4.1 Introduction

The cost of carry model is so far the most widely used model for pricing stock index futures. As analysed in the literature, a traditional inference drawn from the cost of carry model is that spot and futures price changes should be perfectly contemporaneously correlated and there should not be any lead-lag (or causal) relationships between them (Stoll and Whaley, 1990). However, others argue that since transaction costs and market frictions in the spot market are different from that in the futures market, people may prefer to trade in one market relative to the other (Abhyankar, 1995; Fleming et al., 1996). There have been considerable empirical investigations of price dynamics between spot and futures markets. The empirical findings have been mixed, but in most cases it is found that the causality from futures to spot is stronger than the reverse (e.g. Chan, 1992; Alphonse, 2000). The most widely cited explanation for the ‘price discovery’ function of the futures market is lower transaction costs of trading in futures market than in spot market. However, Yadav and Pope (1990, 1994) argue that studies of causal relationships between stock index and futures price using transaction price based indices suffer from infrequent trading problems. Indeed, Shyy et al. (1996) find that CAC-40 futures lead the cash price when transaction price data are used, while cash leads futures when bid/ask quote-midpoint data are used, providing support to the argument of Yadav and Pope (1990, 1994). Given that UK indices are based on quote prices of component shares that should be less subject to infrequent trading problems, it provides a more reasonable context to test causal relationships between stock index and futures price. Furthermore, despite the fact that the London market is an important financial center in the world, there have been far less empirical studies of the UK futures market compared to the US futures market. The above considerations formed the initial motivations for this study.
The 'cost of carry' relationship between stock index futures and the underlying index implies that they are cointegrated in the long run and that the dynamics between them should be studied within the framework (now routine in the literature) of a VECM. The economic intuition is that any deviation from the equilibrium defined by the cost of carry model in a past period will have an impact on price dynamics in one or both markets in the current period. This is known as the 'error correction' effect. Drawing heavily on the work of Green and Joujon (2000), this chapter reports tests of lead-lag relationships between spot and futures prices and the cost of carry model within a single unified VECM, using daily changes of FTSE100 index and futures prices from 28/10/1986 to 30/12/2005. It contributes to the literature in at least three ways.

First, the UK market was studied, for a relatively long data period (nearly 20 years). The quote-based indices in the UK market are likely to reflect more current information than transaction-based indices (such as the US indices) because the latter are potentially subject to infrequent trading problems (Yadav and Pope, 1990, 1994), so the UK market provides an arguably more appropriate context in which to test lead-lag relationships between stock index and futures prices. The data were analysed in two sub-periods. In the first sub-period a dealership trading system was employed by the LSE and an ‘open outcry’ system was employed by the LIFFE, while in the second sub-period an order-driven trading system was employed by the LSE and an electronic platform was employed by the LIFFE. This allowed an examination of the impact of changes in market microstructure on market dynamics. Second, the restrictions implied by the cost of carry theory were carefully derived and imposed on the finally selected VECM, allowing both causality and the cost of carry theory to be tested within the same framework. An explicit specification of the restrictions implied by the cost of carry on the VECM framework suggests that some causal orderings between spot and futures prices are not necessarily inconsistent with the cost of carry model, providing a solution to the general puzzle of the widely documented lead-lag relationships between spot and futures prices, which were formerly interpreted as evidence of market inefficiency inconsistent with cost of carry. Third, in addition to lagged own- and cross-market returns and a lagged error-correction term in a traditional standard VECM for spot and futures returns, the effect on returns of the passage of time (normally ignored in empirical studies) was explicitly accounted for. There are 'artificial' futures price jumps at each contract roll-over when prices from consecutive futures contracts are
spliced into a single time series. The impact of this was explicitly modeled by using a dummy variable approach. While this has been examined by Green and Joujon (2000) for French data, there has been no such investigation for FTSE100 spot and futures prices.

The rest of this chapter is organized as follows: Section 4.2 introduces the data used in this research and provides a preliminary analysis of the data. In section 4.3, a VECM for FTSE100 spot and futures returns is developed and the restrictions implied by the cost of carry model on the VECM are derived, allowing the causality relationships between spot and futures prices and the cost of carry model to be tested within the same VECM. Section 4.4 reports empirical results and section 4.5 draws conclusions.

4.2 Data and Preliminary Analysis

4.2.1 Data and sample descriptions

Daily price dynamics between FTSE100 spot and futures markets over the time period 28/10/1986 — 30/12/2005 were examined in this study. Daily prices were used for two main reasons. First, since futures contracts are resettled daily, daily settlement (closing) prices\(^{14}\) are the prices that determine daily net gains or losses, margin calls and invoice prices for deliveries and hence the cash flows of traders. Therefore, whether the futures contracts are fairly priced or mispriced at the daily settlements seems to be especially important and relevant. Green and Joujon (2000) also argue that a proper test of cost of carry must, at some stage, utilize daily settlement prices. Second, Shiller and Perron (1985) show that the power of many tests commonly used in financial market research does not increase as the number of observations increases, unless this increase is accompanied by an increase in the span of the data. Therefore, we use nearly 20 years of daily settlement prices rather than higher frequency intraday data in this research.

It is commonly believed that the linkage between futures and the underlying spot markets is maintained by arbitrage, which is in turn affected to a great extent by the costs of trading in the two markets. It is therefore important to perform an empirical analysis of spot-futures price dynamics over a time period with a relatively stable

---

\(^{14}\) The official exchange settlement price is usually different from the actual closing price. The settlement price is determined by calculating the weighted average of prices during the closing period, which typically ranges from five to ten minutes and quite often is changed some time after the actual trading ceases. The exact reason is somewhat unclear. However, it seems to be related to the fact that the floor traders and floor brokers do some type of settlement between themselves after the actual close of public trading.
structure of transactions costs. This was used as the principle for selecting the sample period and two sub-samples in this research. Although transactions costs in London have evolved over time, they have been most affected by two discrete microstructure changes at the LSE and one microstructure change at LIFFE. UK financial markets were deregulated in a ‘Big Bang’ on 27/10/1986, fixed commission charges were abolished and trading moved from the floor of the LSE to a screen-based system. 28/10/1986 was therefore set as the start date of our sample period. Another important microstructure transformation occurred at the LSE on 20/10/1997, when the quote-driven dealership system (SEAQ) was largely replaced by an order-driven system (SETS) as the official system for trading the most liquid FTSE100 securities\(^\text{15}\). SETS was launched by the LSE to bring greater speed and efficiency to the market. Under SETS, buy and sell orders placed by each market participant are automatically reported and matched by the electronic order-book. Transaction costs have declined under SETS, helping to narrow the spread between buy and sell prices. In addition, automated matching of buy and sell orders makes the order-driven market more transparent and efficient for participants. 17/10/1997 was therefore tentatively set as the end of the first sub-sample, during which a dealership system was employed by the LSE. Formal tests for sub-sample difference were conducted later and it was suggested that there was indeed a structural break on 20/10/1997 (see Section 4.4.1 for analysis; Table 4.7 for the results of Chow tests; Figure 4.10 and Figure 4.11 for the results of CUSUM tests).

An important microstructure transformation also occurred at LIFFE on 30/11/1998, when the old ‘open outcry’ was replaced by LIFFE CONNECT, described by LIFFE as ‘then the most sophisticated electronic derivatives trading platform in the world’. FTSE100 futures traders have experienced from the new electronic platform substantial benefits such as lower operating costs, remote access to the system and higher transparency. It would therefore be preferable to set 20/10/1997–27/11/1998 as the second sub-sample (an order-driven trading system at the LSE and ‘open outcry’ at LIFFE) and 30/11/1998–30/12/2005 as the third sub-sample (an order-driven system at the LSE and an electronic platform at LIFFE). However, the time period of 20/10/1997–27/11/1998 is too short and there are insufficient observations for empirical work. As will be seen later in Chapter 5 and Chapter 6 respectively, over the sub-

---
\(^{15}\) SEAQ has continued in parallel with SETS, with dealers voluntarily quoting prices over the telephone and trading off-exchange. Therefore traders can choose between the order-driven system and the ‘off-exchange’ dealership system.
sample 30/11/1998–30/12/2005, the estimation of the DCC-TGARCH model failed to achieve convergence and the estimation of the threshold VECM resulted in a very small number of observations (only 32) in an ‘upper regime’ of transactions costs, implying a potential bias in inference caused by tiny sample sizes. To be practicable, 28/10/1986–17/10/1997 was chosen as the first sub-sample (all FTSE100 securities traded through a dealership system and FTSE100 futures traded through ‘open outcry’) and 27/10/1997–30/12/2005 as the second sub-sample (all FTSE100 securities traded through an order-driven system and FTSE100 futures mainly traded on an electronic platform). There is a short time period (until 30/11/1998) at the beginning of the second sub-sample when FTSE100 futures were traded through ‘open outcry’, but the impact was expected to be small (see below for details). Small changes in transactions costs at other times may have affected arbitrage and hence the market dynamics, but this impact should be minor in comparison with the effects of the Big Bang and the moves to SETS and the LIFFE CONNECT.

The FTSE100 is a market-value weighted index of the largest 100 UK companies listed on the LSE and is computed every 15 seconds throughout the day from the mid-point of the inside (i.e. narrowest) spread for each constituent stock. As a quote-based index it may reflect more up-to-date information than transaction-based indices (such as the US indices) since it represents market makers’ (or after 20/10/1997, market participants’) estimates of the current tradable value of component shares (Yadav and Pope, 1994). If these estimates change the index will be updated, even if no transactions occur in constituent shares. As it is composed of the largest 100 shares which are also the most heavily traded shares at the LSE, there should be little or no nonsynchronous trading effect in the index. On the other hand, as the quotes are not necessarily real transaction prices, some apparent arbitrage opportunities between the FTSE100 index and its futures price may simply be statistical illusions.

FTSE100 index futures contracts have four quarterly expirations: March, June, September and December. For any given trading day, the near contract is usually the

---

16 In fact the second sub-sample starts from 27/10/1997 (omitting 5 observations) to allow for the necessary lags in the error-correction model.
17 The inside spread is the difference between the highest bid and the lowest ask prices in the market. Prior to 20 October 1997, the bid and ask prices were quoted by competitive market makers registered for different stocks. When the LSE changed the market microstructure system for trading FTSE100 securities from a dealership system to the electronic order-driven trading system (20 October 1997) the bid prices and ask prices for each stock were actually posted by market participants, who fulfilled the market-making function by placing orders to buy or sell certain quantity on the screen through out the trading period.
most heavily traded. As the last trading day of a contract month approaches, investors begin to roll over their positions from the near contract to the next near contract. Given that each individual futures contract has a limited life span and most futures contracts are heavily traded in the last few months before expiration, a common practice in testing academic hypotheses about futures markets involves linking the price series of individual futures contracts through time so that a longer artificial price history can be created. Many different rollover methods have been utilized and there does not appear to be a consensus regarding the procedure for constructing these artificial series (Ma et al., 1992). First, a decision must be made regarding the selection of a point in time to ‘roll over’ – that is, when to switch from the maturing contract to the next contract. As argued by Green and Joujon (2000), because the cost of carry theory implies that the mispricing generates arbitrage at any time in the life of a contract, it would seem arbitrary to use data that switches to a new contract before maturity. We therefore spliced data from successive futures contracts into a single time series in which observations shift to the next near contract on each expiration day. The final three months’ futures prices before expiration of each futures contract therefore form a single time series of futures price in our study. As pointed out by Ma et al. (1992), heterogeneity of consecutive contracts, as well as the unusual market activities often observed near maturity (e.g. futures markets usually show increased price variance as time to delivery decreases, Samuelson, 1965), can generate significant biases in the time-series properties of the artificial price series. Another problem arising from rollover is that there is an artificial futures price ‘jump’ at each rollover date. A decision needs to be made regarding whether or not a series of prices is to be adjusted to remove the artificial jumps, and how to do it. On the one hand, the jumps may generate seemingly excessive volatility and extremely large price changes, distorting the parameter estimates of the true underlying distributions and bringing into question the validity of the test statistics. A typical method of adjustment is to use a moving average procedure around each rollover to create a series of artificially-smoothed prices. On the other hand, traders argue that no ex post adjustment should be made, since only real transaction prices can be used in practice. This commitment should be followed even when it is known that an explicit bias will be present (Ma et al., 1992). Indeed, since an artificially-smoothed price series does not represent actual transaction prices, it may disguise true arbitrage opportunities. Furthermore, the artificially adjusted price series could effectively import serial dependence into the return process, which could
introduce another bias and distort the conclusions of empirical studies (Working, 1960). In order to account explicitly for artificial price jumps while maintaining real transaction prices, in the manner of Green and Joujon (2000), a dummy variable was defined to model the rollover jumps in this study (see below).

Daily closing prices for the FTSE100 index and daily settlement prices for FTSE100 futures were used in this study. Since the daily closing time on LIFFE for FTSE100 futures is not necessarily the same as the daily closing time on LSE for trading FTSE100 component shares, the study was carried out under the assumption that the daily futures settlement price is synchronous with the daily closing index value. Possible asynchronicity between FTSE100 futures price and the prices of the underlying shares might produce noise in fair value estimates, though it should not lead to systematic differences between the normative index futures price and the actual index futures price, as argued by Yadav and Pope (1990). In order to construct the net cost of carry series \(\text{coc}_t = (r - d)(T - t)\), a measure of risk-free interest rate and dividend forecasts for the 100 stocks in the FTSE100 index are required. Strictly, the interest rate for day \(t\) is for a loan with maturity \((T - t)\). This requires the estimation of the complete term structure of interest rates to allow matching of the maturity of each loan with the time-dependent maturity of the futures contract. In practice, most researchers have used the three or one month Treasury bill yield. However, it can be argued that this could lead to bias due to overlapping of interest rate maturities with the futures expiry date. To overcome this problem, LDMON (London Discount Market Overnight Rate, \(r\)) was used in this study to proxy for the interest rate for a loan with maturity \((T - t)\). The realized ex post dividend yield for the FTSE100 index, which is the most accurate estimate of the forecasted dividend yield of the FTSE100 index, was used as a proxy for the forecasted dividend yield. Both interest rate and dividend yield were

For example, as of 2007, trading of the FTSE100 component shares on the LSE lasts from 08:00-16:29 (when the closing auction starts), and closing values are taken at 16:35 (though the closing value of the index itself is timed at 16:36). On the other hand, LIFFE CONNECT trading hours for FTSE 100 futures last from 08:00-17:30, and the daily settlement time is at 16:30 (from NYSE Euronext website). However, trading hours for FTSE100 index futures on LIFFE have changed occasionally. Trading was from 09:05 to 16:05 for the four years from 28 April 1986 until 23 March 1990, from 08:35 to 16:10 for the eight years from 26 March 1990 to 17 July 1998 and from 08:35 to 16:30 for the few months from 20 July 1998 until December 1998 (see, e.g. Areal and Taylor, 2002).

Given that LDMON is the interest rate for overnight loans, it could be lower or higher than the relevant interest rate for a loan maturing at the futures delivery date, depending on the shape of the yield curve. Therefore, the use of LDMON could bias the estimate of the fair futures price. Theoretically, it is possible to estimate for every day the interest rate for a loan maturing at \(T\) by using the continuous compound interest formula, compounded day by day, from \(t+1\) to \(T\), using each day's overnight rate. However, since the average maturity of futures contracts used in this study is only 1.5 months, the bias introduced by using LDMON is expected to be small and insignificant.
transformed from annual to daily rate. Thus they can be multiplied by the time to maturity (in days) of the relevant futures contract to obtain the net cost of carry.

All the spot data, including the FTSE100 index, the FTSE100 dividend yield and the LDMON, are from DataStream. The FTSE100 futures prices are from Euronext.liffe. The sample includes daily observations from 28/10/1986 to 30/12/2005, excluding non-trading days\(^{20}\) and weekends, giving a total of 4832 observations for the final dataset. The futures and spot prices are defined by \(f_t = \ln F_t\), \(s_t = \ln S_t\), and the basis by \(b_t = f_t - s_t\)\(^{21}\). All daily price changes (or returns) are estimated by the logarithm of the relative prices, i.e. \(\Delta_k f_t = \ln F_t / F_{t-k}\), and \(\Delta_k s_t = \ln S_t / S_{t-k}\). \(k_t\) is the time interval, or the number of days, between contiguous observations. In general, \(k_t = 1\); but \(k_t > 1\) for those time intervals covering weekends or holidays. The observation at one lag is described by \(t - k_t, t - k\) or \(t - 1\) according to context. The evolution of the FTSE100 spot and futures prices are graphed in Figure 4.1 and Figure 4.2 respectively. These show clearly the dot-com bubble around 1995–2000. Following the bursting of the bubble at the beginning of 2000, the market experienced a relatively lengthy recession followed by a recovery from early 2003. The net cost of carry for the two sub-sample periods is shown in Figure 4.3 and Figure 4.4, and the basis is given in Figure 4.5 and Figure 4.6. These show clearly that there is a downward trend in the cost of carry and the basis over the life of each individual futures contract. Accordingly, for the whole dataset the cost of carry and the basis frequently revert back to zero, implying that both should be stationary time series.

\(^{20}\) There are at least 8 days removed each year: New Year's Day, Good Friday (Friday before Easter Sunday), Easter Monday (Monday after Easter Sunday), Early May Bank Holiday (1st Monday in May), Spring Bank Holiday (Last Monday in May), Summer Bank Holiday (Last Monday in August), Christmas Day and Boxing Day.

\(^{21}\) The standard basis is usually defined as \(b_t = F_t - S_t\). However, following many other researchers, in this research the basis is defined as \(b_t = \ln F_t - \ln S_t\).
Figure 4.3 Time plot of the cost of carry for FTSE100 index and futures prices, 28/10/1986 – 17/10/1997

FTSE100 cost of carry, 28/10/1986 – 17/10/1997
Figure 4.5 Time plot of the basis for FTSE100 futures and index prices, 28/10/1986 – 17/10/1997
4.2.2 Normality test

Table 4.1 reports some descriptive statistics for FTSE100 spot and futures returns. The mean of both return series is above zero, with the mean in the first sub-sample slightly higher than that in the second sub-sample. For both sub-samples, the standard deviation of the futures return series is higher than that for the spot return series, consistent with the general finding in the literature that the futures market is usually more volatile than the spot market (MacKinlay and Ramaswamy, 1988). It can also be seen that the skewness value is negative for both futures and spot returns, indicating that both returns have a longer left tail than that of the normal distribution. The negative skewness value is caused by more negative returns than allowed by the normal distribution. The skewness value is more negative for the first sample period than for the second sample period, possibly caused by excessively negative returns during the 1987 stock market crash period. However, the skewness values are not too far away from zero, consistent with limited liability in the spot market and the prevention of accumulated losses by daily re-settlement in the futures market. The kurtosis values for both return series are far greater than 3 (the value for the normal distribution), especially for the first sub-sample, indicating that futures and spot returns both have leptokurtic distributions, exhibiting excess peaks at the mean and thicker tails than the normal distribution. A formal test for normality, the Jarque-Bera test, decisively rejects the null hypothesis of normality for both spot and futures returns at the 1% level for both sub-samples. The non-normal distribution can be explained, at least partially, by ‘volatility clustering’, which will be dealt with in Chapter 5.
Table 4.1 Summary statistics for FTSE100 spot and futures returns

<table>
<thead>
<tr>
<th>FTSE100 Returns</th>
<th>28/10/1986 - 17/10/1997</th>
<th>27/10/1997 - 30/12/2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spot return</td>
<td>Futures return</td>
</tr>
<tr>
<td>Mean</td>
<td>0.000433</td>
<td>0.000435</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>0.009262</td>
<td>0.010864</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.716608</td>
<td>-1.612863</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>29.6032</td>
<td>29.19632</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>83224.14</td>
<td>80579.5</td>
</tr>
<tr>
<td>Prob(Jarque-Bera)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Spot return</td>
<td>Futures return</td>
</tr>
<tr>
<td>Mean</td>
<td>0.000006</td>
<td>0.000057</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>0.012102</td>
<td>0.01258</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.147975</td>
<td>-0.068184</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5.28662</td>
<td>5.467018</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>454.3152</td>
<td>521.7037</td>
</tr>
<tr>
<td>Prob(Jarque-Bera)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

4.2.3 Unit root test

Inclusion of non-stationary variables in a standard regression will cause 'spurious regression' problems and invalidate the standard assumptions for asymptotic analysis (Brooks, 2002). Non-stationarity therefore has to be removed before standard regression techniques can be applied, so testing for unit roots in time series has become a standard practice in economic research. In order to generate robust conclusions, Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and Zivot-Andrews (Zivot) unit root tests were used to test for stationarity. The ADF test is based on the t-ratio for the coefficient $\psi$ in the estimated regression

$$\Delta y_t = c + \psi y_{t-1} + \sum_{i=1}^{p} \omega_i \Delta y_{t-i} + e_t$$  \hspace{1cm} (4.1)$$

where $c$ is a constant term in the testing procedure. The null hypothesis is a unit root ($\psi = 0$) with a drift in the time series $y_t$. The $p$ lags of the dependent variable are included to model any autocorrelation in the dependent variable and hence ensure that the error term $e_t$ is not autocorrelated – a necessary condition for the test to be valid that is not dealt with in the original DF test (Brooks, 2002). The optimal lag order $p$ can be determined by the Schwarz Information Criterion (SIC). The null hypothesis of a
unit root is rejected in favor of the stationary alternative in each case where the $t$-statistic for $\psi$ is more negative than the critical value.

The PP test is similar to the ADF test except that it estimates the non-augmented DF test equation (without lags of the dependent variable in equation (4.1)) and modifies the $t$-ratio of the coefficient so that any serial correlation in the error term does not affect the asymptotic distribution of the test statistic (the asymptotic distribution of the PP modified $t$-ratio is the same as that of the ADF statistic). Compared to the DF test, 'nuisance' serial correlation aside from that generated by the hypothesized unit root has been taken into account by the ADF and the PP unit root tests. Nonetheless, both ADF and PP unit root tests tend to exhibit rather poor behaviour in the presence of certain types of serial correlation (Schwert, 1989).

Unit root tests are biased in favour of the null hypothesis of a unit root if there are structural breaks in the series, which is highly possible if a long sample period is studied. Indeed, the time series plots of the FTSE100 index and futures price over 28/10/1986 – 30/12/2005 (see Figure 4.1 and Figure 4.2) suggest that there might be a discontinuity in the mean. To account for this and check the robustness of the ADF and PP test results, a unit root test developed by Zivot and Andrews (1992) was used that allows for an endogenous one-time (or single) break in the intercept. The null hypothesis of a unit root can be described as $y_t = \mu + y_{t-1} + \epsilon_t$. The alternative hypothesis is that $y_t$ is a trend-stationary process with a one-time break occurring at an unknown point in time, which is to be estimated. The regression equation to test for a unit root is

$$y_t = \mu + \hat{\theta}DU_t(\hat{\lambda}) + \hat{\beta}t + \hat{\alpha}y_{t-1} + \sum_{j=1}^{k} \hat{\delta}_j \Delta y_{t-j} + \epsilon_t \quad (4.2)$$

where $DU_t(\lambda) = 1$ if $t > T\lambda$, and 0 otherwise ($T$ is the number of observations); $\hat{\lambda}$ is the estimated value of the break fraction (if the estimated breakpoint is denoted as $B$, then $\hat{\lambda} = TB / T$). The idea of estimating $\hat{\lambda}$ is to choose the breakpoint that gives the least favorable result for the null hypothesis ($\alpha = 1$); that is, $\lambda$ is chosen to minimize the $t$ statistic for testing $\alpha = 1$. In each case, the null hypothesis of a unit root is rejected
in favour of the alternative of stationarity with structural change at some unknown point if the standard t-statistic for \( \alpha = 1 \) is more negative than the critical value.

The results of unit root test for the series of interest are reported in Table 4.2. It can be seen that, with just 2 exceptions, all these tests indicate that the (logarithm of) futures and spot prices \( (f_t \text{ and } s_t) \), the interest rate \( (r) \) and the dividend yield \( (d) \) are I(1) (the PP test suggests the interest rate to be 1(0) at the 1% level of significance and the Zivot-Andrews test suggests the interest rate to be 1(0) at 5%) and that the daily futures and spot returns \( (\Delta_k f_t \text{ and } \Delta_k s_t) \), the basis \( (b_t) \), and the cost of carry \( ((r - d)(T - t)) \) are I(0).

**Table 4.2 Unit Root Tests, 28/10/1986 – 30/12/2005**

<table>
<thead>
<tr>
<th>Variable</th>
<th>ADF t-stat.</th>
<th>PP adjusted t-stat</th>
<th>Zivot-Andrews minimum t-stat.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_t )</td>
<td>-1.67</td>
<td>-1.66</td>
<td>-4.28</td>
</tr>
<tr>
<td>( s_t )</td>
<td>-1.64</td>
<td>-1.64</td>
<td>-4.25</td>
</tr>
<tr>
<td>( b_t )</td>
<td>-12.18 *</td>
<td>-33.11 *</td>
<td>-12.39 *</td>
</tr>
<tr>
<td>( \Delta_k f_t )</td>
<td>-51.50 *</td>
<td>-69.36 *</td>
<td>-22.94 *</td>
</tr>
<tr>
<td>( \Delta_k s_t )</td>
<td>-67.61 *</td>
<td>-67.59 *</td>
<td>-22.30 *</td>
</tr>
<tr>
<td>( r_t )</td>
<td>-1.28</td>
<td>-4.90 *</td>
<td>-5.06 b</td>
</tr>
<tr>
<td>( d_t )</td>
<td>-1.54</td>
<td>-1.62</td>
<td>-3.24</td>
</tr>
<tr>
<td>( (r - d)(T - t) )</td>
<td>-11.23 *</td>
<td>-12.48 a</td>
<td>-13.18 a</td>
</tr>
</tbody>
</table>

*Notes:*

The test critical values for ADF and PP unit root tests are -3.43 (1%) and -2.86 (5%). The critical values for the Zivot-Andrews unit root test are -5.34 (1%) and -4.80 (5%).

* Rejection of the null hypothesis of a unit root in the relevant time series at 1% significance.

b Rejection of the null hypothesis of a unit root in the concerned time series at 5% significance.
4.2.4 Autocorrelation Test

Table 4.3 reports autocorrelation and partial autocorrelation functions (ACF and PACF) up to 5 lags for spot returns ($\Delta_t s_t$) and futures returns ($\Delta_t f_t$) over the two sub-samples. It can be seen that during 28/10/1986–17/10/1997, both spot and futures return series display evidence of significant autocorrelation up to 4 lags. During 27/10/1997–30/12/2005, both spot and futures return series display evidence of significant (partial) autocorrelation up to 5 lags. The most commonly cited explanations for such linear dependencies in stock returns are non-synchronous trading, partial adjustments, over- or under-reactions, the presence of time-varying risk premia, and the bid/ask bounce effect.

Positive autocorrelation may indicate the presence of a stale price effect. Although the FTSE100 index is a quote-based index and should reflect more current information than a transaction-based index, observed quotation prices used in the index may not be fully up to date or perfectly ‘observed’, due for example to quote adjustment costs. Lags in quote adjustment can reflect trading frictions, since quotes may be adjusted only when trades take place (Theobald and Yallup, 1998). Consequently, differences in price adjustment delays between index stocks may generate positive serial correlation in index returns (Yadav and Pope, 1994).

In an information processing context, the intrinsic value of a stock ‘fully reflects’ information, while the observed price of the stock may only ‘partially reflect’ the full information set. The partial adjustment factors of the price of a stock could be equal to, greater or less than one, corresponding to full, over- or under-adjustments towards intrinsic values (Theobald and Yallup, 1998, 2001). Positive autocorrelations may indicate under-adjustments (or under-reactions) of the price to certain new information and post-event return continuation; negative autocorrelations may indicate over-adjustments (or over-reactions) to certain new information and post-event return reversals (Fama, 1998; Theobald and Yallup, 2004). Autocorrelation in stock and its futures returns may also reflect a time-varying risk premium contained in the context of a rational equilibrium asset pricing model (Fama and French, 1987; Fama, 1991)\(^2\).

\(^2\)Assuming the existence of a risk premium, the GARCH-in-mean model suggests that if volatility is an important factor determining stock returns, volatility clustering may result in serial correlation in stock returns. Thus risk premia may at least partly account for serial correlation in observed stock returns.
Negative autocorrelations in futures returns could also be due to the bid/ask bounce effect. That is, transaction prices bounce randomly between bid and ask prices, producing significant negative serial correlations in the return series (Roll, 1984). However, negative serial correlations in index returns should not be induced by the bid/ask bounce effect, because the FTSE100 index is computed every 15 seconds from the mid-point of the inside (i.e. narrowest) spread for each constituent stock and therefore does not suffer from this problem.

Table 4.3 ACF and PACF of FTSE100 spot and futures returns

<table>
<thead>
<tr>
<th>Lag</th>
<th>Spot Returns ACF</th>
<th>Spot Returns PACF</th>
<th>Futures Returns ACF</th>
<th>Futures Returns PACF</th>
<th>Spot Returns ACF</th>
<th>Spot Returns PACF</th>
<th>Futures Returns ACF</th>
<th>Futures Returns PACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.064</td>
<td>0.064</td>
<td>0.028</td>
<td>0.028</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.019</td>
<td>-0.019</td>
</tr>
<tr>
<td>2</td>
<td>0.007</td>
<td>0.003</td>
<td>-0.041</td>
<td>-0.042</td>
<td>-0.061</td>
<td>-0.061</td>
<td>-0.060</td>
<td>-0.060</td>
</tr>
<tr>
<td>3</td>
<td>0.016</td>
<td>0.015</td>
<td>0.030</td>
<td>0.033</td>
<td>-0.099</td>
<td>-0.099</td>
<td>-0.110</td>
<td>-0.113</td>
</tr>
<tr>
<td>4</td>
<td>0.065</td>
<td>0.063</td>
<td>0.043</td>
<td>0.039</td>
<td>0.036</td>
<td>0.031</td>
<td>0.045</td>
<td>0.037</td>
</tr>
<tr>
<td>5</td>
<td>0.009</td>
<td>0.001</td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.035</td>
<td>-0.047</td>
<td>-0.042</td>
<td>-0.055</td>
</tr>
</tbody>
</table>

Notes:
There are 2776 observations in the sample period 28/10/1986 – 17/10/1997 and 2051 observations in the sample period 27/10/1997 – 30/12/2005. Therefore, the ACF and PACF are significant at the 1% level if it is outside the range [-0.0486, +0.0486], significant at 5% if it is outside the range [-0.0372, +0.0372] and significant at 10% if it is outside the range [-0.0311, +0.0311] during 28/10/1986 – 17/10/1997. The ACF and PACF are significant at 1% if it is outside the range [-0.0565, +0.0565], significant at 5% level if it is outside the range [-0.0433, +0.0433] and significant at 10% if it is outside the range [-0.0362, +0.0362] during 27/10/1997 – 30/12/2005.

*a* significant at the 1% level; *b* significant at 5% level; *c* significant at 10% level.
4.3 Methodology

4.3.1 Cointegration Test

If two or more I(1) series are cointegrated, there exists a linear combination of these series that is I(0). Brenner and Kroner (1995) argue that according to the cost of carry model, if interest rates have a stochastic trend, then spot and forward (futures) prices will not be cointegrated by themselves, and the differential (interest rate less dividend yield) should be included in the cointegrating vector. But their proposition assumes that the time to expiration of the contract, \((T - t)\), is fixed, while the time of expiration, \(T\), is changing. In fact, for futures, the date of delivery, \(T\), is fixed for each contract, but the time to delivery declines as time passes and the contract approaches the delivery date. Here it is argued that no matter what time series properties the interest rate and dividend yield possess, the time series of the cost of carry, \((r - d)(T - t)\), must be stationary if data from successive futures contracts are spliced into a single time series, with rollover at the expiry of each futures contract. The reason is that the time series property of the decreasing time to expiration, \((T - t)\), of each individual futures contract dominates the time series properties of the interest rate and the dividend yield, which are relatively stable over short time period. The cost of carry is therefore forced to revert back to zero over the life of each futures contract (see Figure 4.3 and Figure 4.4). Thus, given the way it is formed in this study, the whole time series of the cost of carry, \((r - d)(T - t)\), will be stationary, as the unit root tests indeed suggest (see Table 4.2). The cost of carry term would therefore not affect the cointegrating relationship between futures and spot prices. Thus the cointegrating regression for I(1) futures and spot prices can be specified as

\[
 f_t = a_0 + a_1 s_t + \varepsilon_t
\]

(4.3)

The constant term \(a_0\) in equation (4.3) can be interpreted as the average cost of carry over the life of a futures contract. The cost of carry relationship requires that \(a_1 = 1\) while the residual \(\varepsilon_t\) in the cointegrating regression should be a zero-mean I(0) stochastic process. Preliminary analysis of the data shows that the basis \((b_t = f_t - s_t)\) is I(0), implying that \(f_t\) and \(s_t\) are cointegrated with cointegrating vector \([1, -1]\). Using
Johansen cointegration test\textsuperscript{23}, the maximum eigenvalue and trace statistics suggest a cointegrating relationship between futures and spot prices with normalized cointegrating vector [1, -0.9957] for the period 28/10/1986–17/10/1997 and [1, -1.0115] for the period 27/10/1997–30/12/2005 (see Table 4.4). Clearly both vectors are close to [1, -1].

\textsuperscript{23} For purposes of the Johansen tests, the lag lengths for the first-differenced endogenous variables were determined on the basis of the VAR model (equations (4.8a) and (4.8b)), the optimal lag length being chosen by the usual information criteria (AIC, SIC, FPE, HQ, LR) and residual diagnostics. The chosen lag length was then used as prior information in the Johansen tests.
### Table 4.4 Cointegration Tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trace statistic</td>
<td>5% Critical Value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>113.47*</td>
<td>15.49</td>
</tr>
<tr>
<td>At most 1</td>
<td>0.08</td>
<td>3.84</td>
</tr>
</tbody>
</table>

*Note: *indicates rejection of the null hypothesis at the 5% significance level.
4.3.2 An Error Correction Model

Cointegration between futures and spot prices has significant implications for modeling the dynamics of individual series since deviations from the long-run equilibrium relationship may affect subsequent price movements in either market. The Johansen tests suggest that the cointegrating vector for futures and spot prices is close to \([1,-1]\), implying that the cointegrating vector simply defines the basis. Indeed, since the basis is a no-arbitrage link between the two markets, it determines whether arbitrage opportunities are available. If the basis is too large or too small, as a result of serious divergence from the long-run equilibrium, arbitrage transactions should occur immediately to drive the price relation between the two markets back to equilibrium. It follows that an error-correction model (ECM) can be used to examine price dynamics between FTSE100 spot and futures markets, with the lagged basis as the I(0) error-correction term (Brooks and Garrett, 2002).

In efficient markets, the equilibrium basis should equal the cost of carry. However, Green and Joujon (2000) argue that it is more reasonable to write the cost of carry relationship as a stochastic relationship, either because the interest rate and/or dividend are stochastic or because of a random error in the basis, possibly caused by differential arrival of information. Assuming the interest rate and dividend yield to be constant until expiration, but with a random error in the basis, they suggest re-writing the cost of carry relationship to incorporate a white noise error term, \(u_t\):

\[
f_t - s_t - (r - d)(T - t) = u_t
\]

or

\[
b_t - (r - d)(T - t) = u_t
\]

The \(u_t\) process of equation (4.4) must be unit-root stationary; otherwise there exist persistent arbitrage opportunities (Tsay, 2005). The cost of carry, \((r - d)(T - t)\), should be trend-stationary for individual contracts. It follows that the basis \(b_t\) is also trend-stationary for individual contracts, with a declining trend governed by the time to expiry. Green and Joujon (2000) suggest that the basis needs to be trend-corrected for individual contracts even though these are spliced together to form a single time series. To induce stationarity in the basis for each contract, the basis is trend-corrected by
allowing for the number of days remaining in the contract at each point in time, which is \( T - t \) at time \( t \). A linear combination of \( b_{t-k} \) and \( T - t + k_t \), which can be interpreted as the lagged trend-corrected basis, is therefore used as the error-correction term.

The time interval \( k_t \) is lengthened during weekends and holidays. *Ceteris paribus*, the longer an asset is held, the greater is the return, so that higher returns should be expected over weekends and holidays than between contiguous weekdays. The time interval \( k_t \) is therefore included in the model to measure the effect of time passing on price dynamics. In particular, stock returns might be affected by closure of the stock market during weekends and holidays. For a single contract, this implies a bivariate error correction model for spot and futures returns:

\[
\Delta_k s_t = \alpha_0 + \sum_{j=1}^{M_1} \alpha_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_{3j} k_{t-j} + \alpha_5 b_{t-k} + \alpha_6 (T - t + k_t) + e_{1t} \\
\Delta_k f_t = \beta_0 + \sum_{j=1}^{N_1} \beta_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{N_2} \beta_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{N_3} \beta_{3j} k_{t-j} + \beta_5 b_{t-k} + \beta_6 (T - t + k_t) + e_{2t}
\]

The futures price series is created here by combining successive contracts into a single time series, with the rollover taking place at the expiry of each futures contract. A problem arises from the rollover because there is an artificial 'jump' in the futures price, and therefore in the basis, at each rollover when switching from one contract to another. To model this jump, following Green and Joujon (2000), the dummy variable \( z_t \) is introduced. To understand the role of \( z_t \), it is helpful first to define \( z_t' \) as:

\[
z_t' = \begin{cases} 
T_g - t & \text{if } T_{g-1} - t + k_t = 0 \\
0 & \text{otherwise}
\end{cases}
\]

Each contract is indexed by \( g \), so that \( T_g \) is the expiration date of contract \( g \). Assuming that contract \( g-1 \) expires at time \( t - k_t \), then \( T_{g-1} = t - k_t \) or \( T_{g-1} - t + k_t = 0 \) and \( b_{t-k} = u_{t-k} \). On the next working day \( t \), contract \( g \) enters the
data and \( b_t = (r - d)(T_g - t) + u_t \). The theoretical change in the basis is therefore
\[ \Delta_k b_t = b_t - b_{t-k} = (r - d)(T_g - t) + \Delta_k u_t, \]
when a new contract enters the data, where \( \Delta_k u_t \) is a white noise error term. Since the interest rate and dividend yield, and hence \( r - d \), are assumed to be constant, it can be seen that \( z'_t \) is proportional to the expected change in the basis when a new contract enters the data and zero at all other times. Thus \( z'_t \) does indeed model the jump in the basis on the day when a new contract enters the data.

However, it is necessary to consider the combined effect of \( z'_t \) and \( k_t \) in modeling the change in the basis. Since \( k_t \) is already in the model, in the absence of contract rollovers, the basis should change by \( -k_t(r - d) \) from day to day, so that \( z'_t \) must take out that part of the rollover jump in the basis not already controlled by \( k_t \).

Therefore the jump in the basis at the time when a new contract enters the data, \( (r - d)(T_g - t) \), has to be modeled by adding back \( k_t \) to get \( z_t = T_g - t + k_t \). At time \( t \), when a new contract enters the series, the expected change in the basis will then be \( (r - d)(T_g - t + k_t) - k_t(r - d) = (r - d)(T_g - t) \), as required. The first term is modeled by \( z_t \) and the second term by \( k_t \). The dummy variable \( z_t \) is therefore defined as

\[
z_t = \begin{cases} T_g - t + k_t & \text{if } T_{g-1} - t + k_t = 0 \\ 0 & \text{otherwise} \end{cases}
\]

If the change in basis is written as \( \Delta_k b_t = \Delta_k f_t - \Delta_k s_t \), it can be seen that the jump in the basis at each rollover is actually the combined effect of the jump in the futures price less the jump in the spot price. \( z_t \) can therefore be used to model jumps in futures and spot prices at each rollover. The final bivariate VECM for the complete time series of spot and futures returns can therefore be specified as

---

\( \Delta_k b_t = b_t - b_{t-k} = (r - d)(T - t) + u_t - (r - d)(T - t + k_t) - u_{t-k} = -k_t(r - d) + \Delta_k u_t, \) where \( \Delta_k u_t \) is a white noise error term.

---

81
\[
\Delta k_s = \alpha_6 + \sum_{j=1}^{M_1} \alpha_{1j} \Delta k_s \Delta f_{i-1} + \sum_{j=1}^{M_2} \alpha_{2j} \Delta f_{i-1} + \sum_{j=0}^{M_3} \alpha_{3j} k_{i-1} + \sum_{j=0}^{M_3} \alpha_{4j} z_{i-1} + \alpha_5 \beta_{i-k} + \alpha_6 (T_k - t + k_i) + e_{it} \quad (4.8a)
\]

\[
\Delta k_f = \beta_6 + \sum_{j=1}^{N_1} \beta_{1j} \Delta k_s \Delta f_{i-1} + \sum_{j=1}^{N_2} \beta_{2j} \Delta f_{i-1} + \sum_{j=0}^{N_3} \beta_{3j} k_{i-1} + \sum_{j=0}^{N_3} \beta_{4j} z_{i-1} + \beta_5 \beta_{i-k} + \beta_6 (T_k - t + k_i) + e_{2t} \quad (4.8b)
\]

where \( e_{it} \) and \( e_{2t} \) are assumed to be identically and independently distributed (i.i.d.) random error terms with zero mean and constant variance. As usual, the VECM of equations (4.8a) and (4.8b) can be used to model both short-run price dynamics (indicated by \( \alpha_{1j}, \alpha_{2j}, \beta_{1j} \) and \( \beta_{2j} \)) and long-run error correction effect (indicated by \( \alpha_5, \alpha_6, \beta_5 \) and \( \beta_6 \)). In addition, distributed lags in \( k_i \) and \( z_i \) are included to match the corresponding effects on the current returns contained in the distributed lags of \( \Delta_k f_i \) and \( \Delta_k s_i \). Because an asset that is held longer should have a greater return, ceteris paribus, \( k_{i-j} \) is expected to have a positive effect on current price changes. \( z_{i-j} \) is used to model artificial futures price ‘jumps’ at contract rollovers, which would cause abrupt ‘jumps’ in the spread between futures and spot prices, or the basis, on those days. Therefore, \( z_{i-j} \) would be expected to have a positive joint effect on the current futures price change. \( z_{i-j} \) might also affect that part of the current spot price change (if any) which is an indirect reflection of artificial futures price jumps. The joint effect of \( z_{i-j} \) on the current futures price changes minus the joint effect of \( z_{i-j} \) on the current spot price changes should be positive, to reflect the wider spread between futures and spot at contract rollovers.

4.3.3 Joint Test of Causality and Cost of Carry

Both the causal relationships between FTSE100 index and futures prices and the validity of the cost of carry relationship between the two markets can be tested within the framework of equations (4.8a) and (4.8b). With respect to causality, futures do not cause spot if \( \alpha_{2j} = 0, \forall j \) and \( \alpha_5 = \alpha_6 = 0 \); while spot does not cause futures if \( \beta_{1j} = 0, \forall j \) and \( \beta_5 = \beta_6 = 0 \). The relevant error-correction coefficients (indicated by
$\alpha_5, \alpha_6$ for spot equation and $\beta_5, \beta_6$ for futures equation) must be insignificant or
causality will be bidirectional (Engle and Granger, 1987; Green and Joujon, 2000). Significant $\alpha_6$ and/or $\beta_6$ would suggest the necessity to induce stationarity in the basis for each individual futures contract. However, it is not possible to predict the sign of $\alpha_6$ and $\beta_6$, or how they should be related to $\alpha_5$ and $\beta_5$. An important reason for this is that the differential (the difference between the interest rate and the dividend yield, $r - d$) is not constant over time (as has been assumed). The time-variation of both interest rate and dividend yield is important in determining the basis (the spread between the futures price and the underlying spot price). All things being equal, the price of index futures will tend to rise relative to the underlying cash index if the interest rate rises, and vice versa. The opposite relationship applies to dividend yields where the index futures price will tend to fall relative to the underlying cash index if the dividend yield rises, and vice versa. From the dataset it can be seen that the interest rate and the dividend yield vary significantly over time, which complicates the relationship between the basis and time to maturity.

Assuming that the cost of carry relationship holds continuously as a stochastic relationship, which is perhaps the most common interpretation of the cost of carry theory, and can be specified as in equation (4.4), this will impose a set of testable cross-equation restrictions on the coefficients of (4.8a) and (4.8b). To work out the restrictions, equation (4.4) is differenced over the unit interval ($k$) between working days and conditional expectations are formed at time $t-k$ for equation (4.4) and its difference:

$$E_{t-k}b_t = (r - d)(T_g - t) \quad (4.9)$$

$$E_{t-k}\Delta_k b_t = \begin{cases} -k_t(r - d) & \text{within a single contract} \\ (r - d)(T_g - t) & \text{if a contract expires at time } t - k \end{cases} \quad (4.10)$$

Because (4.8b) less (4.8a) is equal to the change in the basis between successive working days, it is actually the realization of $\Delta_k b_t$. Hence (4.9) and (4.10) impose cross-equation restrictions on the coefficients of (4.8a) and (4.8b). These can be
evaluated by calculating (4.8b) less (4.8a), forming conditional expectations, and then using (4.9) and (4.10) repeatedly to equate coefficients (Green and Joujon, 2000).

**Derivation of cost-of-carry restrictions for the VECM equations (4.8a) and (4.8b)**

The stochastic continuous cost of carry relationship can be written as

\[ b_t - (T_g - t)R = u_t \]  

(4.11)

where \( R = r - d \) is the difference between the interest rate and dividend yield and \( u_t \) is a white noise error. Forming conditional expectations at time \( (t - k) \) of equation (4.11) and of the first difference of equation (4.11), gives

\[ Er_{-k}b_t = (T_g - t)R \]  

(4.12)

\[ Er_{-k}\Delta b_t = \begin{cases} -k_t R & \text{within a single contract} \\ (T_g - t)R & \text{if a contract expires at time} \ t - k \end{cases} \]  

(4.13)

(4.12) and (4.13) must be true for all values of the coefficients in the VECM (4.8a and 4.8b). Subtracting (4.8a) from (4.8b) will give the following equation, which allows two equation systems for spot and futures price changes to be written as a single equation in changes of the basis:

\[ \Delta b_t = (\beta_0 - \alpha_0) + (\beta_{11} - \alpha_{11})\Delta s_{t-1} + (\beta_{12} - \alpha_{12})\Delta s_{t-2} + (\beta_{13} - \alpha_{13})\Delta s_{t-3} + \ldots + (\beta_{30} - \alpha_{30})k_t + (\beta_{31} - \alpha_{31})k_{t-1} + (\beta_{32} - \alpha_{32})k_{t-2} + (\beta_{33} - \alpha_{33})k_{t-3} + \ldots + (\beta_{40} - \alpha_{40})z_t + (\beta_{41} - \alpha_{41})z_{t-1} + (\beta_{42} - \alpha_{42})z_{t-2} + (\beta_{43} - \alpha_{43})z_{t-3} + \ldots + (\beta_5 - \alpha_5)b_{t-k} + (\beta_6 - \alpha_6)(T_g - t + k_t) + (e_{21} - e_{11}) \]  

(4.14)

Equation (4.14) is actually the realization of \( \Delta b_t \). Hence (4.12) and (4.13) impose restrictions on the coefficients of (4.14) or cross-equation restrictions on the coefficients of (4.8a) and (4.8b).

In any time period
$$\beta_0 - \alpha_0 = 0 \quad (4.15)$$

and

$$\begin{align*}
\beta_{11} - \alpha_{11} &= - (\beta_{21} - \alpha_{21}) \\
\beta_{12} - \alpha_{12} &= - (\beta_{22} - \alpha_{22}) \quad \text{or} \quad \beta_{1j} - \alpha_{1j} = - (\beta_{2j} - \alpha_{2j}) \quad \forall j
\end{align*} \quad (4.16)$$

With restrictions (4.15) and (4.16), equation (4.14) can then be re-written as

$$\Delta_k b_t = (\beta_{21} - \alpha_{21}) \Delta_k b_{t-1} + (\beta_{22} - \alpha_{22}) \Delta_k b_{t-2} + (\beta_{23} - \alpha_{23}) \Delta_k b_{t-3} + \ldots$$

$$+ (\beta_{30} - \alpha_{30}) k_t + (\beta_{31} - \alpha_{31}) k_{t-1} + (\beta_{32} - \alpha_{32}) k_{t-2} + (\beta_{33} - \alpha_{33}) k_{t-3} + \ldots$$

$$+ (\beta_{40} - \alpha_{40}) z_t + (\beta_{41} - \alpha_{41}) z_{t-1} + (\beta_{42} - \alpha_{42}) z_{t-2} + (\beta_{43} - \alpha_{43}) z_{t-3} + \ldots$$

$$+ (\beta_{5} - \alpha_{5}) b_{t-k} + (\beta_{6} - \alpha_{6})(T_g - t + k_t) + (e_{2t} - e_{1t}) \quad (4.17)$$

First consider the case that all observations are within a single contract. Then $z_t = 0$ and $z_{t-j} = 0 \quad \forall j$. Given equations (4.12) and (4.13), forming conditional expectations of equation (4.17) within a single contract gives

$$-k_t R = (\beta_{21} - \alpha_{21})(-k_{t-1} R) + (\beta_{22} - \alpha_{22})(-k_{t-2} R) + (\beta_{23} - \alpha_{23})(-k_{t-3} R) + \ldots$$

$$+ (\beta_{30} - \alpha_{30}) k_t + (\beta_{31} - \alpha_{31}) k_{t-1} + (\beta_{32} - \alpha_{32}) k_{t-2} + (\beta_{33} - \alpha_{33}) k_{t-3} + \ldots$$

$$+ (\beta_{4} - \alpha_{4}) (T_g - t + k_t) R + (\beta_{6} - \alpha_{6})(T_g - t + k_t) \quad (4.18)$$

For equation (4.18) to be true for all $f_t, s_t, k_t$, the restrictions implied are

$$-R = \beta_{30} - \alpha_{30} \quad (4.19)$$

and

$$\begin{align*}
(\beta_{21} - \alpha_{21}) R &= (\beta_{31} - \alpha_{31}) \\
(\beta_{22} - \alpha_{22}) R &= (\beta_{32} - \alpha_{32}) \quad \text{or} \quad -R = -(\beta_{3j} - \alpha_{3j})/(\beta_{2j} - \alpha_{2j}) \quad \forall j
\end{align*} \quad (4.20)$$

and
\[-R = (\beta_6 - \alpha_6)/(\beta_5 - \alpha_5) \quad (4.21)\]

For observations crossing contracts, if contract \( g - 1 \) expires at time \( t - k \), and contract \( g \) enters the series at time \( t \), then \( b_{t-k} = (T_{g-1} - t + k)R \), \( E_{t-k}b_t = (T_g - t)R \), \( E_{t-k}\Delta_k b_t = (T_g - t)R \), \( T_{g-1} - t + k_i = 0 \), \( z_i = T_g - t + k_i \) and \( z_{i-j} = 0 \forall j \quad (0 < j < 90) \) (lags must be within one contract). Given the above, forming conditional expectations of equation (4.17) gives

\[
(T_g - t)R = (\beta_{21} - \alpha_{21})(-k_{i-1}R) + (\beta_{22} - \alpha_{22})(-k_{i-2}R) + (\beta_{23} - \alpha_{23})(-k_{i-3}R) + ... \\
+ (\beta_{30} - \alpha_{30})k_i + (\beta_{31} - \alpha_{31})k_{i-1} + (\beta_{32} - \alpha_{32})k_{i-2} + (\beta_{33} - \alpha_{33})k_{i-3} + ... \quad (4.22)
\]

The terms in \( k_{i-j} \) \((j > 0)\) and the error correction terms as well as \((T_g - t + k)\) involving \( \beta_5, \alpha_5, \beta_6, \alpha_6 \) have the same relationship as before. This leaves

\[
(T_g - t)R = (\beta_{30} - \alpha_{30})k_i + (\beta_{40} - \alpha_{40})(T_g - t + k_i) \quad (4.23)
\]

For consistency, it must be that \(-R = \beta_{30} - \alpha_{30}\), which implies that

\[
(T_g - t + k_i)R = (\beta_{40} - \alpha_{40})(T_g - t + k_i) \quad (4.24)
\]

Hence

\[-R = -(\beta_{40} - \alpha_{40}) \quad (4.25)\]

If contract \( g - 1 \) expires at \( t - k \), \( k_{i-1} \) and contract \( g \) enters the series at time \( t - k \), in which case \( E_{t-k}\Delta_k b_t = -k_iR \), \( T_g - t + k_i + k_{i-1} = 0 \), \( b_{t-k} = \Delta_k b_{i-1} = (T_g - t + k_i)R \), \( z_{i-1} = T_g - t + k_i + k_{i-1} \) and \( z_{i-j} = 0 \forall j, j \neq 1 \). Forming conditional expectations of equation (4.17) will then give

\[-k_iR = (\beta_{21} - \alpha_{21})(T_g - t + k_i)R + (\beta_{22} - \alpha_{22})(-k_{i-2}R) + (\beta_{23} - \alpha_{23})(-k_{i-3}R) + ... \\
+ (\beta_{30} - \alpha_{30})k_i + (\beta_{31} - \alpha_{31})k_{i-1} + (\beta_{32} - \alpha_{32})k_{i-2} + (\beta_{33} - \alpha_{33})k_{i-3} + ... \quad (4.26)
\]

\[
+ (\beta_{41} - \alpha_{41})(T_g - t + k_i + k_{i-1}) + (\beta_{5} - \alpha_{5})(T_g - t + k_i)R + (\beta_{6} - \alpha_{6})(T_g - t + k_i)
\]

86
The terms in \( k_{r-j} \) \(( j > 1)\) and the error correction terms as well as \( (T_g - t + k_r)\) involving \( \beta_s, \alpha_s, \beta_6, \alpha_6 \) have the same relationship as before. It is also the case that 

\[-R = \beta_{30} - \alpha_{30} \]

This leaves

\[
(\beta_{21} - \alpha_{21})(T_g - t + k_r)R + (\beta_{31} - \alpha_{31})k_{r-1} + (\beta_{41} - \alpha_{41})(T_g - t + k_r + k_{r-1}) = 0 \quad (4.27)
\]

But \(-R = -(\beta_{31} - \alpha_{31})/(\beta_{21} - \alpha_{21}), \) which implies that

\[
(\beta_{21} - \alpha_{21})(T_g - t + k_r + k_{r-1})R + (\beta_{41} - \alpha_{41})(T_g - t + k_r + k_{r-1}) = 0 \quad (4.28)
\]

Hence

\[-R = (\beta_{41} - \alpha_{41})/(\beta_{21} - \alpha_{21}) \quad (4.29)\]

Doing the same thing for a contract expiring at \( t - k_r - k_{r-1} - k_{r-2} \), at \( t - k_r - k_{r-1} - k_{r-2} - k_{r-3} \), and etc., gives the further restrictions

\[
\left\{ 
\begin{align*}
-R &= (\beta_{42} - \alpha_{42})/(\beta_{22} - \alpha_{22}) \\
-R &= (\beta_{43} - \alpha_{43})/(\beta_{23} - \alpha_{23}) \quad \text{or} \quad -R = (\beta_{4j} - \alpha_{4j})/(\beta_{2j} - \alpha_{2j}) \quad (4.30)
\end{align*}
\right.
\]

Putting (4.15), (4.16), (4.19), (4.20), (4.21), (4.25) and (4.30) together and replacing \( R \) by \((r - d)\) gives the restrictions 1 to 6 imposed by the stochastic continuous cost of carry relationship on equations (4.8a) and (4.8b), which are listed in Table 4.5 (restriction 7 follows from equations (4.31a) and (4.31b) to account for 1987 stock market crash effect, see below). Different restrictions have different implications. For example, restriction 1 and restriction 7 imply that the expected return for investing in spot and in futures markets should be equal under any market conditions (even when market crashes), while restriction 3 indicates that the time passing itself should have an equivalent effect on spot and on futures price changes. These are reasonable given that the two markets are actually different places for trading the same underlying asset. The short-run inter- and intra-market dynamics for spot and futures prices can be
represented by restriction 2, while the long-run relationship between spot and futures prices can be represented by restrictions 4, 5 and 6 jointly. Specifically, restrictions 4 and 5 are associated with the jump in the basis when a new contract enters the data, which can be used to examine the extent to which the basis is correctly priced at 3 months from delivery. Restriction 6 indicates the stability of the long-run relationship between the basis and the time to delivery of the futures contracts.

Table 4.5 Testable restrictions imposed by the continuous stochastic cost of carry relationship (equation 4.4) on selected VECM equations (4.8a) and (4.8b)

1. $\beta_0 - \alpha_0 = 0$
2. $\beta_{1j} - \alpha_{1j} = - (\beta_{2j} - \alpha_{2j}) \quad \forall j$
3. $\beta_{30} - \alpha_{30} = (\beta_{3j} - \alpha_{3j}) / (\beta_{2j} - \alpha_{2j}) \quad [= -(r - d)] \quad \forall j$
4. $\beta_{30} - \alpha_{30} = - (\beta_{40} - \alpha_{40}) \quad [= -(r - d)]$
5. $\beta_{30} - \alpha_{30} = (\beta_{4j} - \alpha_{4j}) / (\beta_{2j} - \alpha_{2j}) \quad [= -(r - d)] \quad \forall j$
6. $\beta_{30} - \alpha_{30} = (\beta_{6} - \alpha_{6}) / (\beta_{5} - \alpha_{5}) \quad [= -(r - d)]$
7. $\beta_{7j} - \alpha_{7j} = 0 \quad \forall j$

As argued by Green and Joujon (2000), once the restrictions imposed by cost of carry are set out, it transpires that it is not true that causality from futures to spot or from spot to futures is necessarily inconsistent with cost of carry. The reason is that the cost of carry model predicts movements (or changes) in the basis and not in either the spot or the futures price on their own. Therefore, with respect to the cost of carry theory, it is the difference between equations (4.8b) and (4.8a) that is of significance, not the coefficients in each individual equation. However, the cost of carry model is inconsistent with different lag lengths in spot and futures equations and one-way causality from the variable with the longer lag length. The same lag length is therefore imposed in both spot and futures equations in the empirical analysis.
4.4 Empirical results

4.4.1 Causality and cointegration

To confirm the conjecture that the dynamics between FTSE100 spot and futures prices should be conducted over sub-samples corresponding to different trading systems employed by both the LSE and LIFFE, the VECM for spot and futures returns (equations (4.8a) and (4.8b)) was first estimated over the whole sample period. Chow tests were then used to examine whether indeed there were structural breaks on the two important dates, 20/10/1997 and 30/11/1998. Several procedures were used to finalize the model specification and estimation. First of all, it is necessary to select reasonable lag orders to capture the features of the data, such as serial correlation in the spot and futures return series and their inter-market dependence. For this purpose information criteria (including sequential modified LR test, Final Prediction Error, Akaike Information Criterion, Schwarz Information Criterion, and Hannan-Quinn Information Criterion) were used. It is worth noting that SIC embodies a much stiffer penalty term and tends to be more conservative than AIC in selecting the lag length, while HQ is somewhere in between. In the present case identical lags are used for variables $\Delta_k s_t$, $\Delta_k f_t$, $k_t$ and $z_t$ in both equations. The number of parameters goes up quickly but, beyond the first lag or two, most of the new additions tend to be unimportant, causing SIC to reject longer lags in favor of shorter ones. Being opposite to SIC, LR is likely to select a very large model. Overall, no criterion is definitely superior to others. The equations are likely to need longer lags to handle the dynamics of the data adequately but a model containing irrelevant lags of a variable (and therefore unnecessary parameters) will lead to increased coefficient standard errors, making it more difficult to find significant relationships in the data (Brooks, 2002). To form both a parsimonious and a relatively adequate model, both information criteria and a Ljung-Box Q test on residuals have been used to check model adequacy, while restricting the model to a maximum of 5 lags. Because daily data are used in this research, 5 lags correspond to a weekly effect, which we believe is long enough to handle any short-run dynamics of FTSE100 spot and futures prices given that both FTSE100 component shares and FTSE100 futures are highly frequently traded.

A bivariate VAR comprising equations (4.8a) and (4.8b) with 5 lags was initially estimated using Ordinary Least Squares (OLS). The lowest optimal lag order suggested
by any information criterion (usually SIC and/or HQ) was first chosen for the VAR. The Ljung-Box Q statistic was then applied to the residuals to check the assumption of no serial or cross-correlation. Where there was evidence of significant serial and/or cross-correlation in the residuals up to 5 lags (compared with 2 standard error bounds), one lag was added and the augmented model was estimated. The Ljung-Box Q test was then applied to the residuals of the augmented model to check model adequacy. The final lag order of the model was chosen where residuals displayed no significant serial or cross-correlation up to 5 lags. The lag orders (for the whole sample) selected by different information criteria are reported in Table 4.6, which shows that SIC selects 3 lags and all other information criteria select 4 lags. The Ljung-Box Q test on residuals shows that at least 4 lags (for the whole sample) had to be included to ensure no significant serial or cross correlation in the residuals up to 5 lags. Serial and cross correlation up to 5 lags in the residuals estimated from the finalized VAR of 4 lags are displayed in Figure 4.7.
Table 4.6 VAR lag order selected by different information criteria

28/10/1986 – 30/12/2005

<table>
<thead>
<tr>
<th>Lag</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SIC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>NA</td>
<td>1.35e-09</td>
<td>-14.7452</td>
<td>-14.7103</td>
<td>-14.7330</td>
</tr>
<tr>
<td>1</td>
<td>638.4975</td>
<td>1.19e-09</td>
<td>-14.8762</td>
<td>-14.8360</td>
<td>-14.8621</td>
</tr>
<tr>
<td>3</td>
<td>89.0084</td>
<td>1.11e-09</td>
<td>-14.9403</td>
<td>-14.8893*</td>
<td>-14.9224</td>
</tr>
<tr>
<td>4</td>
<td>25.6716*</td>
<td>1.11e-09</td>
<td>-14.9440*</td>
<td>-14.8876</td>
<td>-14.9242*</td>
</tr>
<tr>
<td>5</td>
<td>7.6528</td>
<td>1.11e-09</td>
<td>-14.9439</td>
<td>-14.8822</td>
<td>-14.9223</td>
</tr>
</tbody>
</table>

Key:
* indicates the lag order selected by the criterion
LR: sequential modified LR test statistic
FPE: Final prediction error
AIC: Akaike information criterion
SIC: Schwarz information criterion
HQ: Hannan-Quinn information criterion
Figure 4.7 Serial and cross-correlations of the residuals estimated from the VAR of equations (4.8a) and (4.8b) (4 lags), 28/10/1986 – 30/12/2005

Autocorrelations with 2 Std.Err. Bounds

Key:

*DS* is the first difference of spot price, or spot price change

*DF* is the first difference of futures price, or futures price change.
After the introduction of SETS by the LSE on 20/10/1997 and LIFFE CONNECT by LIFFE on 30/11/1998, transaction costs in both markets should have declined. This could have significant impact on index arbitrage and hence the relationship between FTSE100 cash and futures prices. It is therefore necessary to check whether the estimated coefficients are stable over the whole sample, because if an estimated model fails to show stable coefficients, inference using it may be suspect. Since these two important changes in microstructure are already known, Chow tests seem to be both sufficiently powerful and suitable for detecting coefficient instability. The null hypothesis of Chow tests is that coefficients are stable over the whole sample, with the alternative hypothesis of a break at a known point in the sample (20/10/1997 and 30/11/1998 in this case). The results of both a Chow breakpoint test and a Chow forecast test are reported in Table 4.7. It can be seen that the null hypothesis of no break at 20/10/1997 or 30/11/1998 is decisively rejected by both tests. As a result, it seems natural to analyse the FTSE100 spot-futures price dynamics over three sub-samples: 28/10/1986 to 17/10/1997, 20/10/1997 to 27/11/1998 and 30/11/1998 to 30/12/2005. However, as argued above, the middle time period is too short for empirical work. For the sub-sample 30/11/1998 to 30/12/2005, as will be seen later, estimation of the DCC-TGARCH in Chapter 5 failed to achieve convergence; estimation of the threshold VECM in Chapter 6 resulted in a very small number of observations (only 32) in the upper regime, implying potential bias in inference. The date when SEAQ was replaced by SETS (20/10/1997) at the LSE was therefore used to separate the whole sample into two sub-samples: 28/10/1986–17/10/1997 and 27/10/1997–30/12/2005. The basic linear VECM was nonetheless estimated over 30/11/1998–30/12/2005 and the main results were found to be qualitatively identical with those for 27/10/1997–30/12/2005. This suggests that the mixed market microstructure underlying the second sub-sample used in this research (a mix of ‘open outcry’ and electronic platform at LIFFE) should not have particularly strong implications for the general behaviour of FTSE100 spot and futures.
Table 4.7 Chow tests for stability of estimated coefficients of equations (4.8a) and (4.8b) (4 lags), 28/10/1986 – 30/12/2005

<table>
<thead>
<tr>
<th>Null hypothesis: 20/10/1997 is not a break date</th>
<th>Spot equation (4.8a)</th>
<th>Futures equation (4.8b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chow breakpoint test</td>
<td>F-statistic</td>
<td>P-value</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>2.9598</td>
<td>0.00001</td>
<td>4.2277</td>
</tr>
<tr>
<td>Chow forecast test</td>
<td>F-statistic</td>
<td>P-value</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>1.7813</td>
<td>0.00000</td>
<td>1.3836</td>
</tr>
</tbody>
</table>

Null hypothesis: 30/11/1998 is not a break date

<table>
<thead>
<tr>
<th>Spot equation (4.8a)</th>
<th>Futures equation (4.8b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chow breakpoint test</td>
<td>F-statistic</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>2.7416</td>
<td>0.00003</td>
</tr>
<tr>
<td>Chow forecast test</td>
<td>F-statistic</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>1.5408</td>
<td>0.00000</td>
</tr>
</tbody>
</table>

Note: The null hypothesis is that coefficients are stable over the whole sample. The alternative hypothesis is that there is a structural break at either 20/10/1997 or 30/11/1998.

The procedure of selecting lag order discussed above was repeated for the two subsamples. For 28/10/1986–17/10/1997, the lag orders selected by different information criteria are reported in Table 4.8, which shows that SIC selects 3 lags and all other information criteria select 4 lags. The Ljung-Box Q test shows that at least 4 lags must be included to ensure no significant serial or cross-correlation in the residuals for up to 5 lags. Serial and cross-correlation for up to 5 lags in the residuals estimated from the finalized VAR of 4 lags is presented in Figure 4.8. The lag orders selected by different information criteria for 27/10/1997–30/12/2005 are reported in Table 4.9, which shows that SIC and HQ select 4 lags and all other information criteria select 5 lags. The Ljung-Box Q test shows that at least 5 lags must be included to ensure no significant serial or cross-correlation in the residuals. Serial and cross-correlation for up to 5 lags in the residuals estimated from the finalized VAR of 5 lags are presented in Figure 4.9.
Table 4.8 VAR lag order selected by different information criteria

28/10/1986 – 17/10/1997

<table>
<thead>
<tr>
<th>Lag</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SIC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>NA</td>
<td>1.24e-09</td>
<td>-14.8285</td>
<td>-14.7729</td>
<td>-14.8085</td>
</tr>
<tr>
<td>1</td>
<td>399.6466</td>
<td>1.08e-09</td>
<td>-14.9707</td>
<td>-14.9065</td>
<td>-14.9475</td>
</tr>
<tr>
<td>2</td>
<td>128.4377</td>
<td>1.03e-09</td>
<td>-15.0144</td>
<td>-14.9417</td>
<td>-14.9882</td>
</tr>
<tr>
<td>3</td>
<td>35.9522</td>
<td>1.02e-09</td>
<td>-15.0246</td>
<td>-14.9433</td>
<td>-14.9952</td>
</tr>
<tr>
<td>4</td>
<td>17.9170*</td>
<td>1.02e-09*</td>
<td>-15.0282*</td>
<td>-14.9384</td>
<td>-14.9958*</td>
</tr>
<tr>
<td>5</td>
<td>6.5599</td>
<td>1.02e-09</td>
<td>-15.0277</td>
<td>-14.9293</td>
<td>-14.9922</td>
</tr>
</tbody>
</table>

Key:
* indicates the lag order selected by the criterion

LR: sequential modified LR test statistic
FPE: Final prediction error
AIC: Akaike information criterion
SIC: Schwarz information criterion
HQ: Hannan-Quinn information criterion
Figure 4.8 Serial and cross-correlation in the residuals estimated from the VAR of equations (4.8a) and (4.8b) (4 lags), 28/10/1986 – 17/10/1997

Autocorrelations with 2 Std.Err. Bounds

Key:

DS is the first difference of spot price, or spot price change

DF is the first difference of futures price, or futures price change.
Table 4.9 VAR lag order selected by different information criteria

27/10/1997 – 30/12/2005

<table>
<thead>
<tr>
<th>Lag</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SIC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>NA</td>
<td>1.08e-09</td>
<td>-14.9737</td>
<td>-14.8912</td>
<td>-14.9434</td>
</tr>
<tr>
<td>4</td>
<td>33.1937</td>
<td>8.00e-10</td>
<td>-15.2704</td>
<td>-15.1440*</td>
<td>-15.2241*</td>
</tr>
<tr>
<td>5</td>
<td>13.5856*</td>
<td>7.98e-10*</td>
<td>-15.2732*</td>
<td>-15.1358</td>
<td>-15.2228</td>
</tr>
</tbody>
</table>

Key:

* indicates the lag order selected by the criterion

LR: sequential modified LR test statistic
FPE: Final prediction error
AIC: Akaike information criterion
SIC: Schwarz information criterion
HQ: Hannan-Quinn information criterion
Figure 4.9 Serial and cross-correlations of the residuals estimated from the VAR of equations (4.8a) and (4.8b) (5 lags), 27/10/1997 – 30/12/2005

Autocorrelations with 2 Std.Err. Bounds

Key:

$DS$ is the first difference of spot price, or spot price change.

$DF$ is the first difference of futures price, or futures price change.
A CUSUM test was used to check the stability of the estimated coefficients in the two sub-samples. The CUSUM test is based on recursive residuals estimated from the regression. Under the null hypothesis that coefficients are stable, the cumulated sums of the recursive residuals should act like a random walk. If there is a structural break, however, they will tend to drift above the bounding lines, which are normally set for a 5% significance level. The CUSUM test (see Figure 4.10 and Figure 4.11) suggests that the coefficients are fairly stable over both sub-samples.

**Figure 4.10 The plots of CUSUM against 5% critical bounds for residuals of equation (4.8a) and equation (4.8b), 28/10/1986 – 17/10/1997**
Figure 4.11 The plots of CUSUM against 5% critical bounds for residuals of equation (4.8a) and equation (4.8b), 27/10/1997 – 30/12/2005
In a multivariate regression model, the errors in different equations may be correlated. Indeed, it was found that the residuals of OLS estimation of equations (4.8a) and (4.8b) were highly contemporaneously correlated across equations, with a correlation coefficient of 0.9453 for the first period 28/10/1986–17/10/1997 and 0.9819 for the second period 27/10/1997–30/12/2005. OLS estimates each equation independently by minimizing the sum of squared residuals (RSS) and does not account for cross-equation correlations. However, according to the general theory of the least squares method, which takes the covariation of errors into account, multiple equations should be solved simultaneously. Otherwise, minimum variance of the errors in the estimated regression parameters cannot be achieved. The efficiency of the estimation may be improved by adjusting for this cross-equation correlation. Proposed by Zellner (1962), Seemingly Unrelated Regression (SUR), also called joint generalized least squares (JGLS) or Zellner estimation, is a generalization of OLS in which multiple equations are estimated simultaneously while allowing for contemporaneous correlation between equations. Like OLS, the SUR method assumes that all the regressors are independent variables, but SUR uses the cross-equation correlation in residuals to improve the regression estimates. In particular, the SUR method requires an initial OLS estimation of each regression separately and the OLS residuals are then used to estimate the cross-equation covariance matrix of the residuals, \( \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon}_{t} \hat{\epsilon}_{t}^{\prime} \), which is then used to obtain the minimal variance of the errors in a multivariate regression context (Zellner, 1962; Greene, 2003). The SUR method was therefore used to estimate the VECM of (4.8a) and (4.8b), with 4 lags for the first period and 5 lags for the second period.

For the period 28/10/1986–17/10/1997, the time plots of residuals estimated from equations (4.8a) and (4.8b) (displayed in Figure 4.12 and Figure 4.13 respectively) show outliers around the stock market crash of 19/10/1987 (‘Black Monday’). This indicates that to avoid spurious inference, the 1987 crash needs to be accommodated explicitly in the model (see below for equations (4.31a) and (4.31b)). Indeed, Antoniou and Garrett (1993) also found that the pricing relationship between FTSE100 spot and futures markets have actually broken on 19/10/1987, although it was then restored on 20/10/1987. For the period 27/10/1997–30/12/2005, the time plots of residuals estimated from equations (4.8a) and (4.8b) (displayed in Figure 4.14 and Figure 4.15 respectively) indicate that the residual series are uniformly very close to zero.
Figure 4.12 The plot of residuals estimated from equation (4.8a) for spot returns, 28/10/1986 – 17/10/1997

Spot residuals, 03/11/1986 – 17/10/1997
Figure 4.13 The plot of residuals estimated from equation (4.8b) for futures returns, 28/10/1986 – 17/10/1997

Futures residuals, 03/11/1986 – 17/10/1997
Figure 4.14 The plot of residuals estimated from equation (4.8a) for spot returns, 27/10/1997 – 30/12/2005
Figure 4.15 The plot of residuals estimated from equation (4.8b) for futures returns, 27/10/1997 – 30/12/2005

Futures residual, 03/11/1997 – 30/12/2005
Given the obvious evidence of stock market crash and the reported evidence by Antoniou and Garrett (1993) that the linkage between FTSE100 spot and futures markets broke on 19/10/1987, a dummy variable $DM$ taking the value 1 on 19/10/1987 and 0 otherwise was defined and added to equations (4.8a) and (4.8b) to capture the 1987 stock market crash effect for the sample period 28/10/1986–17/10/1997. Since both futures and spot prices fell dramatically for several consecutive days following the crash, lags of $DM$ were also included to model any persistent effects it may have had. The VECM augmented by including $DM$ and lags of $DM$ can be specified as

$$\Delta_k s_t = \alpha_0 + \sum_{j=1}^{M_1} \alpha_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_{3j} k_{t-j} + \sum_{j=0}^{M_4} \alpha_{4j} z_{t-j} + \alpha_5 b_{t-k} + \alpha_6 (T_g - t + k_t) + \sum_{j=0}^{M_4} \alpha_{7j} DM_{t-j} + e_{1t}$$

(4.31a)

$$\Delta_k f_t = \beta_0 + \sum_{j=1}^{N_1} \beta_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{N_2} \beta_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{N_3} \beta_{3j} k_{t-j} + \sum_{j=0}^{N_4} \beta_{4j} z_{t-j} + \beta_5 b_{t-k} + \beta_6 (T_g - t + k_t) + \sum_{j=0}^{N_4} \beta_{7j} DM_{t-j} + e_{2t}$$

(4.31b)

The number of lags was chosen to ensure that no obvious outlier in the residuals was identified around the crash (4 lags of $DM$ were finally chosen). Due to the inclusion of $DM$ and lags of $DM$, the cost of carry model imposes another restriction (restriction 7 in Table 4.5) in addition to restrictions 1 to 6 on the coefficients of (4.31a) and (4.31b). Equations (4.31a) and (4.31b) were again estimated by SUR for the sample period 28/10/1986–17/10/1997. The time plots of the estimated residuals are displayed in Figure 4.16 and Figure 4.17 respectively. They show that there is no obvious outlier around the time of the 1987 stock market crash and that the residual series are uniformly very close to zero. Thus reliable inference for the period 28/10/1986–17/10/1997 can be made using results estimated from equations (4.31a) and (4.31b).
Figure 4.16 The plot of residuals estimated from equation (4.31a) for spot returns, 28/10/1986 – 17/10/1997
Figure 4.17 The plot of residuals estimated from equation (4.31b) for futures returns, 28/10/1986 – 17/10/1997

Futures residuals, 03/11/1986 – 17/10/1997
Table 4.10 reports the SUR estimation results of equations (4.31a) and (4.31b) for the period 28/10/1986-17/10/1997. It can be seen that there is bidirectional causality between futures and spot markets, with causality from spot to futures being much stronger. The stronger causality from spot to futures can be explained by the fact that the FTSE100 index is a quote-based index, which in theory can reflect information instantly because market makers (or market participants) are able to adjust their quotes for the index component shares immediately in response to new information shocks. On the other hand, the FTSE100 futures price can reflect new information only if transactions on the futures contract have taken place. But transaction costs in the futures market are usually lower than in the spot market, which can sometimes result in the lead of the futures price over the spot price. It can also be seen that those coefficients of lagged cross-market returns that are significant are all positive, implying that past price movements in the spot/futures market have positive impact on current price movements in the futures/spot market. This is consistent with expectation, since where one market leads another in its response to new information, the lagged market should respond positively to the price movement in the leading market to keep the spread between the two markets within a no-arbitrage band.

It is noticed that both spot and futures returns tend to be mean-reverting, indicated by the significant negative impact of lagged returns on the current return. But it is also evident that there is much stronger evidence of mean-reversion in futures than in spot returns. This possibly indicates that, if the futures price overreacts to new information, subsequent corrections in the futures price are driven by the true value contained in the information. Overreaction of the futures price to new information is possible because of transaction advantages of trading in futures markets such as lower transaction costs, higher liquidity and small capital outlay for the margin, allowing leverage. Furthermore, position building can be very rapid because the purchase of one futures contract represents a claim on the whole index portfolio. If the futures price overreacts to certain information, mean-reversion is necessary for the markets to be efficient in the long run. Another common explanation for the mean-reversion in futures returns comes from the bid/ask bounce effect in futures prices (Roll, 1984). Stock indices are unlikely to be subject to bid-ask bounce because effects present in the prices of the component shares are likely to be offsetting when aggregated. Furthermore, because the computation of
the FTSE100 index is based on mid-point quotes for component shares, the FTSE100 index actually does not suffer from this problem.

Table 4.10 SUR estimation of equations (4.31a) and (4.31b)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>-0.003705</td>
<td>-2.307332</td>
<td>-0.005815</td>
<td>-3.101861</td>
</tr>
<tr>
<td>$\Delta_k s_{t-1}$</td>
<td>-0.112686</td>
<td>-1.949060 $^c$</td>
<td>0.368936</td>
<td>5.465959 $^a$</td>
</tr>
<tr>
<td>$\Delta_k s_{t-2}$</td>
<td>0.040244</td>
<td>0.653919</td>
<td>0.303943</td>
<td>4.230354 $^a$</td>
</tr>
<tr>
<td>$\Delta_k s_{t-3}$</td>
<td>0.012679</td>
<td>0.212238</td>
<td>0.145279</td>
<td>2.083098 $^b$</td>
</tr>
<tr>
<td>$\Delta_k s_{t-4}$</td>
<td>0.071496</td>
<td>1.419857</td>
<td>0.127427</td>
<td>2.167649 $^b$</td>
</tr>
<tr>
<td>$\Delta_k f_{t-1}$</td>
<td>0.165395</td>
<td>3.245963 $^a$</td>
<td>-0.275689</td>
<td>-4.634489 $^a$</td>
</tr>
<tr>
<td>$\Delta_k f_{t-2}$</td>
<td>-0.025309</td>
<td>-0.465253</td>
<td>-0.292848</td>
<td>-4.611190 $^a$</td>
</tr>
<tr>
<td>$\Delta_k f_{t-3}$</td>
<td>-0.043836</td>
<td>-0.830645</td>
<td>-0.174616</td>
<td>-2.834232 $^a$</td>
</tr>
<tr>
<td>$\Delta_k f_{t-4}$</td>
<td>-0.018185</td>
<td>-0.398508</td>
<td>-0.078554</td>
<td>-1.474517</td>
</tr>
<tr>
<td>$k_t$</td>
<td>-2.60E-05</td>
<td>-0.103062</td>
<td>0.00017</td>
<td>0.576921</td>
</tr>
<tr>
<td>$k_{t-1}$</td>
<td>0.000741</td>
<td>2.702191 $^a$</td>
<td>0.001081</td>
<td>3.379507 $^a$</td>
</tr>
<tr>
<td>$k_{t-2}$</td>
<td>0.000573</td>
<td>2.055980 $^b$</td>
<td>0.00096</td>
<td>2.953817 $^a$</td>
</tr>
<tr>
<td>$k_{t-3}$</td>
<td>0.000786</td>
<td>2.873384 $^a$</td>
<td>0.001089</td>
<td>3.411606 $^a$</td>
</tr>
<tr>
<td>$k_{t-4}$</td>
<td>0.000710</td>
<td>2.825309 $^a$</td>
<td>0.00082</td>
<td>2.795292 $^a$</td>
</tr>
<tr>
<td>$z_t$</td>
<td>-2.59E-06</td>
<td>-0.177202</td>
<td>0.000115</td>
<td>6.719164 $^a$</td>
</tr>
<tr>
<td>$z_{t-1}$</td>
<td>1.29E-06</td>
<td>0.081719</td>
<td>5.44E-05</td>
<td>2.947861 $^a$</td>
</tr>
<tr>
<td>$z_{t-2}$</td>
<td>8.86E-07</td>
<td>0.055674</td>
<td>2.87E-05</td>
<td>1.542975</td>
</tr>
<tr>
<td>$z_{t-3}$</td>
<td>1.39E-05</td>
<td>0.882413</td>
<td>3.00E-05</td>
<td>1.630347</td>
</tr>
<tr>
<td>$z_{t-4}$</td>
<td>-1.30E-06</td>
<td>-0.084851</td>
<td>1.11E-05</td>
<td>0.619335</td>
</tr>
<tr>
<td>$b_{t-k_{t}}$</td>
<td>0.005115</td>
<td>0.167270</td>
<td>-0.120918</td>
<td>-3.386841 $^a$</td>
</tr>
<tr>
<td>$T_g - t + k_t$</td>
<td>2.51E-06</td>
<td>0.324718</td>
<td>1.39E-05</td>
<td>1.540666</td>
</tr>
<tr>
<td>$DM$</td>
<td>-0.113117</td>
<td>-13.52158 $^a$</td>
<td>-0.169685</td>
<td>-17.37410 $^a$</td>
</tr>
<tr>
<td>$DM(-1)$</td>
<td>-0.116113</td>
<td>-13.04729 $^a$</td>
<td>-0.109275</td>
<td>-10.51769 $^a$</td>
</tr>
<tr>
<td>$DM(-2)$</td>
<td>0.077178</td>
<td>8.400474 $^a$</td>
<td>0.084496</td>
<td>7.877844 $^a$</td>
</tr>
<tr>
<td>$DM(-3)$</td>
<td>-0.067394</td>
<td>-7.276797 $^a$</td>
<td>-0.090519</td>
<td>-8.371844 $^a$</td>
</tr>
<tr>
<td>$DM(-4)$</td>
<td>-0.013071</td>
<td>-1.395689</td>
<td>-0.032227</td>
<td>-2.947421 $^a$</td>
</tr>
</tbody>
</table>

Notes:
- $^a$ significant at the 1% level
- $^b$ significant at the 5% level
- $^c$ significant at the 10% level
The coefficients for $k_{t-j}$ ($j=1, 2, 3, 4$) are positive and significant in both spot and futures equations. This indicates that the passage of time has a positive, though quantitatively small, effect on price changes in both markets. This is consistent with expectation, because returns over longer time intervals such as weekends and holidays should be higher than returns over shorter time intervals, *ceteris paribus*. Furthermore, $z_{t}$ and $z_{t-1}$ are positive and significant in the futures return equation, suggesting that the ‘artificial’ price jumps at contract rollovers should not be ignored in explaining the futures price dynamics.

The error correction term, or the lagged basis, is significant in the futures equation, further indicating the stronger lead of the spot price over the futures price. The stronger causality from spot to futures price, indicating stronger lead of the spot price over the futures price in response to new information, may have caused temporary divergence from equilibrium between the two markets in the period analysed. Subsequent adjustments in the futures market would then have been necessary to drive the futures price (and the relationship between the two markets) to a new equilibrium. It could also be that the ease of trading in futures markets may have caused the futures price to have overreacted to new information, causing the divergence from equilibrium. Subsequent adjustment in the futures market would therefore have been necessary to correct the overreaction of the futures price and the divergence from equilibrium. The coefficient on the error-correction term ($b_{t-1}$) is -0.120918, suggesting that about 12 percent of the divergence from equilibrium was corrected in the futures market within one day. The negative sign implies that if the basis was too large/small compared to the equilibrium value, the futures price tended to fall/rise to recover the no-arbitrage relationship. However, the time to maturity term is found to be insignificant.

The dummy variable $DM$ and its lags, excluding $DM(-2)$, are negative and significant in both equations, reflecting a general fall in all asset prices during the crash period. The parameter on $DM(-2)$ is positive because there was a price increase on the second trading day following ‘Black Monday’, as can be seen from the dataset. The magnitudes of the coefficients on $DM$ and its lags in the futures equation are in general greater than those in the spot equation (with the exception of $DM(-1)$), implying that the FTSE100 futures price dropped more than the spot price and that a one-to-one no-
arbitrage relationship between the futures and spot markets was disrupted during the crash period. However, the coefficients on $DM(-1)$ and $DM(-2)$ in the spot and futures equations have the same sign and are very close in size, suggesting that the no-arbitrage link between the two markets was restored on the two trading days immediately following the ‘Black Monday’. This is consistent with the results of Antoniou and Garrett (1993), who have reported that the no-arbitrage link between the two markets was disrupted on 19/10/1987 and restored on 20/10/1987.

Table 4.11 reports the SUR estimation results of equations (4.8a) and (4.8b) for 27/10/1997–30/12/2005. Unidirectional causality from spot to futures is found during this sample period, with lagged spot returns having a positive impact on the current futures return. Again, the futures return series display very strong evidence of mean-reversion, reflecting either a tendency of futures prices to overreact to new information, with subsequent correction, or a bid-ask bounce effect in the futures price. The coefficients for $k_t$ and $k_{t-3}$ are significant in the spot return equation, but only at the 10% level, while all coefficients of $k_t$ and its lags are non-significant in the futures return equation. For $z_t$ and its lags in the futures return equation, only $z_{t-4}$ is marginally significant (positive) at the 10% level. These results suggest that compared to other factors in the model, the time passing effect and the ‘artificial’ price jumps at contract rollovers are not so important in explaining spot and futures price dynamics during this period as in the first period. It is noticeable that the error correction term is insignificant in both the spot and futures equations, suggesting either that there is no particular divergence from long-run equilibrium during this sample period, possibly because there is sufficient continual adjustment in one or both markets, or that such divergence has been rapidly eliminated by short-run price adjustments rather than by the actions of arbitrageurs. The time to maturity term is also found to be insignificant in this sample period.
Table 4.11 SUR estimation of equations (4.8a) and (4.8b)

27/10/1997 – 30/12/2005

<table>
<thead>
<tr>
<th>Variables</th>
<th>( \Delta_k s_t ) Coefficient</th>
<th>( \Delta_k s_t ) t-statistic</th>
<th>( \Delta_k f_t ) Coefficient</th>
<th>( \Delta_k f_t ) t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.002641</td>
<td>-1.096183</td>
<td>-0.00263</td>
<td>-1.060476</td>
</tr>
<tr>
<td>( \Delta_k s_{t-1} )</td>
<td>-0.13492</td>
<td>-1.010392</td>
<td>0.545386</td>
<td>3.967574</td>
</tr>
<tr>
<td>( \Delta_k s_{t-2} )</td>
<td>-0.129191</td>
<td>-0.851298</td>
<td>0.344372</td>
<td>2.204376</td>
</tr>
<tr>
<td>( \Delta_k s_{t-3} )</td>
<td>0.140193</td>
<td>0.906831</td>
<td>0.468587</td>
<td>2.944415</td>
</tr>
<tr>
<td>( \Delta_k s_{t-4} )</td>
<td>0.200836</td>
<td>1.385146</td>
<td>0.365815</td>
<td>2.450882</td>
</tr>
<tr>
<td>( \Delta_k s_{t-5} )</td>
<td>0.128048</td>
<td>1.095548</td>
<td>0.170983</td>
<td>1.421079</td>
</tr>
<tr>
<td>( \Delta_k f_{t-1} )</td>
<td>0.1277</td>
<td>0.977865</td>
<td>-0.538835</td>
<td>-4.008222</td>
</tr>
<tr>
<td>( \Delta_k f_{t-2} )</td>
<td>0.063916</td>
<td>0.431038</td>
<td>-0.407307</td>
<td>-2.668319</td>
</tr>
<tr>
<td>( \Delta_k f_{t-3} )</td>
<td>-0.23767</td>
<td>-1.575203</td>
<td>-0.565443</td>
<td>-3.640486</td>
</tr>
<tr>
<td>( \Delta_k f_{t-4} )</td>
<td>-0.171009</td>
<td>-1.206185</td>
<td>-0.330459</td>
<td>-2.264228</td>
</tr>
<tr>
<td>( \Delta_k f_{t-5} )</td>
<td>-0.179283</td>
<td>-1.576722</td>
<td>-0.22708</td>
<td>-1.940004</td>
</tr>
<tr>
<td>( k_t )</td>
<td>0.000715</td>
<td>1.649874 ( ^e )</td>
<td>0.000694</td>
<td>1.555750</td>
</tr>
<tr>
<td>( k_{t-1} )</td>
<td>0.000195</td>
<td>0.484479</td>
<td>0.000145</td>
<td>0.351404</td>
</tr>
<tr>
<td>( k_{t-2} )</td>
<td>-0.000325</td>
<td>-0.778256</td>
<td>-0.000357</td>
<td>-0.830378</td>
</tr>
<tr>
<td>( k_{t-3} )</td>
<td>0.000694</td>
<td>1.652878 ( ^e )</td>
<td>0.000664</td>
<td>1.536251</td>
</tr>
<tr>
<td>( k_{t-4} )</td>
<td>0.000342</td>
<td>0.843698</td>
<td>0.000421</td>
<td>1.008335</td>
</tr>
<tr>
<td>( k_{t-5} )</td>
<td>-0.000145</td>
<td>-0.332817</td>
<td>-4.79E-05</td>
<td>-0.106963</td>
</tr>
<tr>
<td>( z_t )</td>
<td>-1.69E-05</td>
<td>-0.689417</td>
<td>3.08E-05</td>
<td>1.218574</td>
</tr>
<tr>
<td>( z_{t-1} )</td>
<td>1.26E-06</td>
<td>0.049237</td>
<td>4.02E-05</td>
<td>1.524904</td>
</tr>
<tr>
<td>( z_{t-2} )</td>
<td>-1.80E-05</td>
<td>-0.701283</td>
<td>9.47E-06</td>
<td>0.358415</td>
</tr>
<tr>
<td>( z_{t-3} )</td>
<td>-1.16E-05</td>
<td>-0.453169</td>
<td>8.69E-06</td>
<td>0.32972</td>
</tr>
<tr>
<td>( z_{t-4} )</td>
<td>4.06E-05</td>
<td>1.602942</td>
<td>4.91E-05</td>
<td>1.88313 ( ^e )</td>
</tr>
<tr>
<td>( z_{t-5} )</td>
<td>-1.40E-06</td>
<td>-0.056877</td>
<td>-2.36E-06</td>
<td>-0.093549</td>
</tr>
<tr>
<td>( b_{t-k} )</td>
<td>0.021335</td>
<td>0.267169</td>
<td>-0.064238</td>
<td>-0.78143</td>
</tr>
<tr>
<td>( T_o - t + k_t )</td>
<td>1.26E-05</td>
<td>1.004284</td>
<td>9.38E-06</td>
<td>0.723724</td>
</tr>
</tbody>
</table>

Notes:

\( ^a \) significant at the 1% level

\( ^b \) significant at 5% level

\( ^c \) significant at 10% level.
4.4.2 Wald test of the cost of carry model

A Wald test was used to test the restrictions imposed by the cost of carry model on equations (4.31a) and (4.31b) for 28/10/1986–17/10/1997 and on equations (4.8a) and (4.8b) for 27/10/1997–30/12/2005 (see Table 4.5 for the restrictions). The Wald test results are reported in Table 4.12. Joint restrictions imposed by the cost of carry model are decisively rejected for 28/10/1986–17/10/1997, but they are jointly accepted at the 5% level for 27/10/1997–30/12/2005. The implication is that the no-arbitrage cost of carry relationship tends to hold in the second sub-sample. However, rejection of the cost of carry relationship in the first sub-sample does not necessarily imply the existence of profitable arbitrage opportunities, because of transaction costs and arbitrage risks. Furthermore, the observed violation of the cost of carry relationship could be illusory, because the reported FTSE100 index is a quote-based value and is not necessarily a real tradable price.

Individual restrictions can be tested separately, allowing possible reasons to be diagnosed for rejecting the cost of carry relationship in the first sub-sample and accepting it in the second sub-sample. Consistent with the joint test results, most restrictions are decisively rejected individually for 28/10/1986–17/10/1997. However, only one sub-restriction under restriction 2 is rejected while all other restrictions are accepted at a high level of confidence for 27/10/1997–30/12/2005. As discussed earlier, different restrictions have different implications. Short-run dynamics of spot and futures prices are represented by restriction 2 and long-run relationship between spot and futures prices are represented by restrictions 4, 5 and 6 jointly. Given that 3 out of 4 sub-restrictions under restriction 2 are accepted for the period 28/10/1986–17/10/1997, while other restrictions (with the exception of one sub-restriction under restriction 3) are rejected, it is possible to conclude that while the two markets tend to function properly in the short run, they have failed to maintain the long-run equilibrium relationship that is implied by the cost-of-carry model used here.

In summary, the results of the Wald test of the cost of carry restrictions over the two sub-samples are consistent with the previous findings: during 28/10/1986–17/10/1997 there is evidence of divergence from long-run equilibrium that is found to be corrected in the futures market, and indeed the cost of carry relationship is rejected; during 27/10/1997–30/12/2005 there is no obvious evidence of divergence from long-run
equilibrium or that any divergence was rapidly eliminated by short-run price adjustments rather than by the actions of arbitrageurs, and indeed the cost of carry relationship is accepted. The result suggests a progressive maturation of the market, in the sense that the FTSE100 spot and futures markets have become more efficiently linked to each other in recent history, probably due to the employment of more efficient trading systems by both LSE and LIFFE.
Table 4.12 Wald test of the restrictions imposed by the cost of carry model on the selected VECM

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint tests (restrictions 1<del>7 from Table 4.5 for 28/10/1986 – 17/10/1997 and restrictions 1</del>6 from Table 4.5 for 27/10/1997 – 30/12/2005)</td>
<td>x²-stat. 394.3941</td>
<td>x²-stat. 26.4948</td>
</tr>
<tr>
<td></td>
<td>P-value 0.0000 *</td>
<td>P-value 0.0890</td>
</tr>
<tr>
<td><strong>Restriction 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_0 - \alpha_0 = 0 )</td>
<td>x²-stat. 10.2024</td>
<td>x²-stat. 0.0005</td>
</tr>
<tr>
<td></td>
<td>P-value 0.0014 *</td>
<td>P-value 0.9816</td>
</tr>
<tr>
<td><strong>Restriction 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{ij} - \alpha_{ij} = -(\beta_{2j} - \alpha_{2j}) \quad \forall j )</td>
<td>x²-stat. 25.5236</td>
<td>x²-stat. 10.1800</td>
</tr>
<tr>
<td></td>
<td>P-value 0.0000 *</td>
<td>P-value 0.0014 *</td>
</tr>
<tr>
<td></td>
<td>x²-stat. 0.2294</td>
<td>x²-stat. 0.2926</td>
</tr>
<tr>
<td></td>
<td>P-value 0.6319</td>
<td>P-value 0.5886</td>
</tr>
<tr>
<td></td>
<td>x²-stat. 0.0528</td>
<td>x²-stat. 0.0207</td>
</tr>
<tr>
<td></td>
<td>P-value 0.8182</td>
<td>P-value 0.8856</td>
</tr>
<tr>
<td></td>
<td>x²-stat. 0.3577</td>
<td>x²-stat. 1.6490</td>
</tr>
<tr>
<td></td>
<td>P-value 0.5498</td>
<td>x²-stat. 1.2807</td>
</tr>
<tr>
<td></td>
<td>P-value 0.1991</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>P-value 0.2578</td>
</tr>
</tbody>
</table>

*Note: * denotes rejection of the null hypothesis, that is, the relevant cost of carry restriction is rejected.
Table 4.12 Wald test of the restrictions imposed by the cost of carry model on the selected VECM (continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{30} - \alpha_{30} = -\frac{\beta_{31} - \alpha_{31}}{\beta_{21} - \alpha_{21}} )</td>
<td>( x^2)-stat. 7.3984</td>
<td>( x^2)-stat. 0.2107</td>
</tr>
<tr>
<td>( \beta_{30} - \alpha_{30} = -\frac{\beta_{32} - \alpha_{32}}{\beta_{22} - \alpha_{22}} )</td>
<td>( x^2)-stat. 10.3658</td>
<td>( x^2)-stat. 0.0769</td>
</tr>
<tr>
<td>( \beta_{30} - \alpha_{30} = -\frac{\beta_{33} - \alpha_{33}}{\beta_{23} - \alpha_{23}} )</td>
<td>( x^2)-stat. 5.8415</td>
<td>( x^2)-stat. 0.0852</td>
</tr>
<tr>
<td>( \beta_{30} - \alpha_{30} = -\frac{\beta_{34} - \alpha_{34}}{\beta_{24} - \alpha_{24}} )</td>
<td>( x^2)-stat. 0.8776</td>
<td>( x^2)-stat. 1.0484</td>
</tr>
<tr>
<td>( \beta_{30} - \alpha_{30} = -\frac{\beta_{35} - \alpha_{35}}{\beta_{25} - \alpha_{25}} )</td>
<td>( x^2)-stat. 1.0513</td>
<td>( x^2)-stat. 0.3052</td>
</tr>
<tr>
<td>( \beta_{30} - \alpha_{30} = -\frac{\beta_{40} - \alpha_{40}}{\beta_{40} - \alpha_{40}} )</td>
<td>( x^2)-stat. 9.2842</td>
<td>( x^2)-stat. 0.1007</td>
</tr>
</tbody>
</table>

* denotes rejection of the null hypothesis, that is, the relevant cost of carry restriction is rejected.
Table 4.12 Wald test of the restrictions imposed by the cost of carry model on the selected VECM (continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x^2$-stat.</td>
<td>P-value</td>
</tr>
<tr>
<td><strong>Restriction 5</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{30} - \alpha_{30} = (\beta_{4j} - \alpha_{4j})/ (\beta_{2j} - \alpha_{2j})$</td>
<td>$\beta_{30} - \alpha_{30} = (\beta_{41} - \alpha_{41})/ (\beta_{21} - \alpha_{21})$</td>
<td>9.1941</td>
</tr>
<tr>
<td>$\forall j$</td>
<td>P-value</td>
<td>0.0024 *</td>
</tr>
<tr>
<td>$\beta_{30} - \alpha_{30} = (\beta_{42} - \alpha_{42})/ (\beta_{22} - \alpha_{22})$</td>
<td>$\beta_{30} - \alpha_{30} = (\beta_{43} - \alpha_{43})/ (\beta_{23} - \alpha_{23})$</td>
<td>7.9548</td>
</tr>
<tr>
<td>P-value</td>
<td>0.0048 *</td>
<td>P-value</td>
</tr>
<tr>
<td>$\beta_{30} - \alpha_{30} = (\beta_{44} - \alpha_{44})/ (\beta_{24} - \alpha_{24})$</td>
<td>$\beta_{30} - \alpha_{30} = (\beta_{45} - \alpha_{45})/ (\beta_{25} - \alpha_{25})$</td>
<td>7.6076</td>
</tr>
<tr>
<td>P-value</td>
<td>0.0058 *</td>
<td>P-value</td>
</tr>
<tr>
<td><strong>Restriction 6</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{30} - \alpha_{30} = (\beta_{6} - \alpha_{6})/ (\beta_{5} - \alpha_{5})$</td>
<td>$\beta_{30} - \alpha_{30} = (\beta_{5} - \alpha_{5})$</td>
<td>7.4398</td>
</tr>
<tr>
<td>P-value</td>
<td>0.0064 *</td>
<td>P-value</td>
</tr>
</tbody>
</table>

*Note:* * denotes rejection of the null hypothesis, that is, the relevant cost of carry restriction is rejected.
Table 4.12 Wald test of the restrictions imposed by the cost of carry model on the selected VECM (continued)

<table>
<thead>
<tr>
<th>Restrictions of the cost of carry</th>
<th>28/10/1986 – 17/10/1997</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x^2$-stat.</td>
</tr>
<tr>
<td>$\beta_{70} - \alpha_{70} = 0$</td>
<td>270.1983</td>
</tr>
<tr>
<td>$\beta_{71} - \alpha_{71} = 0$</td>
<td>3.4891</td>
</tr>
<tr>
<td>$\beta_{72} - \alpha_{72} = 0$</td>
<td>3.7493</td>
</tr>
<tr>
<td>$\beta_{73} - \alpha_{73} = 0$</td>
<td>36.8438</td>
</tr>
<tr>
<td>$\beta_{74} - \alpha_{74} = 0$</td>
<td>24.7206</td>
</tr>
</tbody>
</table>

**Restriction 7**

$\beta_{7j} - \alpha_{7j} = 0 \quad \forall j$

*Note:* * denotes rejection of the null hypothesis, that is, the relevant cost of carry restriction is rejected.
4.4.3 Causality and cointegration under accepted cost of carry restrictions

The Wald test above shows that some restrictions imposed by the cost of carry model on the selected VECM for FTSE100 spot and futures prices are accepted, especially during the second period within which almost all the cost of carry restrictions are accepted (except that one sub-restriction under restriction 2 is rejected). To further examine the price dynamics between FTSE100 spot and futures markets, equations (4.31a) and (4.31b) for 28/10/1986–17/10/1997 and equations (4.8a) and (4.8b) for 27/10/1997–30/12/2005 were again estimated, but with those accepted cost of carry restrictions (see Table 4.12) imposed during the estimation. The estimation of a restricted VECM should require fewer parameters to be estimated than that of the unrestricted VECM because some parameters can be replaced, according to the accepted relationships implied by the cost of carry restrictions, by combination of other parameters that already exist in the model. But this introduces a problem. For those parameters that are replaced by combinations of other parameters, it is difficult to know whether it is significant. To overcome this problem, two-step estimation procedure was employed. In the first step, all the parameters in the spot equation were maintained as before, but some parameters in the futures equation were replaced, according to the accepted relationships implied by the cost of carry restrictions, by combination of other parameters that already exist in the model. This ensures that both the value itself and the standard error are explicitly reported for all the parameters in the spot equation. Then in the second step, all the parameters in the futures equation were maintained as before, but some parameters in the spot equation were replaced, according to the accepted relationships implied by the cost of carry restrictions, by combination of other parameters that already exist in the model. Accordingly, this ensures that both the value itself and the standard error are explicitly reported for all the parameters in the futures equation (see Appendix 4.1 for details). Here we report the results from the first step estimation for the spot equation and the results from the second step estimation for the futures equation, though the two step estimations should give exactly the same results. The results of the restricted VECM are reported in Table 4.13 for the first period and Table 4.14 for the second period. It can be seen that the results are almost the same as before: both the estimated value and the significance level for individual parameters are very close under restricted and unrestricted VECM analysis, suggesting that the
restrictions derived from the cost of carry model on the selected VECM are reasonable. In particular, here the results again suggest that during the first period, there is divergence from equilibrium, which is corrected in the futures market; however, during the second period, there is no obvious evidence of divergence from equilibrium. This further confirms that the no-arbitrage cost of carry relationship is maintained well between FTSE100 spot and futures markets in the more recent period, implying more efficient market in this period. To avoid repetition, we would not analyse in detail all the parameters.
Table 4.13 SUR estimation of the restricted equations (4.31a) and (4.31b)

28/10/1986 – 17/10/1997

<table>
<thead>
<tr>
<th>Variables</th>
<th>$\Delta_k s_t$</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>$\Delta_k f_t$</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>-0.003427</td>
<td>-2.168167 $^b$</td>
<td></td>
<td>-0.00505</td>
<td>-2.96073 $^a$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k s_{t-1}$</td>
<td>-0.114615</td>
<td>-1.985257 $^b$</td>
<td></td>
<td>0.363616</td>
<td>5.432371 $^a$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k s_{t-2}$</td>
<td>0.049103</td>
<td>0.804245</td>
<td></td>
<td>0.328374</td>
<td>4.788052 $^a$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k s_{t-3}$</td>
<td>0.015928</td>
<td>0.268795</td>
<td></td>
<td>0.154239</td>
<td>2.318948 $^b$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k s_{t-4}$</td>
<td>0.076813</td>
<td>1.534884</td>
<td></td>
<td>0.142091</td>
<td>2.505099 $^b$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k f_{t-1}$</td>
<td>0.163749</td>
<td>3.219211 $^a$</td>
<td></td>
<td>-2.80E-01</td>
<td>-4.758298 $^a$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k f_{t-2}$</td>
<td>-0.031984</td>
<td>-0.589394</td>
<td></td>
<td>-0.311255</td>
<td>-4.970953 $^a$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k f_{t-3}$</td>
<td>-0.048120</td>
<td>-0.914202</td>
<td></td>
<td>-0.186431</td>
<td>-3.071911 $^a$</td>
<td></td>
</tr>
<tr>
<td>$\Delta_k f_{t-4}$</td>
<td>-0.020978</td>
<td>-0.460398</td>
<td></td>
<td>-8.63E-02</td>
<td>-1.63332</td>
<td></td>
</tr>
<tr>
<td>$k_t$</td>
<td>-5.31E-05</td>
<td>-0.211794</td>
<td></td>
<td>9.50E-05</td>
<td>0.332892</td>
<td></td>
</tr>
<tr>
<td>$k_{t-1}$</td>
<td>0.000702</td>
<td>2.582332 $^a$</td>
<td></td>
<td>0.000974</td>
<td>3.20452 $^a$</td>
<td></td>
</tr>
<tr>
<td>$k_{t-2}$</td>
<td>0.000538</td>
<td>1.949837 $^c$</td>
<td></td>
<td>0.000865</td>
<td>2.811016 $^a$</td>
<td></td>
</tr>
<tr>
<td>$k_{t-3}$</td>
<td>0.000747</td>
<td>2.757167 $^a$</td>
<td></td>
<td>0.000982</td>
<td>3.249083 $^a$</td>
<td></td>
</tr>
<tr>
<td>$k_{t-4}$</td>
<td>0.000653</td>
<td>2.671930 $^a$</td>
<td></td>
<td>6.62E-04</td>
<td>2.709294 $^a$</td>
<td></td>
</tr>
<tr>
<td>$z_t$</td>
<td>-2.47E-06</td>
<td>-0.168648</td>
<td></td>
<td>0.000115</td>
<td>6.739717 $^a$</td>
<td></td>
</tr>
<tr>
<td>$z_{t-1}$</td>
<td>1.75E-06</td>
<td>0.110563</td>
<td></td>
<td>5.57E-05</td>
<td>3.020877 $^a$</td>
<td></td>
</tr>
<tr>
<td>$z_{t-2}$</td>
<td>1.67E-06</td>
<td>0.104944</td>
<td></td>
<td>3.08E-05</td>
<td>1.663605 $^c$</td>
<td></td>
</tr>
<tr>
<td>$z_{t-3}$</td>
<td>1.43E-05</td>
<td>0.909897</td>
<td></td>
<td>3.12E-05</td>
<td>1.697778 $^c$</td>
<td></td>
</tr>
<tr>
<td>$z_{t-4}$</td>
<td>-4.78E-07</td>
<td>-0.031096</td>
<td></td>
<td>1.34E-05</td>
<td>0.749079</td>
<td></td>
</tr>
<tr>
<td>$b_{t-k_t}$</td>
<td>0.003468</td>
<td>0.113521</td>
<td></td>
<td>-0.125461</td>
<td>-3.534291 $^a$</td>
<td></td>
</tr>
<tr>
<td>$T_a - t + k_t$</td>
<td>2.83E-06</td>
<td>0.365857</td>
<td></td>
<td>1.48E-05</td>
<td>1.639284</td>
<td></td>
</tr>
<tr>
<td>$DM$</td>
<td>-0.113187</td>
<td>-13.52974 $^a$</td>
<td></td>
<td>-0.169876</td>
<td>-17.39867 $^a$</td>
<td></td>
</tr>
<tr>
<td>$DM(-1)$</td>
<td>-0.120004</td>
<td>-13.86985 $^a$</td>
<td></td>
<td>-0.120004</td>
<td>-13.86985 $^a$</td>
<td></td>
</tr>
<tr>
<td>$DM(-2)$</td>
<td>0.073015</td>
<td>8.174501 $^a$</td>
<td></td>
<td>0.073015</td>
<td>8.174501 $^a$</td>
<td></td>
</tr>
<tr>
<td>$DM(-3)$</td>
<td>-0.067002</td>
<td>-7.253870 $^a$</td>
<td></td>
<td>-0.089437</td>
<td>-8.400957 $^a$</td>
<td></td>
</tr>
<tr>
<td>$DM(-4)$</td>
<td>-0.013331</td>
<td>-1.428302</td>
<td></td>
<td>-0.032942</td>
<td>-3.07363 $^a$</td>
<td></td>
</tr>
</tbody>
</table>

Notes:

* significant at the 1% level

$^b$ significant at the 5% level

$c$ significant at the 10% level.
Table 4.14 SUR estimation of the restricted equations (4.8a) and (4.8b)

27/10/1997 – 30/12/2005

<table>
<thead>
<tr>
<th>Variables</th>
<th>$\Delta k s_t$</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>$\Delta k f_t$</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.002644</td>
<td>-1.098977</td>
<td>-0.002644</td>
<td>-1.098977</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k s_{t-1}$</td>
<td>-0.137221</td>
<td>-1.027612</td>
<td>0.534737</td>
<td>3.893101*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k s_{t-2}$</td>
<td>-0.132683</td>
<td>-0.874368</td>
<td>0.32821</td>
<td>2.105849 b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k s_{t-3}$</td>
<td>0.136625</td>
<td>0.883823</td>
<td>0.452075</td>
<td>2.848021*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k s_{t-4}$</td>
<td>0.195661</td>
<td>1.349549</td>
<td>0.341862</td>
<td>2.296038 b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k s_{t-5}$</td>
<td>0.127151</td>
<td>1.087966</td>
<td>0.166829</td>
<td>1.390419</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k f_{t-1}$</td>
<td>0.129762</td>
<td>0.993644</td>
<td>-5.29E-01</td>
<td>-3.940113 a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k f_{t-2}$</td>
<td>0.066763</td>
<td>0.450239</td>
<td>-0.39413</td>
<td>-2.584715 a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k f_{t-3}$</td>
<td>-0.234273</td>
<td>-1.552712</td>
<td>-0.549723</td>
<td>-3.543673 a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k f_{t-4}$</td>
<td>-0.167357</td>
<td>-1.180441</td>
<td>-0.313558</td>
<td>-2.151035 b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta k f_{t-5}$</td>
<td>-0.177045</td>
<td>-1.557064</td>
<td>-0.216724</td>
<td>-1.853865 c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_t$</td>
<td>0.000707</td>
<td>1.634644</td>
<td>6.59E-04</td>
<td>1.522443</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{t-1}$</td>
<td>0.0002</td>
<td>0.497081</td>
<td>0.000168</td>
<td>0.417396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{t-2}$</td>
<td>-0.000322</td>
<td>-0.773014</td>
<td>-3.45E-04</td>
<td>-0.826648</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{t-3}$</td>
<td>0.000698</td>
<td>1.664874 c</td>
<td>6.83E-04</td>
<td>1.628365</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{t-4}$</td>
<td>0.000319</td>
<td>0.786435</td>
<td>3.12E-04</td>
<td>0.768919</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{t-5}$</td>
<td>-0.000172</td>
<td>-0.39598</td>
<td>-1.74E-04</td>
<td>-0.400411</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_t$</td>
<td>-1.67E-05</td>
<td>-0.680523</td>
<td>3.18E-05</td>
<td>1.280823</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{t-1}$</td>
<td>-6.57E-07</td>
<td>-0.025665</td>
<td>3.13E-05</td>
<td>1.215057</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{t-2}$</td>
<td>-1.94E-05</td>
<td>-0.75698</td>
<td>2.96E-06</td>
<td>0.114746</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{t-3}$</td>
<td>-1.30E-05</td>
<td>-0.507513</td>
<td>2.32E-06</td>
<td>0.090538</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{t-4}$</td>
<td>4.02E-05</td>
<td>1.589688</td>
<td>4.73E-05</td>
<td>1.865434 c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{t-5}$</td>
<td>-5.99E-07</td>
<td>-0.024433</td>
<td>1.33E-06</td>
<td>0.053997</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_{t-k}$</td>
<td>0.020817</td>
<td>0.260679</td>
<td>-0.066636</td>
<td>-0.811159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_{\alpha-t+k}$</td>
<td>1.47E-05</td>
<td>1.170159</td>
<td>1.90E-05</td>
<td>1.502916</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

- * significant at the 1% level
- b significant at the 5% level
- c significant at the 10% level.
4.5 Conclusions

This chapter examines short-run causality and long-run cointegration between FTSE100 index and futures prices and tests the validity of the cost of carry model for pricing FTSE100 futures contracts within a single VECM framework. The empirical study was performed in two sub-samples. The first sub-sample covers 28/10/1986–17/10/1997, when all FTSE100 securities were traded through a dealership system and FTSE100 futures were traded through 'open outcry'. The second sub-sample covers 27/10/1997–30/12/2005, when all FTSE100 securities were traded through an order-driven trading system, and FTSE100 futures were traded on an electronic platform (but only after 30/11/1998).

For the first sub-period, 28/10/1986–17/10/1997, there is bidirectional causality between FTSE100 spot and futures, with the causality from spot to futures being much stronger than the reverse. The FTSE100 futures price changes also have a stronger mean-reverting tendency than the spot price changes. The cost of carry model seems to be invalid for pricing FTSE100 futures, as indicated by decisive rejections of the cost of carry restrictions imposed on the finally selected VECM framework. Accordingly, there is evidence of divergence from the equilibrium defined by the cost of carry model. This divergence is found to be mainly corrected in the futures market. These findings somehow imply a form of inefficiency of the FTSE100 futures market during this period. To have more supportive empirical evidence, these issues are going to be further explored in later chapters with more factors (e.g. conditional variance, regime-dependent price dynamics) taken into account. For the moment, possible explanations for these findings are summarized below.

First, our finding regarding lead-lag relationships between stock index and futures price seems to be controversial with what is generally documented in the literature. That is, while it is widely reported that index futures prices tend to lead the underlying cash indices, we find that the causality from FTSE100 cash price to futures price is much stronger than the reverse. As analysed above, since the FTSE100 index is a quote-based index, it should reflect information more quickly than the transaction-based futures price because market makers (or market participants) can update their quotes for the index component shares immediately in response to any new information, whereas information can only be impounded in the futures price after transactions have taken
place in the futures market. Our finding provides support to the argument of Yadav and Pope (1990, 1994) that studies of lead-lag relationship between stock index and futures prices using transaction price based indices may suffer from measurement errors induced by infrequent trading problems. The lagged response of the FTSE100 futures price to new information could have caused temporary divergence from equilibrium between the two markets. Subsequent adjustment in the futures market, reflected in strong evidence of mean-reversion in futures price changes and the significant error correction term in the futures equation, was therefore necessary to drive the futures price and the diverged relationship between the two markets to a new equilibrium. Another possible reason is that due to the ease of trading in futures markets, the futures price may have sometimes overreacted to new information, causing divergence from equilibrium of both the futures price and the relationship between the two markets. Subsequent adjustment in the futures market then occurred to correct both the overreaction and the divergence from equilibrium, reflected in the strong evidence of mean-reversion in futures price changes and in the significant error correction term in the futures equation. Another possibility is that the mean-reversion in futures price changes could reflect bid-ask bounce in futures prices.

The evidence on short-run dynamics for FTSE100 spot and futures prices documented in the second sub-period is slightly different from that of the first sub-period. During the second sub-period 27/10/1997–30/12/2005, there is unidirectional causality from spot to futures, further indicating that the quote-based FTSE100 index can reflect information more quickly than the transaction-based FTSE100 futures price. Only futures price changes are significantly mean-reverting, due to either bid-ask bounce or overreaction and subsequent correction in futures prices. On the other hand, the evidence on the long-run equilibrium relationship during the second sub-sample is quite the opposite of that found during the first sub-sample. No evidence of divergence from the cost of carry relationship is found, either because there was sufficient continual adjustment in one or both markets, or because such divergence was rapidly eliminated by short-run price adjustments rather than by the actions of arbitrageurs over this time period. Accordingly, the cost of carry model is found to be valid for pricing FTSE100 futures, indicated by acceptance of the cost of carry restrictions at a 5% level of significance.
The different empirical evidence regarding the long-run cost of carry relationship over the two sub-samples suggests that the two markets became more efficiently related to each other, and hence that fewer profitable arbitrage opportunities were present in the more recent sub-sample period. This can be explained by more efficient trading systems and improved trading facilities in the second sub-period, when arbitrage transactions should have been much easier and cheaper than before. There are at least three reasons for this. First, on 20/10/1997 the LSE changed its trading system for the most liquid shares (including all FTSE 100 securities) from SEAQ (a dealership system) to SETS (a fully electronic, order-driven trading system). Transaction costs have declined under the new system and the automated matching of buy and sell orders makes the order-driven market more transparent and efficient for participants. Second, on 30/11/1998, trading of FTSE100 futures at LIFFE was transformed from 'open outcry' to LIFFE CONNECT, an electronic platform. FTSE100 futures traders experience several benefits from this new platform. For example, it allows for quick reaction to market changes. Traders can watch futures prices change in real time around the world and can execute trades within milliseconds. Electronic trading is also likely to generate fewer errors in orders than open outcry trading. This lowers the costs of buying and selling FTSE100 futures in LIFFE. Furthermore, in January 2002 the purchase of LIFFE by Euronext was completed. Euronext.liffe creates a single market for derivatives by replacing multiple trading venues with a single market supported by a state-of-the-art electronic trading system, which reduces costs for both Euronext.liffe itself and its customers, and makes cross-border trading easier and cheaper. Third, Barclays Global Investors (BGI) introduced the exchange traded fund (ETF) in April 2000 for the FTSE100, to track the FTSE 100 index. Among many other advantages, ETF purchases made on the LSE are not directly subject to stamp duty and short sales of ETF are possible. This makes purchase and sale of the FTSE100 index portfolio in the spot market much quicker, easier and cheaper than before. The improved trading facilities and lower transaction costs in both spot and futures markets during the sample period 27/10/1997–30/12/2005 have therefore made arbitrage transactions cheaper and easier than before and allowed more investors to make arbitrage transactions between the FTSE100 index and futures, leading to a more efficient link between FTSE100 spot and futures markets.
Appendix 4.1 Estimation of the restricted VECM for spot and futures returns

In the first step, all parameters in the spot equation remained unchanged while some parameters in the futures equation were replaced by combinations of existing parameters according to accepted relationships implied by the cost of carry model (see Table 4.12 for details). Therefore, for the first sub-sample, a system of equations as follows was estimated.

\[
\begin{align*}
\Delta_k s_t &= \alpha_0 + \alpha_{11} \Delta_k s_{t-1} + \alpha_{12} \Delta_k s_{t-2} + \alpha_{13} \Delta_k s_{t-3} + \alpha_{14} \Delta_k s_{t-4} \\
&+ \alpha_{21} \Delta_k f_{t-1} + \alpha_{22} \Delta_k f_{t-2} + \alpha_{23} \Delta_k f_{t-3} + \alpha_{24} \Delta_k f_{t-4} \\
&+ \alpha_{30} k_t + \alpha_{31} k_{t-1} + \alpha_{32} k_{t-2} + \alpha_{33} k_{t-3} + \alpha_{34} k_{t-4} \\
&+ \alpha_{40} z_t + \alpha_{41} z_{t-1} + \alpha_{42} z_{t-2} + \alpha_{43} z_{t-3} + \alpha_{44} z_{t-4} \\
&+ \alpha_{50} \pi_t + \alpha_{51} \pi_{t-1} + \alpha_{52} \pi_{t-2} + \alpha_{53} \pi_{t-3} + \alpha_{54} \pi_{t-4} + \epsilon_{1t}
\end{align*}
\]

In the second step, all parameters in the futures equation remained unchanged while some parameters in the spot equation were replaced by combinations of existing parameters according to accepted relationships implied by the cost of carry model (see Table 4.12 for details). That is, a system of equations as follows was estimated.

\[
\begin{align*}
\Delta_k f_t &= \beta_0 + \beta_{11} \Delta_k s_{t-1} + (\alpha_{12} - \beta_{22} + \alpha_{22}) \Delta_k s_{t-2} + (\alpha_{13} - \beta_{23} + \alpha_{23}) \Delta_k s_{t-3} \\
&+ (\alpha_{14} - \beta_{24} + \alpha_{24}) \Delta_k s_{t-4} + \beta_{21} \Delta_k f_{t-1} + \beta_{22} \Delta_k f_{t-2} + \beta_{23} \Delta_k f_{t-3} + \beta_{24} \Delta_k f_{t-4} \\
&+ \beta_{30} k_t + \beta_{31} k_{t-1} + \beta_{32} k_{t-2} + \beta_{33} k_{t-3} + (\alpha_{34} - (\beta_{30} - \alpha_{30}) (\beta_{24} - \alpha_{24})) k_{t-4} \\
&+ \beta_{40} z_t + \beta_{41} z_{t-1} + \beta_{42} z_{t-2} + \beta_{43} z_{t-3} + \beta_{44} z_{t-4} \\
&+ \beta_{50} \pi_t + \beta_{51} \pi_{t-1} + \beta_{52} \pi_{t-2} + \beta_{53} \pi_{t-3} + \beta_{54} \pi_{t-4} + \epsilon_{2t}
\end{align*}
\]
\[
\Delta_k s_t = a_0 + a_{11} \Delta_k s_{t-1} + (\beta_{12} + \beta_{22} - \alpha_{22}) \Delta_k s_{t-2} + (\beta_{13} + \beta_{23} - \alpha_{23}) \Delta_k s_{t-3} \\
+ (\beta_{14} + \beta_{24} - \alpha_{24}) \Delta_k s_{t-4} + \alpha_{21} \Delta k f_{t-1} + \alpha_{22} \Delta k f_{t-2} + \alpha_{23} \Delta k f_{t-3} + \alpha_{24} \Delta k f_{t-4} \\
+ \alpha_{30} k_{t} + \alpha_{31} k_{t-1} + \alpha_{32} k_{t-2} + \alpha_{33} k_{t-3} + \alpha_{34} (\beta_{34} + (\beta_{30} - \alpha_{30})(\beta_{24} - \alpha_{24})) k_{t-4} \\
+ \alpha_{40} z_{t} + \alpha_{41} z_{t-1} + \alpha_{42} z_{t-2} + \alpha_{43} z_{t-3} + \alpha_{44} z_{t-4} \\
+ \alpha_{5} b_{t-k} + \alpha_{6} (T_g - t + k_{t}) \\
+ \alpha_{70} D M_t + \beta_{71} D M_{t-1} + \beta_{72} D M_{t-2} + \alpha_{73} D M_{t-3} + \alpha_{74} D M_{t-4} + e_{1t}
\]
\[
\Delta_k f_t = \beta_0 + \beta_{11} \Delta k s_{t-1} + \beta_{12} \Delta k s_{t-2} + \beta_{13} \Delta k s_{t-3} + \beta_{14} \Delta k s_{t-4} \\
+ \beta_{21} \Delta k f_{t-1} + \beta_{22} \Delta k f_{t-2} + \beta_{23} \Delta k f_{t-3} + \beta_{24} \Delta k f_{t-4} \\
+ \beta_{30} k_{t} + \beta_{31} k_{t-1} + \beta_{32} k_{t-2} + \beta_{33} k_{t-3} + \beta_{34} k_{t-4} \\
+ \beta_{40} z_{t} + \beta_{41} z_{t-1} + \beta_{42} z_{t-2} + \beta_{43} z_{t-3} + \beta_{44} z_{t-4} \\
+ \beta_{5} b_{t-k} + \beta_{6} (T_g - t + k_{t}) \\
+ \beta_{70} D M_t + \beta_{71} D M_{t-1} + \beta_{72} D M_{t-2} + \beta_{73} D M_{t-3} + \beta_{74} D M_{t-4} + e_{2t}
\]

For the first sub-sample, the estimation results are reported in Table A1.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>Variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_0)</td>
<td>-0.003427</td>
<td>-2.168167 b</td>
<td>(a_0)</td>
<td>-0.003427</td>
<td>-2.168167 b</td>
</tr>
<tr>
<td>(a_{11})</td>
<td>-0.114615</td>
<td>-1.985257 b</td>
<td>(a_{11})</td>
<td>-0.114615</td>
<td>-1.985257 b</td>
</tr>
<tr>
<td>(a_{12})</td>
<td>0.049103</td>
<td>0.804245</td>
<td>(a_{21})</td>
<td>0.163749</td>
<td>3.219211 a</td>
</tr>
<tr>
<td>(a_{13})</td>
<td>0.015928</td>
<td>0.268795</td>
<td>(a_{22})</td>
<td>-0.031984</td>
<td>-0.589394</td>
</tr>
<tr>
<td>(a_{14})</td>
<td>0.076813</td>
<td>1.534884</td>
<td>(a_{23})</td>
<td>-0.048120</td>
<td>-0.914202</td>
</tr>
<tr>
<td>(a_{21})</td>
<td>0.163749</td>
<td>3.219211 a</td>
<td>(a_{24})</td>
<td>-0.020978</td>
<td>-0.460398</td>
</tr>
<tr>
<td>(a_{22})</td>
<td>-0.031984</td>
<td>-0.589394</td>
<td>(a_{30})</td>
<td>-5.31E-05</td>
<td>-0.211794</td>
</tr>
<tr>
<td>(a_{23})</td>
<td>-0.048120</td>
<td>-0.914202</td>
<td>(a_{31})</td>
<td>0.000702</td>
<td>2.582332 a</td>
</tr>
<tr>
<td>(a_{24})</td>
<td>-0.020978</td>
<td>-0.460398</td>
<td>(a_{32})</td>
<td>0.000538</td>
<td>1.949837 c</td>
</tr>
<tr>
<td>(a_{30})</td>
<td>-5.31E-05</td>
<td>-0.211794</td>
<td>(a_{33})</td>
<td>0.000747</td>
<td>2.757167 a</td>
</tr>
<tr>
<td>(a_{31})</td>
<td>0.000702</td>
<td>2.582332 a</td>
<td>(a_{40})</td>
<td>-2.47E-06</td>
<td>-0.168648</td>
</tr>
<tr>
<td>(a_{32})</td>
<td>0.000538</td>
<td>1.949837 c</td>
<td>(a_{41})</td>
<td>1.75E-06</td>
<td>0.110563</td>
</tr>
<tr>
<td>(a_{33})</td>
<td>0.000747</td>
<td>2.757167 a</td>
<td>(a_{42})</td>
<td>1.67E-06</td>
<td>0.104944</td>
</tr>
<tr>
<td>(a_{34})</td>
<td>0.000653</td>
<td>2.671930 a</td>
<td>(a_{43})</td>
<td>1.43E-05</td>
<td>0.909897</td>
</tr>
<tr>
<td>(a_{40})</td>
<td>-2.47E-06</td>
<td>-0.168648</td>
<td>(a_{44})</td>
<td>-4.78E-07</td>
<td>-0.031096</td>
</tr>
<tr>
<td>(a_{41})</td>
<td>1.75E-06</td>
<td>0.110563</td>
<td>(a_5)</td>
<td>0.003468</td>
<td>0.113521</td>
</tr>
<tr>
<td>(a_{42})</td>
<td>1.67E-06</td>
<td>0.104944</td>
<td>(a_6)</td>
<td>2.83E-06</td>
<td>0.365857</td>
</tr>
<tr>
<td>(a_{43})</td>
<td>1.43E-05</td>
<td>0.909897</td>
<td>(a_{70})</td>
<td>-0.113187</td>
<td>-13.52974 a</td>
</tr>
</tbody>
</table>

Table A1 SUR estimation of the restricted VECM, 28/10/1986 – 17/10/1997
From Table A1 we can see that for the same parameters, the first step estimation gives exactly the same results as the second step estimation. Therefore, the results from the first step estimation for the spot equation and the results from the second step estimation for the futures equation can be used jointly to report both the value and the t-statistic for each parameter of the system.
Similarly, for the second sub-sample, in the first step, all parameters in the spot equation remained unchanged while some parameters in the futures equation were replaced by combinations of existing parameters according to accepted relationships implied by the cost of carry model (see Table 4.12 for details). That is, a system of equations as follows was estimated.

\[ \Delta_k s_t = \alpha_0 + \alpha_{11} \Delta_k s_{t-1} + \alpha_{12} \Delta_k s_{t-2} + \alpha_{13} \Delta_k s_{t-3} + \alpha_{14} \Delta_k s_{t-4} + \alpha_{15} \Delta_k s_{t-5} + \alpha_{21} \Delta_k f_{t-1} + \alpha_{22} \Delta_k f_{t-2} + \alpha_{23} \Delta_k f_{t-3} + \alpha_{24} \Delta_k f_{t-4} + \alpha_{25} \Delta_k f_{t-5} + \alpha_{30} k_t + \alpha_{31} k_{t-1} + \alpha_{32} k_{t-2} + \alpha_{33} k_{t-3} + \alpha_{34} k_{t-4} + \alpha_{35} k_{t-5} + \alpha_{40} z_t + \alpha_{41} z_{t-1} + \alpha_{42} z_{t-2} + \alpha_{43} z_{t-3} + \alpha_{44} z_{t-4} + \alpha_{45} z_{t-5} + \alpha_5 b_{t-k} + \alpha_6 (T_g - t + k_t) + \epsilon_{1t} \]

\[ \Delta_k f_t = \beta_0 + \beta_{11} \Delta_k s_{t-1} + (\alpha_{12} - \beta_{22} + \alpha_{22}) \Delta_k s_{t-2} + (\alpha_{13} - \beta_{23} + \alpha_{23}) \Delta_k s_{t-3} + (\alpha_{14} - \beta_{24} + \alpha_{24}) \Delta_k s_{t-4} + (\alpha_{15} - \beta_{25} + \alpha_{25}) \Delta_k s_{t-5} + \beta_{21} \Delta_k f_{t-1} + \beta_{22} \Delta_k f_{t-2} + \beta_{23} \Delta_k f_{t-3} + \beta_{24} \Delta_k f_{t-4} + \beta_{25} \Delta_k f_{t-5} + \beta_{30} k_t + (\alpha_{31} - (\beta_{30} - \alpha_{30})(\beta_{21} - \alpha_{21})) k_{t-1} + (\alpha_{32} - (\beta_{30} - \alpha_{30})(\beta_{22} - \alpha_{22})) k_{t-2} + (\alpha_{33} - (\beta_{30} - \alpha_{30})(\beta_{23} - \alpha_{23})) k_{t-3} + (\alpha_{34} - (\beta_{30} - \alpha_{30})(\beta_{24} - \alpha_{24})) k_{t-4} + (\alpha_{35} - (\beta_{30} - \alpha_{30})(\beta_{25} - \alpha_{25})) k_{t-5} + (\alpha_{40} - \beta_{30} + \alpha_{30}) z_t + (\alpha_{41} + (\beta_{30} - \alpha_{30})(\beta_{21} - \alpha_{21})) z_{t-1} + (\alpha_{42} + (\beta_{30} - \alpha_{30})(\beta_{22} - \alpha_{22})) z_{t-2} + (\alpha_{43} + (\beta_{30} - \alpha_{30})(\beta_{23} - \alpha_{23})) z_{t-3} + (\alpha_{44} + (\beta_{30} - \alpha_{30})(\beta_{24} - \alpha_{24})) z_{t-4} + (\alpha_{45} + (\beta_{30} - \alpha_{30})(\beta_{25} - \alpha_{25})) z_{t-5} + \beta_5 b_{t-k} + ((\beta_{30} - \alpha_{30})(\beta_5 - \alpha_5) + \alpha_6)(T_g - t + k_t) + \epsilon_{2t} \]

In the second step, all parameters in the futures equation remained unchanged while some parameters in the spot equation were replaced by combinations of existing parameters according to accepted relationships implied by the cost of carry model (see Table 4.12 for details). That is, a system of equations as follows was estimated.
\[ \Delta_t s_t = \beta_0 + \alpha_{11}\Delta_t s_{t-1} + (\beta_{12} + \beta_{22} - \alpha_{22})\Delta_t s_{t-2} + (\beta_{13} + \beta_{23} - \alpha_{23})\Delta_t s_{t-3} \\
+ (\beta_{14} + \beta_{24} - \alpha_{24})\Delta_t s_{t-4} + (\beta_{15} + \beta_{25} - \alpha_{25})\Delta_t s_{t-5} \\
+ \alpha_{30}k_t + (\beta_{31} + (\beta_{21} - \alpha_{21}))k_{t-1} \\
+ (\beta_{32} + (\beta_{22} - \alpha_{22}))k_{t-2} \\
+ (\beta_{33} + (\beta_{23} - \alpha_{23}))k_{t-3} \\
+ (\beta_{34} + (\beta_{24} - \alpha_{24}))k_{t-4} \\
+ (\beta_{35} + (\beta_{25} - \alpha_{25}))k_{t-5} \\
+ (\beta_{30} + \beta_{40} - \alpha_{30})z_t \\
+ (\beta_{41} - (\beta_{21} - \alpha_{21}))z_{t-1} \\
+ (\beta_{42} - (\beta_{22} - \alpha_{22}))z_{t-2} \\
+ (\beta_{43} - (\beta_{23} - \alpha_{23}))z_{t-3} \\
+ (\beta_{44} - (\beta_{24} - \alpha_{24}))z_{t-4} \\
+ (\beta_{45} - (\beta_{25} - \alpha_{25}))z_{t-5} \\
+ \alpha_5 b_{t-k} + (\beta_6 - (\beta_{30} - \alpha_{30})(\beta_5 - \alpha_5))(T_g - t + k_t) + e_{1t} \]

\[ \Delta_t f_t = \beta_0 + \beta_{11}\Delta_t s_{t-1} + \beta_{12}\Delta_t s_{t-2} + \beta_{13}\Delta_t s_{t-3} + \beta_{14}\Delta_t s_{t-4} + \beta_{15}\Delta_t s_{t-5} \\
+ \beta_{21}\Delta_t f_{t-1} + \beta_{22}\Delta_t f_{t-2} + \beta_{23}\Delta_t f_{t-3} + \beta_{24}\Delta_t f_{t-4} + \beta_{25}\Delta_t f_{t-5} \\
+ \beta_{30}k_t + \beta_{31}k_{t-1} + \beta_{32}k_{t-2} + \beta_{33}k_{t-3} + \beta_{34}k_{t-4} + \beta_{35}k_{t-5} \\
+ \beta_{40}z_t + \beta_{41}z_{t-1} + \beta_{42}z_{t-2} + \beta_{43}z_{t-3} + \beta_{44}z_{t-4} + \beta_{45}z_{t-5} \\
+ \beta_5 b_{t-k} + \beta_6(T_g - t + k_t) + e_{2t} \]

The estimation results for the second sub-sample are reported in Table A2. Again, from Table A2 we can see that for the same parameters, the first step estimation gives exactly the same results as the second step estimation. Therefore, the results from the first step estimation for the spot equation and the results from the second step estimation for the futures equation can be used jointly to report both the value and the t-statistic for each parameter of the system.
Table A2 SUR estimation of the restricted VECM, 27/10/1997 - 30/12/2005

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>Variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_0 )</td>
<td>-0.002644</td>
<td>-1.098977</td>
<td>( \alpha_{11} )</td>
<td>-0.137221</td>
<td>-1.027612</td>
</tr>
<tr>
<td>( \alpha_{11} )</td>
<td>-0.137221</td>
<td>-1.027612</td>
<td>( \alpha_{21} )</td>
<td>0.129762</td>
<td>0.993644</td>
</tr>
<tr>
<td>( \alpha_{12} )</td>
<td>-0.132683</td>
<td>-0.874368</td>
<td>( \alpha_{22} )</td>
<td>0.066763</td>
<td>0.45024</td>
</tr>
<tr>
<td>( \alpha_{13} )</td>
<td>0.136625</td>
<td>0.883823</td>
<td>( \alpha_{23} )</td>
<td>-0.234273</td>
<td>-1.552712</td>
</tr>
<tr>
<td>( \alpha_{14} )</td>
<td>0.195661</td>
<td>1.349549</td>
<td>( \alpha_{24} )</td>
<td>-0.167357</td>
<td>-1.180441</td>
</tr>
<tr>
<td>( \alpha_{15} )</td>
<td>0.127151</td>
<td>1.087966</td>
<td>( \alpha_{25} )</td>
<td>-0.177045</td>
<td>-1.557064</td>
</tr>
<tr>
<td>( \alpha_{21} )</td>
<td>0.129762</td>
<td>0.993644</td>
<td>( \alpha_{30} )</td>
<td>0.000707</td>
<td>1.634644</td>
</tr>
<tr>
<td>( \alpha_{22} )</td>
<td>0.066763</td>
<td>0.450239</td>
<td>( \alpha_{31} )</td>
<td>0.000707</td>
<td>1.634644</td>
</tr>
<tr>
<td>( \alpha_{23} )</td>
<td>-0.234273</td>
<td>-1.552712</td>
<td>( \beta_0 )</td>
<td>-0.002644</td>
<td>-1.098977</td>
</tr>
<tr>
<td>( \alpha_{24} )</td>
<td>-0.167357</td>
<td>-1.180441</td>
<td>( \beta_{11} )</td>
<td>0.534737</td>
<td>3.893101</td>
</tr>
<tr>
<td>( \alpha_{25} )</td>
<td>-0.177045</td>
<td>-1.557064</td>
<td>( \beta_{12} )</td>
<td>0.32821</td>
<td>2.105849</td>
</tr>
<tr>
<td>( \alpha_{30} )</td>
<td>0.000707</td>
<td>1.634644</td>
<td>( \beta_{13} )</td>
<td>0.452075</td>
<td>2.848021</td>
</tr>
<tr>
<td>( \alpha_{31} )</td>
<td>0.0002</td>
<td>0.497081</td>
<td>( \beta_{14} )</td>
<td>0.341862</td>
<td>2.296038</td>
</tr>
<tr>
<td>( \alpha_{32} )</td>
<td>-0.000322</td>
<td>-0.773014</td>
<td>( \beta_{15} )</td>
<td>0.166829</td>
<td>1.390419</td>
</tr>
<tr>
<td>( \alpha_{33} )</td>
<td>0.000698</td>
<td>1.664874</td>
<td>( \beta_{21} )</td>
<td>-5.29E-01</td>
<td>-3.940113</td>
</tr>
<tr>
<td>( \alpha_{34} )</td>
<td>0.000319</td>
<td>0.786435</td>
<td>( \beta_{22} )</td>
<td>-0.39413</td>
<td>-2.584715</td>
</tr>
<tr>
<td>( \alpha_{35} )</td>
<td>-0.000172</td>
<td>-0.39598</td>
<td>( \beta_{23} )</td>
<td>-0.549723</td>
<td>-3.543673</td>
</tr>
<tr>
<td>( \alpha_{40} )</td>
<td>-1.67E-05</td>
<td>-0.680523</td>
<td>( \beta_{24} )</td>
<td>-0.313558</td>
<td>-2.151035</td>
</tr>
<tr>
<td>( \alpha_{41} )</td>
<td>-6.57E-07</td>
<td>-0.025665</td>
<td>( \beta_{25} )</td>
<td>-0.216724</td>
<td>-1.853865</td>
</tr>
<tr>
<td>( \alpha_{42} )</td>
<td>-1.94E-05</td>
<td>-0.75698</td>
<td>( \beta_{30} )</td>
<td>6.59E-04</td>
<td>1.522443</td>
</tr>
<tr>
<td>( \alpha_{43} )</td>
<td>-1.30E-05</td>
<td>-0.507513</td>
<td>( \beta_{31} )</td>
<td>0.000168</td>
<td>0.417396</td>
</tr>
<tr>
<td>( \alpha_{44} )</td>
<td>4.02E-05</td>
<td>1.589688</td>
<td>( \beta_{32} )</td>
<td>-3.45E-04</td>
<td>-0.826648</td>
</tr>
<tr>
<td>( \alpha_{45} )</td>
<td>-5.99E-07</td>
<td>-0.024433</td>
<td>( \beta_{33} )</td>
<td>6.83E-04</td>
<td>1.628365</td>
</tr>
<tr>
<td>( \alpha_{5} )</td>
<td>0.020817</td>
<td>0.260679</td>
<td>( \beta_{34} )</td>
<td>3.12E-04</td>
<td>0.768919</td>
</tr>
<tr>
<td>( \alpha_{6} )</td>
<td>1.47E-05</td>
<td>1.170159</td>
<td>( \beta_{35} )</td>
<td>-1.74E-04</td>
<td>-0.400411</td>
</tr>
<tr>
<td>( \beta_{11} )</td>
<td>0.534737</td>
<td>3.893101</td>
<td>( \beta_{40} )</td>
<td>3.18E-05</td>
<td>1.280823</td>
</tr>
<tr>
<td>( \beta_{21} )</td>
<td>-0.529291</td>
<td>-3.940114</td>
<td>( \beta_{41} )</td>
<td>3.13E-05</td>
<td>1.215057</td>
</tr>
<tr>
<td>( \beta_{22} )</td>
<td>-0.39413</td>
<td>-2.584715</td>
<td>( \beta_{42} )</td>
<td>2.96E-06</td>
<td>0.114746</td>
</tr>
<tr>
<td>( \beta_{23} )</td>
<td>-0.549723</td>
<td>-3.543673</td>
<td>( \beta_{43} )</td>
<td>2.32E-06</td>
<td>0.090538</td>
</tr>
<tr>
<td>( \beta_{24} )</td>
<td>-0.313558</td>
<td>-2.151035</td>
<td>( \beta_{44} )</td>
<td>4.73E-05</td>
<td>1.865434</td>
</tr>
<tr>
<td>( \beta_{25} )</td>
<td>-0.216724</td>
<td>-1.853865</td>
<td>( \beta_{45} )</td>
<td>1.33E-06</td>
<td>0.053997</td>
</tr>
<tr>
<td>( \beta_{30} )</td>
<td>0.000659</td>
<td>1.522443</td>
<td>( \beta_{5} )</td>
<td>-0.066636</td>
<td>-0.811159</td>
</tr>
<tr>
<td>( \beta_{5} )</td>
<td>-0.066636</td>
<td>-0.811159</td>
<td>( \beta_{6} )</td>
<td>1.90E-05</td>
<td>1.502916</td>
</tr>
</tbody>
</table>
5.1 Introduction

In the last chapter, the relationship between the FTSE100 index and futures was examined within a linear VECM framework. An important assumption of the linear VECM is that both futures and spot price series have constant variance. However, financial time series usually contain time-varying volatility. The existence of conditional variance means that the linear model may be misspecified and hence may provide inconclusive inference on the relationship between the two markets. Given the close link between information and volatility, it has been widely argued that a test of causality-in-mean captures only part of the short-run dynamics between index futures and cash price, and a test of causality-in-variance may provide an alternative perspective to investigate dynamics between the two markets (Chan et al., 1991; Abhyankar, 1998). Furthermore, while stock index futures are usually believed to provide an important vehicle for investors to hedge stock market movements, the US Brady Commission Report (1988) expressed concern about the role of index futures in the 1987 stock market crash. Futures transactions have actually been widely blamed for exacerbating rather than dampening extraordinary movements in stock prices. The detrimental economic and financial consequences of the alleged market volatility have prompted some analysts and regulators to call for limitations on futures trading activity (Edwards, 1988; Lee and Ohk, 1992; Yu, 2001). Less concern was evident in the UK, but Antoniou and Garrett (1993) found that the expected arbitrage relationships between spot and futures did break down on the day of the crash itself. These considerations have not only generated substantial practical as well as academic interest in volatility transmissions/spillovers between index futures and cash markets, they also have important implications for regulation of the futures market.

This chapter examines the volatility processes of the FTSE100 spot and futures markets and possible volatility spillovers between them over 28/10/1986–30/12/2005. The linear VECM developed in the last chapter is extended by using a DCC-TGARCH specification to model the residual variance-covariance processes. This is formulated so as to account for several factors. First, the DCC-TGARCH can model 'volatility
clustering’ and ‘leverage effects’, which are two important features of financial time series that are widely documented in the literature. Second, the model can capture the conditional correlation between spot and futures markets, with important implications for hedging strategies using futures contracts. Given the close arbitrage link between spot and futures markets, their volatilities as well as their returns would be expected to be highly positively correlated. It is the high positive correlation between spot and futures markets that makes hedging with futures contracts a simple and perhaps the most widely used strategy for reducing and managing risk—a hedge is achieved by taking opposite positions in spot and futures markets simultaneously, so that any loss sustained from an adverse price movement in one market should to some degree be offset by a favorable price movement in the other. Most institutional investors, for example, would use stock index futures to protect their portfolios against market risk. But, if the conditional variances of the index and futures returns change over time, we would also expect the correlation between the two series to be time-varying. Clearly therefore, accurate estimates of the conditional correlation between spot and futures returns as well as of the conditional variances of the returns is important for such hedging because all are important inputs for the calculation of dynamic hedge ratios, $h_t = \rho_{SF,t} \sigma_{S,t} / \sigma_{F,t}$ (Brooks, 2002). Third, unrestricted multivariate GARCH models are often not parsimonious and their estimation usually proves to be a formidable task, even though they can make more efficient use of information in the variance-covariance matrix. The DCC-GARCH model of Engle (2002), which is designed to be estimated in two-steps and is easy to implement, has the flexibility of univariate GARCH models coupled with parsimonious parametric models for the conditional correlation. The DCC-GARCH has therefore been used to model the conditional variance and covariance processes of spot and futures returns.

Since DCC-GARCH does not allow for dynamic dependence between volatility series, it does not provide a direct observation of causality-in-variance. Cheung and Ng (1996) have developed a two-step CCF test that involves estimation of univariate time-series models in the first step and calculation of the cross-correlation function (CCF) of the squared-standardized residuals in the second step to examine causality-in-variance. The first step of the CCF test, formulating and estimating univariate GARCH models, can be taken from the first step of estimating the DCC-TGARCH model, so the latter provides a reasonable framework within which the CCF test can be carried out. The
CCF of the squared-standardized residuals estimated from the DCC-TGARCH model can then be calculated to investigate causality-in-variance between the two series.

The rest of this chapter is organised as follows: section 5.2 analyses the volatility-clustering feature of FTSE100 spot and futures prices and acknowledges the limitations of the linear VECM developed in chapter 4 in modeling FTSE100 spot-futures price dynamics. Section 5.3 introduces the DCC-TGARCH model and the CCF test for causality-in-variance used in this study. The empirical results are reported in section 5.4 and section 5.5 draws conclusions.

5.2 Limitations of the linear VECM analysis

The basic linear VECM developed in the last chapter to analyse the dynamic relationship between FTSE100 spot and futures markets are as follows

\[ \Delta_k s_t = \alpha_0 + \sum_{j=1}^{M_1} \alpha_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_{3j} k_{t-j} + \sum_{j=0}^{M_4} \alpha_{4j} z_{t-j} + \alpha_5 b_{l-k} + \alpha_6 (T_{g} - t + k_t) + e_{1t} \]  

\[ (5.1a) \]

\[ \Delta_k f_t = \beta_0 + \sum_{j=1}^{N_1} \beta_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{N_2} \beta_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{N_3} \beta_{3j} k_{t-j} + \sum_{j=0}^{N_4} \beta_{4j} z_{t-j} + \beta_5 b_{l-k} + \beta_6 (T_{g} - t + k_t) + e_{2t} \]  

\[ (5.1b) \]

where \( e_{1t} \) and \( e_{2t} \) are assumed to be identically and independently distributed (i.i.d.) random error terms with zero mean and constant variance. An important feature of financial time series, however, is ‘volatility clustering’, a form of heteroskedasticity in which large (small) changes in asset prices (of either sign) tend to follow large (small) changes (of either sign). Volatility may therefore be time-varying and positively correlated to its level in the immediately preceding periods. If the assumption of constant variance is violated, then the presence of heteroskedasticity could affect the standard errors of the least-squares estimates and thus any hypothesis tests which follow. Different approaches were therefore used to examine whether ‘volatility clustering’ is a feature of FTSE100 spot and futures price changes during the period.
5.2.1 Visual inspection of time plot of return series

First of all, inspecting the time plot of price changes (i.e. the return series) directly can tell whether the volatility varies over time. Both spot and futures price changes (in logs) are plotted against time over the two sub-samples (Figure 5.1 and Figure 5.3 for FTSE100 spot returns during the two sub-samples; Figure 5.4 and Figure 5.6 for FTSE100 futures returns during the two sub-samples). During the first sub-sample period, the market has had an overall rising trend, though both spot and futures markets have fallen dramatically during the 1987 stock market crash period, indicated by several very large negative returns during the crash week (also one very large positive return on 21/10/1987). Once the 1987 crash period (19/10/1989–23/10/1987) is omitted from the original time series (see Figure 5.2 for spot returns and Figure 5.5 for futures returns), both markets appear to be much more stable in general, evidenced by relatively small positive and negative returns. During the second sub-sample period, the London stock market has experienced a roughly two years rising period due to the dotcom bubble of the late 1990s, followed by a three-year-recession due to the burst of the dotcom bubble at the beginning of 2000, which was then followed by a recovery from the recession from the beginning of 2003. Accordingly, both markets appear to have higher volatility during the three years recession period, when many large positive and large negative returns are observed during a short space of time. On the other hand, both markets appear to be relatively tranquil during the last three years of the sub-sample when the market has experienced a long recovery from the recession. Overall, it seems that the market is more volatile during falling periods than during rising periods, and an obvious feature of both markets during both sub-samples is that large/small price changes tend to follow large/small price changes, implying a pattern of ‘volatility clustering’.
Figure 5.1 Time plot of daily FTSE 100 spot returns (in logs), 28/10/1986 – 17/10/1997

FTSE100 spot returns, 28/10/1986 – 17/10/1997
Figure 5.2 Time plot of daily FTSE 100 spot returns (in logs), 28/10/1986 – 17/10/1997 (19/10/1987 – 23/10/1987 not included)
Figure 5.3 Time plot of daily FTSE 100 spot returns (in logs), 27/10/1997 – 30/12/2005

FTSE100 spot returns, 27/10/1997 – 30/12/2005
Figure 5.4 Time plot of daily FTSE 100 futures returns (in logs), 28/10/1986 – 17/10/1997
Figure 5.5 Time plot of daily FTSE 100 futures returns (in logs), 28/10/1986 – 17/10/1997 (19/10/1987 – 23/10/1987 not included)
Figure 5.6 Time plot of daily FTSE 100 futures returns (in logs), 27/10/1997 – 30/12/2005
5.2.2 Autocorrelations in squared returns

If ‘volatility clustering’ is present, it should also be signaled by serially correlated squared returns, which are often used to construct a coarse measure of the variance. For comparison, the autocorrelation and partial autocorrelation functions (ACF and PACF) up to 5 lags for both spot and futures returns and squared returns are reported in Table 5.1. A noticeable characteristic is that serial correlations in the squared returns are more prevalent than those in the returns for both price series. All ACFs and PACFs of squared returns are significant at the 1% level up to 5 lags, with the exception that the PACF at lag 5 for futures squared returns is only significant at the 10% level. Serial correlation in squared returns can be attributed to time-varying volatility.

5.2.3 ARCH LM test

Given that the preliminary data analysis indicates ‘volatility clustering’ in both FTSE100 spot and futures price series, an ARCH LM test was performed to check for autoregressive conditional heteroskedasticity (ARCH) effects in the residuals of equations (5.1a) and (5.1b). To test the null hypothesis that there is no ARCH up to order q in the residuals, the following regression was estimated

\[ e_t^2 = \lambda_0 + \sum_{s=1}^{q} \lambda_s e_{t-s}^2 + \nu_t \] (5.2)

where \( e_t \) is the residual from equations (5.1a) and (5.1b). Table 5.2 reports the F- and Obs*R-squared statistics testing the null hypothesis of no ARCH effect up to order 5 \((q = 5)\) in the residuals of an OLS estimation of equations (5.1a) and (5.1b), which involves 4 lags for the first sample period and 5 lags for the second sample period. The F-statistic is a test for the joint significance of all lagged squared residuals. The Obs*R-squared statistic is Engle’s LM test statistic, computed as the number of observations times the \( R^2 \) from the test regression. The LM test statistic, \( TR^2 \), is asymptotically distributed as \( X^2(q) \) under quite general conditions, that is, \( TR^2 \sim X^2(q) \) (Brooks, 2002). It can be seen that in both sub-periods, both the F- and LM-statistics suggest the presence of ARCH effects in the FTSE100 spot and futures returns. Given the strong evidence of ARCH effects in the residuals, it is important to explicitly model the time-varying volatility in both spot and futures returns to support valid hypothesis testing.
Table 5.1 ACF and PACF of the returns and squared returns

<table>
<thead>
<tr>
<th></th>
<th>Spot Returns</th>
<th>Futures Returns</th>
<th>Spot Squared Returns</th>
<th>Futures Squared Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACF</td>
<td>PACF</td>
<td>ACF</td>
<td>PACF</td>
</tr>
<tr>
<td>lags</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.064*</td>
<td>0.064*</td>
<td>0.028</td>
<td>0.028</td>
</tr>
<tr>
<td>2</td>
<td>0.007</td>
<td>0.003</td>
<td>-0.041 b</td>
<td>-0.042 b</td>
</tr>
<tr>
<td>3</td>
<td>0.016</td>
<td>0.015</td>
<td>0.030</td>
<td>0.033 c</td>
</tr>
<tr>
<td>4</td>
<td>0.065*</td>
<td>0.063*</td>
<td>0.043 b</td>
<td>0.039 b</td>
</tr>
<tr>
<td>5</td>
<td>0.009</td>
<td>0.001</td>
<td>-0.004</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Spot Returns</th>
<th>Futures Returns</th>
<th>Spot Squared Returns</th>
<th>Futures Squared Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACF</td>
<td>PACF</td>
<td>ACF</td>
<td>PACF</td>
</tr>
<tr>
<td>lags</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.019</td>
<td>-0.019</td>
</tr>
<tr>
<td>2</td>
<td>-0.061 a</td>
<td>-0.061 a</td>
<td>-0.060 a</td>
<td>-0.06 a</td>
</tr>
<tr>
<td>3</td>
<td>-0.099 a</td>
<td>-0.099 a</td>
<td>-0.110 a</td>
<td>-0.113 a</td>
</tr>
<tr>
<td>4</td>
<td>0.036</td>
<td>0.031</td>
<td>0.045 b</td>
<td>0.037 c</td>
</tr>
<tr>
<td>5</td>
<td>-0.035</td>
<td>-0.047 b</td>
<td>-0.042 c</td>
<td>-0.055 b</td>
</tr>
</tbody>
</table>

Notes:
There are 2776 observations in period 28/10/1986 – 17/10/1997 and 2051 observations in period 27/10/1997 – 30/12/2005. Therefore, the ACF and PACF are significant at the 1% level if outside the range [-0.0486, +0.0486], significant at 5% if outside the range [-0.0372, +0.0372] and significant at 10% if outside the range [-0.0311, +0.0311] during 28/10/1986 – 17/10/1997. During 27/10/1997 – 30/12/2005, the ACF and PACF are significant at the 1% level if outside the range [-0.0565, +0.0565], significant at 5% if outside the range [-0.0433, +0.0433] and significant at 10% if outside the range [-0.0362, +0.0362].

* significant at the 1% level
b significant at the 5% level
c significant at the 10% level.
Table 5.2 ARCH-LM test on OLS residuals of equations (5.1a) and (5.1b)

<table>
<thead>
<tr>
<th></th>
<th>28/10/1986 - 17/10/1997</th>
<th>27/10/1997 - 30/12/2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spot residuals</td>
<td>Futures residuals</td>
</tr>
<tr>
<td>F-statistic</td>
<td>44.46046</td>
<td>26.8418</td>
</tr>
<tr>
<td>Prob. F(5, 2761)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Obs*R-squared</td>
<td>206.1844</td>
<td>128.2658</td>
</tr>
<tr>
<td>Prob. Chi-square(5)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: The null hypothesis is that there is no ARCH effect up to order 5 in the residuals.

5.3 Methodology

5.3.1 GARCH class of models

The presence of successive periods of relative volatility and stability in financial time series can be interpreted as 'autocorrelation in volatility'. This has motivated the creation of the ARCH model by Engle (1982). To understand how the model works, a definition of the conditional variance of a random variable, $e_t$, is required. The conditional variance of $e_t$ may be denoted $\sigma_t^2$, which can be written as

$$
\sigma_t^2 = \text{var}(e_t | e_{t-1}, e_{t-2}, \ldots) = E[(e_t - E(e_t))^2 | e_{t-1}, e_{t-2}, \ldots] \quad (5.3)
$$

It is usually assumed that $E(e_t) = 0$, so

$$
\sigma_t^2 = \text{var}(e_t | e_{t-1}, e_{t-2}, \ldots) = E[e_t^2 | e_{t-1}, e_{t-2}, \ldots] \quad (5.4)
$$

Equation (5.3) states that the conditional variance of a zero mean normally distributed random variable $e_t$ is equal to the conditional expected value of the square of $e_t$.

Under an ARCH($q$) model, therefore, the 'autocorrelation in volatility' is modeled by
allowing the conditional variance of the error term, $\sigma_{i}^{2}$, to depend on the $q$ lags of squared errors

$$\sigma_{i}^{2} = \delta_{0} + \sum_{i=1}^{q} \delta_{i} e_{i-i}^{2}$$

(5.5)

However, ARCH models have some limitations. First, the value of $q$, the number of lags of the squared error required to capture all the dependence in the conditional variance, might be very large. This would result in a large ARCH model that was not parsimonious. Second, given that a conditional variance must always be strictly positive, all of the coefficients in the ARCH model are usually required to be non-negative. However, ceteris paribus, the more parameters there are in the ARCH model, the more likely it is that one or more of them will have negative estimated values. To overcome the limitations of ARCH models, Bollerslev (1986) proposed the GARCH (generalized ARCH) conditional variance specification that allows for a parsimonious parameterisation of the lag structure. The GARCH model allows the conditional variance to be dependent on its previous own lags ($\sigma_{i-i}^{2}$) as well as lags of squared errors ($e_{i-i}^{2}$). In general, a GARCH(1,1) model will be sufficient to capture volatility clustering in the data, and higher order models are rarely estimated or even entertained in the academic finance literature. A GARCH(1,1) model can be written as

$$\sigma_{i}^{2} = \delta_{0} + \delta_{1} e_{i-1}^{2} + \delta_{2} \sigma_{i-1}^{2}$$

(5.6)

GARCH models are more widely used than ARCH models because they are more parsimonious and avoid over-fitting. Consequently, the model is less likely to breach non-negativity constraints on parameters (Brooks, 2002).

The GARCH(1,1) model of equation (5.6) is a univariate model, describing the dynamic process of the conditional variance for a single asset. For a group of assets, several different multivariate GARCH specifications have been proposed in the literature to model changes over time of both covariation between assets and individual conditional variances. Examples of multivariate GARCH models include the VECH, the diagonal VECH and the BEKK models (Brooks, 2002). However, unrestricted multivariate GARCH models are often not parsimonious and their estimation may be
very difficult. The exact interpretation of the impact of individual coefficients in an unrestricted multivariate GARCH model is also difficult. The DCC-GARCH model developed more recently by Engle (2002) has the flexibility of univariate GARCH models coupled with parsimonious parametric models for the conditional correlation. Due to its ease in implementation, the DCC-GARCH model has been employed in many studies to examine conditional variance-covariance processes in groups of financial assets.

5.3.2 VECM for the mean and DCC-TGARCH (1,1) for the variance-covariance

Based on the study conducted in Chapter 4, the VECM of equations (5.1a) and (5.1b) were also used here to model the mean of the return series for FTSE100 spot and futures prices. In order to identify both the conditional variance process for each of the FTSE100 spot and futures return series and the conditional correlation between them, the residual vector from equations (5.1a) and (5.1b) was first assumed to follow a DCC-GARCH (1,1) process. The DCC-GARCH model is based on the assumption that asset returns are conditionally multivariate normal with zero mean and covariance matrix

\[ H_t = \text{Cov}(E_t \mid \phi_{t-k_t}) \]

where \( \phi_{t-k_t} \) denotes all information available at time \( t-k_t \). The residual vector \( E_t = (e_{1t}, e_{2t})' \) from equations (5.1a) and (5.1b) can therefore be written as

\[ E_t \mid \phi_{t-k_t} \sim N(0, H_t) \]  \hspace{1cm} (5.7)

and

\[ H_t = D_t R_t D_t' = \begin{bmatrix} \sigma_{1t}^2 & 0 \\ 0 & \sigma_{2t}^2 \end{bmatrix} \begin{bmatrix} 1 & \rho_{12,t} \\ \rho_{12,t} & 1 \end{bmatrix} \begin{bmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{bmatrix} \]  \hspace{1cm} (5.8)

where \( \sigma_{1t}^2 \) and \( \sigma_{2t}^2 \) are conditional variances of spot and futures returns respectively. Both are assumed to follow a univariate GARCH (1,1) process. \( R_t \) is the conditional correlation matrix with the component \( \rho_{12,t} \) being the conditional correlation between spot and futures markets. The processes of \( \sigma_{1t}^2, \sigma_{2t}^2 \) and \( \rho_{12,t} \) are defined below.

147
It is well known that financial data often display a 'leverage effect', i.e. a negative shock to financial asset prices is likely to cause volatility to rise by more than a positive shock of the same magnitude. To account for 'volatility clustering' as well as a possible asymmetric impact of lagged innovations on current volatility, the univariate GARCH (1,1) model was further extended as in Glosten et al. (1993). That is, zero was used as a threshold value to separate the impacts of past shocks. The final threshold GARCH (1,1) model, or TGARCH(1,1), for $\sigma_{it}^2 \ (i = 1,2)$ is

$$
\sigma_{it}^2 = \psi_{i0} + \psi_{i1} e_{i,t-k}^2 + \psi_{i2} e_{i,t-k}^2 I_i + \psi_{i3} \sigma_{i,t-k}^2
$$

where $I_i \ (i = 1,2)$ is the indicator function with $I_i = 1$ if $e_{i,t-k} < 0$ and $I_i = 0$ otherwise. A positive value of $\psi_{i2}$ means that negative residuals tend to increase the variance more than positive residuals. For the TGARCH (1,1) model of equation (5.9), $\psi_{i0} > 0$, $\psi_{i1} + \psi_{i2} > 0$ and $0 < \psi_{i3} < 1$ are sufficient conditions to ensure positive definite conditional variance: $\sigma_{it}^2 > 0$. The short-run persistence of positive shocks is given by $\psi_{i1}$ and negative shocks by $\psi_{i1} + \psi_{i2}$. Under the assumption that the conditional shocks follow a symmetric distribution, the average short-run persistence is $\psi_{i1} + \psi_{i2} / 2$ and the average long-run persistence is $\psi_{i1} + \psi_{i2} / 2 + \psi_{i3}$. Ling and McAleer (2002) show that the necessary and sufficient condition for the existence of the second moment of $e_{it}$ (in other words $E(e_{t}^2) < \infty$) in the TGARCH(1,1) model is $\psi_{i1} + \psi_{i2} / 2 + \psi_{i3} < 1$ (Lanza et al., 2006).

The univariate TGARCH (1, 1) models for the spot and futures return series are estimated in the first step. The estimates of $\sigma_{it}^2 \ (i = 1,2)$ from the first step are used to calculate the standardized residuals, $e_{it} = e_{it} / \sigma_{it} \ (i = 1,2)$, which are then used in the second step to calculate the conditional correlation between the two markets. The proposed dynamic covariance and correlation structure can be specified respectively as equation (5.10) and (5.11):

$$
Q_t = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 \bar{\Xi}_{t-k} \bar{\Xi}_{t-k} + \theta_2 Q_{t-k}
$$

(5.10)
\[ R_t = Q_t^{-1} Q_t^* Q_t^*^{-1} \] (5.11)

where \( Q_t \) is the conditional covariance matrix of the standardized residuals, with spot and futures conditional variances \( \sigma_{it}^2 (i = 1, 2) \) and conditional covariance \( \sigma_{12t} \):

\[
Q_t = \begin{bmatrix}
\sigma_{it}^2 & \sigma_{12t} \\
\sigma_{12t} & \sigma_{2t}^2
\end{bmatrix}
\]

\( \overline{Q} \) is the unconditional covariance matrix of the standardized residuals. In the estimation procedure \( \overline{Q} \) is replaced by the sample analogue \( T^{-1} \sum_{t=1}^{T} \Xi_t \Xi_t^t \), where \( \Xi_t = (e_{1t}, e_{2t})' \) is the standardized residual vector; and \( \theta_1 \) and \( \theta_2 \) are parameters. \( Q_t^* \) is a 2x2 diagonal matrix composed of the squared root of the diagonal elements of \( Q_t \), that is, \( Q_t^* = \begin{bmatrix} \sigma_{it} & \sigma_{12t} \\
\sigma_{12t} & \sigma_{2t}
\end{bmatrix} \). Therefore, \( Q_t^*^{-1} \) is a diagonal matrix with \( 1/\sigma_{it} \) (i = 1, 2) on the leading diagonal. The conditional correlation matrix is therefore \( R_t = \begin{bmatrix} 1 & \rho_{12t} \\
\rho_{12t} & 1
\end{bmatrix} \), where \( \rho_{12t} = \sigma_{12t} / \sigma_{it} \sigma_{2t} \). If \( \theta_1 \) and \( \theta_2 \) are zero, the Constant Conditional Correlation (CCC) model developed by Bollerslev (1990) is obtained and where they are non-zero, an ARMA-type structure for the conditional correlation emerges.

5.3.3 CCF test for causality-in-variance

Since the DCC-TGARCH model does not allow for dynamic dependence between volatility series, it is not possible to view causality-in-variance directly in this framework. However, the two-step estimation procedure of the DCC-TGARCH model conforms to the two stage procedure to test for causality-in-variance based on the cross correlation function (CCF) of squared-standardized residuals proposed by Cheung and Ng (1996). The first stage of the CCF test involves the estimation of univariate time-series models, allowing for time variation in both conditional means and conditional variances. In the second stage the series of squared residuals standardized by conditional variances are constructed. The CCF of these squared-standardized residuals
is then used to test the null hypothesis of no causality-in-variance. Cheung and Ng (1996) applied the CCF test in two empirical examples, each one illustrating that the CCF method is useful in determining causality in variance.

The two-step CCF approach and the two-step estimation procedure of the DCC-TGARCH model share the advantage of not involving simultaneous modelling of both intra- and inter-series dynamics of the variance processes and are therefore easy to implement. The CCF test was therefore used in combination with the DCC-TGARCH(1,1) to test for causality-in-variance in this study. The first-step of the CCF test was borrowed from the first-step model formulation and estimation of the DCC-TGARCH (1,1). The estimates of $\sigma_{i,t}$ ($i = 1, 2$) from the first step were used in the second step to calculate the standardized residuals, $e_{it} = e_{it}/\sigma_{it}$ ($i = 1, 2$). The CCF of the squared-standardized residuals ($e_{it}^2$) at different lags $k_i$, which can be designated as $CCF^2(k_i)$, was then used to examine causality-in-variance between spot and futures markets. Under the null hypothesis of no causality-in-variance, $(\sqrt{T}CCF^2(k_1), \sqrt{T}CCF^2(k_2), \sqrt{T}CCF^2(k_3), \ldots, \sqrt{T}CCF^2(k_m))$ converge to $N(0, I_m)$ as $T \to \infty$, where $T$ is the number of observations, $k_1, k_2, k_3, \ldots, k_m$ are $m$ different integers and $I_m$ is an $m \times m$ identity matrix. There is no evidence of causality-in-variance if the $CCF^2(k_i)$, at all possible leads and lags, are not significantly different from zero (Cheung and Ng, 1996; Cheung and Fung, 1997). For example, the hypothesis that spot does not cause futures in variance implies zero cross correlation between $e_{1,s}^2$ and $e_{2,t}^2$ ($s < t$) for all $s$ and $t$, and vice versa. To test for a causal relationship at a specified lag $k_i$, $\sqrt{T} \cdot CCF^2(k_i)$ can be compared with the standard normal distribution. $CCF^2(k_i)$ is significantly different from zero at the 1% level if it is outside the band $(-2.58 \cdot 1/\sqrt{T}, +2.58 \cdot 1/\sqrt{T})$, significantly different from zero at the 5% level if it is outside the band $(-1.96 \cdot 1/\sqrt{T}, +1.96 \cdot 1/\sqrt{T})$ and significantly different from zero at the 10% level if it is outside the band $(-1.64 \cdot 1/\sqrt{T}, +1.64 \cdot 1/\sqrt{T})$. 
As pointed out by Cheung and Ng (1996), the existence of any serial correlation in $\varepsilon_{it}$ or $\varepsilon_{it}^2$ can affect the size of the proposed CCF test for causality in variance (because the serial correlation pattern in $\varepsilon_{it}$ or $\varepsilon_{it}^2$ could introduce a similar serial correlation pattern in the cross correlation function between the two series of squared-standardized residuals). The time-series model specified in the first stage should ‘accurately’ account for serial correlation in the data. The Ljung-Box Q statistics calculated from standardized residuals and their squares were used to check whether the selected model adequately described serial correlation in the first and second moments.

5.4 Empirical results

5.4.1 Results of the DCC-TGARCH (1,1) model

The empirical analysis based on the DCC-TGARCH (1,1) was performed using the same mean equations (equations (5.1a) and (5.1b)) and within the same two sub-samples (28/10/1986–17/10/1997 and 27/10/1997–30/12/2005) as in chapter 4. To compare with Chapter 4, the same lag structures for the mean equations, that is, 4 lags for the first sub-sample and 5 lags for the second sub-sample, were used in this chapter (Ljung-Box Q tests on residuals were also used later to check for model adequacy). The DCC-TGARCH models were estimated using maximum-likelihood under the multivariate normal distribution assumption, although this was unsuccessful for the second sub-sample and a multivariate t-distribution assumption had to be used instead (see page 156).

The results in Chapter 4 suggest that the 1987 stock market crash had a significant impact on both FTSE100 spot and futures price dynamics. Therefore, for the first sub-sample, equations (4.11a) and (4.11b) of Chapter 4 were initially used as the mean equations of the DCC-TGARCH model, that is, a dummy variable, $DM_i$, and its lags ($DM_{t-i}$, $i=0, 1, 2, 3, 4$) were included in the mean equations to capture the 1987 stock market crash effect. However, the estimation of the DCC-TGARCH failed to converge if $DM_i$ and/or its lags were included. The non-convergence may be caused by a multicollinearity problem introduced by the inclusion of the dummy variable and/or its lags, making it impossible to estimate all the coefficients. Unfortunately, the exact collinear relationship is difficult to detect, as discussed by Brooks (2002). To capture
the 1987 stock market crash effect while avoiding the influence of the dummy variable on the estimation of the DCC-TGARCH model, the crash effect was removed from the original spot and futures return series and the DCC-TGARCH model was estimated using the adjusted dataset. That is, the spot return series was adjusted by subtracting the coefficients $\alpha_{7j}$ of the SUR estimation of equations (4.11a) and (4.11b) from the original returns on relevant dates, while the futures return series was similarly adjusted by subtracting the coefficients $\beta_{7j}$ from the original returns on relevant dates. With the dummy variable ($DM$) and 4 lags included (covering 5 trading days of the October 1987 crash), the adjusted spot and futures return series have exactly the same data as the original spot and futures return series on all days except on the 5 trading days from 19/10/1987 to 23/10/1987 (on these five days, the adjusted spot returns can be specified as $\Delta_k s_t - \alpha_{7j} DM$ and the adjusted futures returns can be specified as $\Delta_k f_t - \beta_{7j} DM$, with $j = 0, 1, 2, 3, 4$). With the dataset already adjusted for the 1987 crash effect, equations (5.1a) and (5.1b) were then used as the mean equations for the DCC-TGARCH model. Using the adjusted spot and futures returns during 28/10/1986–17/10/1997, achieving convergence was no longer a problem when estimating the DCC-TGARCH model by maximum likelihood.

Ljung-Box Q tests were used to check for serial correlation in residuals. For both sub-samples, these tests reveal no evidence of serial correlation in either standardized residuals or squared-standardized residuals up to 5 lags, suggesting that the selected model and the lag structure (4 lags for the first sub-sample and 5 lags for the second sub-sample) fit the data reasonably well (see Table 5.3).
Table 5.3 Ljung-Box Q test on standardized residuals ($\varepsilon_{it}$) and their squared values ($\varepsilon_{it}^2$) estimated from the fitted DCC-TGARCH (1,1) model

<table>
<thead>
<tr>
<th>lag</th>
<th>Spot residuals $\varepsilon_{it}$</th>
<th>Futures residuals $\varepsilon_{2t}$</th>
<th>Spot squared residuals $\varepsilon_{it}^2$</th>
<th>Futures squared residuals $\varepsilon_{2t}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q-statistic</td>
<td>p-value</td>
<td>Q-statistic</td>
<td>p-value</td>
</tr>
<tr>
<td>1</td>
<td>0.007</td>
<td>0.932</td>
<td>0.037</td>
<td>0.847</td>
</tr>
<tr>
<td>2</td>
<td>0.268</td>
<td>0.874</td>
<td>0.099</td>
<td>0.952</td>
</tr>
<tr>
<td>3</td>
<td>0.313</td>
<td>0.958</td>
<td>0.416</td>
<td>0.937</td>
</tr>
<tr>
<td>4</td>
<td>0.335</td>
<td>0.987</td>
<td>0.477</td>
<td>0.976</td>
</tr>
<tr>
<td>5</td>
<td>0.665</td>
<td>0.985</td>
<td>1.454</td>
<td>0.918</td>
</tr>
</tbody>
</table>

27/10/1997 – 30/12/2005

<table>
<thead>
<tr>
<th>lag</th>
<th>Spot residuals $\varepsilon_{it}$</th>
<th>Futures residuals $\varepsilon_{2t}$</th>
<th>Spot squared residuals $\varepsilon_{it}^2$</th>
<th>Futures squared residuals $\varepsilon_{2t}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q-statistic</td>
<td>p-value</td>
<td>Q-statistic</td>
<td>p-value</td>
</tr>
<tr>
<td>1</td>
<td>0.244</td>
<td>0.622</td>
<td>0.038</td>
<td>0.845</td>
</tr>
<tr>
<td>2</td>
<td>0.408</td>
<td>0.816</td>
<td>0.206</td>
<td>0.902</td>
</tr>
<tr>
<td>3</td>
<td>0.755</td>
<td>0.860</td>
<td>0.866</td>
<td>0.834</td>
</tr>
<tr>
<td>4</td>
<td>0.776</td>
<td>0.942</td>
<td>0.912</td>
<td>0.923</td>
</tr>
<tr>
<td>5</td>
<td>0.780</td>
<td>0.978</td>
<td>0.912</td>
<td>0.969</td>
</tr>
</tbody>
</table>

Notes: For the finally selected DCC-TGARCH(1,1) model, 4 lags are included in the mean equations for 28/10/1986–17/10/1997 and 5 lags are included in the mean equations for 27/10/1997–30/12/2005. The reported Ljung-Box Q-statistics and the p-values for the Q-statistics for a specified lag provide a joint test of serial correlation in the standardized residuals or squared-standardized residuals up to the specified lag.
The maximum likelihood estimators and the corresponding t-statistics for the DCC-TGARCH (1,1) model for the first period 28/10/1986–17/10/1997 are reported in Table 5.4. The findings are in general consistent with what has been found in Chapter 4. There is bidirectional causality-in-mean between futures and spot markets, with causality from spot to futures much stronger than the reverse. As anticipated, past price movements in the spot/futures market have a positive impact on current price movements in the futures/spot market, indicated by the fact that coefficients on lagged cross market returns are positively significant (except for $\Delta_t s_{t-4}$ in the spot equation, which has a negative coefficient and is significant at the 5% level). This reflects the fact that where one market leads another in its response to new information, the lagged market should respond positively to the price movement in the leading market to keep the spread between the two markets within a no-arbitrage band. Both the spot and futures return series tend to be mean-reverting, indicated by the significant negative impact of lagged own market returns on the current return (except that $\Delta_t s_{t-4}$ in the spot equation is positively significant at the 5% level). It can also be seen that mean-reversion in futures returns is stronger than that in spot returns. As analysed in Chapter 4, mean-reversion in returns could reflect overreaction of both markets to new information, with subsequent corrections driven by the true value of the information. In this regard, mean-reversion is necessary for the markets to be efficient in the long run. Mean-reversion in futures returns could also be explained by bid-ask bounce in futures prices.

The coefficients for $k_{t-j}$ ($j=1, 2, 3, 4$) are positive and significant in both spot and futures equations, indicating that the passage of time has a positive, though quantitatively small, effect on price changes in both markets. This provides support to our conjecture that returns over longer time intervals such as weekends and holidays should be higher than returns over shorter time intervals, ceteris paribus. There is also evidence of significant $z_t$ and/or its lags in both futures and spot equations, suggesting that the 'artificial' price jumps at contract rollovers should not be ignored in explaining spot-futures price dynamics during this sample period (the observed negative effect of $z_t$ on spot price changes may actually reflect a transmission of futures price jumps at contract rollovers to the spot market). As anticipated, those $z_{t-j}$ that are significant in
the futures equation are all positive while the joint effect of $z_{t-j}$ on the current futures price changes minus the joint effect of $z_{t-j}$ on the current spot price changes is also positive, reflecting a jump in futures price and a wider spread between futures and spot prices at contract rollovers.

The error correction term is negative and significant in the futures return equation, suggesting that any divergence from equilibrium between the two markets is corrected in the futures market. The negative sign implies that if the basis was too large/small compared to the equilibrium value, the futures price would tend to fall/rise to recover the no-arbitrage relationship. As explained in Chapter 4, the stronger causality from spot to futures price, indicating stronger lead of the spot price over the futures price in response to new information, may have caused temporary divergence from equilibrium between the two markets in the period analysed. Subsequent adjustments in the futures market would then have been necessary to drive the futures price (and the relationship between the two markets) to a new equilibrium. This could also reflect the fact that the ease of trading in the futures market has caused the FTSE100 futures price to overreact to new information. This has caused temporary divergence from equilibrium and subsequent adjustment in the futures market to correct the overreaction of the futures price and the divergence from equilibrium. However, the coefficient on the error-correction term ($b_{t-1}$) is -0.0763, compared to -0.120918 reported in Chapter 4, suggesting that less than 8% (compared to more than 12% reported in Chapter 4) of the divergence from equilibrium has been corrected in the futures market within one day.

The time to maturity term is negatively significant (though only the 10% level) in the spot return equation, providing support to the conjecture that it might be necessary to induce stationarity in the basis for individual futures contracts.

Turning next to the variance equation, it can be seen that the conditions $\psi_{10} > 0$, $\psi_{11} + \psi_{12} > 0$ and $0 < \psi_{13} < 1$ are all satisfied, ensuring a positive definite conditional variance. Both spot and futures variances are highly persistent. The long-run persistence in variance, measured by $\psi_{11} + \psi_{12}/2 + \psi_{13}$, is 0.9704 for spot returns and 0.9788 for futures returns. And the condition $\psi_{11} + \psi_{12}/2 + \psi_{13} < 1$ is satisfied for both return series, ensuring the existence of the second moment of $e_{ii}$ ($i=1, 2$). As
anticipated, the parameter of the term measuring the leverage effect is positive and significant in both spot and futures variance equations, suggesting that the volatilities of both markets indeed react more strongly to bad news than to good news. The coefficients in the covariance equation are highly significant, indicating time-varying correlation between the two markets. The persistence in the conditional correlation, measured by $\theta_1 + \theta_2$, is 0.9862, suggesting high persistence in the conditional correlation between the two markets.
Table 5.4 Results of DCC-TGARCH(1, 1), 28/10/1986 – 17/10/1997

\[
\Delta_k s_t = \alpha_0 + \sum_{j=1}^{M_1} \alpha_1 \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_2 \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_3 k_{t-j} + \sum_{j=0}^{M_4} \alpha_4 z_{t-j} + \\
\alpha_5 b_{t-k} + \alpha_6 (T_g - t + k_t) + e_t
\]

\[
\Delta_k f_t = \beta_0 + \sum_{j=1}^{N_1} \beta_1 \Delta_k s_{t-j} + \sum_{j=1}^{N_2} \beta_2 \Delta_k f_{t-j} + \sum_{j=0}^{N_3} \beta_3 k_{t-j} + \sum_{j=0}^{N_4} \beta_4 z_{t-j} + \\
\beta_5 b_{t-k} + \beta_6 (T_g - t + k_t) + e_{2t}
\]

<table>
<thead>
<tr>
<th>variables</th>
<th>Spot Return Equation (5.1a)</th>
<th></th>
<th>Futures Return Equation (5.1b)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficients</td>
<td>t-statistics</td>
<td>Coefficients</td>
<td>t-statistics</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.003299</td>
<td>-2.40 \textsuperscript{b}</td>
<td>-0.005439</td>
<td>-3.56 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k s_{t-1})</td>
<td>-0.1292</td>
<td>-2.43 \textsuperscript{b}</td>
<td>0.3397</td>
<td>6.09 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k s_{t-2})</td>
<td>0.0637</td>
<td>1.30</td>
<td>0.3250</td>
<td>5.77 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k s_{t-3})</td>
<td>0.0679</td>
<td>1.26</td>
<td>0.1882</td>
<td>3.01 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k s_{t-4})</td>
<td>0.1054</td>
<td>2.38 \textsuperscript{b}</td>
<td>0.1824</td>
<td>3.77 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k f_{t-1})</td>
<td>0.1774</td>
<td>3.70 \textsuperscript{a}</td>
<td>-0.2589</td>
<td>-4.89 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k f_{t-2})</td>
<td>-0.0507</td>
<td>-1.18</td>
<td>-0.3259</td>
<td>-6.42 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k f_{t-3})</td>
<td>-0.0540</td>
<td>-1.15</td>
<td>-0.1787</td>
<td>-3.21 \textsuperscript{a}</td>
</tr>
<tr>
<td>(\Delta_k f_{t-4})</td>
<td>-0.0870</td>
<td>-2.21 \textsuperscript{b}</td>
<td>-0.1640</td>
<td>-3.60 \textsuperscript{a}</td>
</tr>
<tr>
<td>(k_t)</td>
<td>0.000041</td>
<td>0.20</td>
<td>0.000262</td>
<td>1.11</td>
</tr>
<tr>
<td>(k_{t-1})</td>
<td>0.000822</td>
<td>3.39 \textsuperscript{a}</td>
<td>0.001148</td>
<td>4.26 \textsuperscript{a}</td>
</tr>
<tr>
<td>(k_{t-2})</td>
<td>0.000597</td>
<td>2.52 \textsuperscript{b}</td>
<td>0.000959</td>
<td>3.57 \textsuperscript{a}</td>
</tr>
<tr>
<td>(k_{t-3})</td>
<td>0.000740</td>
<td>3.14 \textsuperscript{a}</td>
<td>0.001044</td>
<td>3.98 \textsuperscript{a}</td>
</tr>
<tr>
<td>(k_{t-4})</td>
<td>0.000580</td>
<td>2.78 \textsuperscript{a}</td>
<td>0.000838</td>
<td>3.56 \textsuperscript{a}</td>
</tr>
<tr>
<td>(z_t)</td>
<td>-0.000026</td>
<td>-2.65 \textsuperscript{a}</td>
<td>0.000068</td>
<td>6.42 \textsuperscript{a}</td>
</tr>
<tr>
<td>(z_{t-1})</td>
<td>-0.000010</td>
<td>-0.63</td>
<td>0.000030</td>
<td>1.64 \textsuperscript{c}</td>
</tr>
<tr>
<td>(z_{t-2})</td>
<td>0.000011</td>
<td>0.65</td>
<td>0.000032</td>
<td>1.65 \textsuperscript{c}</td>
</tr>
<tr>
<td>(z_{t-3})</td>
<td>0.000023</td>
<td>1.50</td>
<td>0.000040</td>
<td>2.19 \textsuperscript{b}</td>
</tr>
<tr>
<td>(z_{t-4})</td>
<td>0.000005</td>
<td>0.41</td>
<td>-0.000010</td>
<td>0.72</td>
</tr>
<tr>
<td>(b_{t-k})</td>
<td>0.0289</td>
<td>1.10</td>
<td>-0.0763</td>
<td>-2.46 \textsuperscript{b}</td>
</tr>
<tr>
<td>(T_g - t + k)</td>
<td>-0.000013</td>
<td>-1.92 \textsuperscript{c}</td>
<td>-0.000008</td>
<td>-1.10</td>
</tr>
</tbody>
</table>

Notes:
\textsuperscript{a} significant at the 1% level; \textsuperscript{b} significant at the 5% level; \textsuperscript{c} significant at the 10% level.
Table 5.4 Results of DCC-TGARCH(1, 1), 28/10/1986 – 17/10/1997 (continued)

Conditional variance equation: 
$$\sigma_i^2 = \psi_{i0} + \psi_{i1} e_{i,j-k}^2 + \psi_{i2} e_{i,j-k}^2 I_{i,j-k} + \psi_{i3} \sigma_{i,j-k}^2$$

Conditional covariance equation: 
$$Q_i = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 \Xi_{i,k} + \theta_2 Q_{i,k}$$

Conditional correlation equation: 
$$R_i = Q_i^{-1} Q_{i,k}^{-1}$$

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Variable</th>
<th>Spot Variance Equation (5.9)</th>
<th>Futures Variance Equation (5.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coefficient</td>
<td>t-statistics</td>
</tr>
<tr>
<td>$\psi_{i0}$</td>
<td>Intercept</td>
<td>0.000002</td>
<td>6.53 a</td>
</tr>
<tr>
<td>$\psi_{i1}$</td>
<td>$e_{i,j-k}^2$</td>
<td>0.0567</td>
<td>9.66 a</td>
</tr>
<tr>
<td>$\psi_{i2}$</td>
<td>$e_{i,j-k}^2 I_{i,j-k}$</td>
<td>0.0408</td>
<td>4.81 a</td>
</tr>
<tr>
<td>$\psi_{i3}$</td>
<td>$\sigma_{i,j-k}^2$</td>
<td>0.8933</td>
<td>103.13 a</td>
</tr>
<tr>
<td>$\psi_{i1} + \psi_{i2}$</td>
<td></td>
<td>0.0975</td>
<td></td>
</tr>
<tr>
<td>$\psi_{i1} + \psi_{i2} / 2 + \psi_{i3}$</td>
<td></td>
<td>0.9704</td>
<td></td>
</tr>
</tbody>
</table>

Conditional Covariance Equation (5.10)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Xi_{i,k} \Xi_{i,k}$</td>
<td>0.0751</td>
</tr>
<tr>
<td>$Q_{i,k}$</td>
<td>0.9111</td>
</tr>
<tr>
<td>$\theta_1 + \theta_2$</td>
<td>0.9862</td>
</tr>
</tbody>
</table>

Note:
* significant at the 1% level.
The maximum likelihood estimators and the corresponding \( t \)-statistics for the DCC-TGARCH (1,1) model for the second period 27/10/1997–30/12/2005\(^{25}\) are reported in Table 5.5. Here, the evidence on short-run lead-lag relationships between spot and futures prices is different from that found using linear VECM in Chapter 4. While the linear VECM analysis suggests that there is only unidirectional causality from spot to futures, the analysis based on the DCC-TGARCH (1, 1) reveals bidirectional causality-in-mean between spot and futures markets, though causality from spot to futures seems to be stronger than the reverse. As anticipated, and consistent with earlier results, past price movements in the spot/futures market have positive impact on current price movements in the futures/spot market. This is indicated by the fact that all the significant coefficients of lagged cross market returns are positive. While the linear VECM analysis suggests that only the futures return series has a mean-reverting pattern, the DCC-TGARCH suggests that both spot and futures return series tend to be mean-reverting, indicated by the negative and significant impact of lagged own market returns on the current return. As analysed above, mean-reversion is required for the market to be efficient in the long run, though overreactions may be allowed in the short run. Further, the futures return series shows stronger evidence of mean-reversion than the spot return series, reflecting possibly stronger overreaction to new information in the futures price, due to the ease of trading in futures markets, or the fact that bid/ask bounce affects the futures price but not the index.

As reported in Chapter 4, \( k_t \) and \( k_{t-3} \) are positively significant at the 10% level in spot equation, though here \( k_{t-3} \) is also significant at the 10% level in the futures equation. This indicates that the passage of time had a positive, though quantitatively small, effect on price changes in both markets during this sample period. The findings of the linear VECM and DCC-TGARCH analyses regarding \( z_t \) and its lags seem to be inconsistent. While the linear VECM analysis in Chapter 4 indicates positively significant (at the 10% level) \( z_{t-4} \) in the futures equation, here none of the \( z_{t-j} \) (\( j = 0, 1, 2, 3, 4 \)) is significant in the futures equation, though \( z_t \) is found to be negatively significant (at the 5% level) in the spot equation. This seems to be puzzling at first

---

\(^{25}\) The maximum likelihood estimation of the DCC-TGARCH (1,1) model during this sample period was carried out under the assumption of multivariate \( t \)-distribution for spot and futures returns, because the estimation failed to get convergence under the multivariate normal distribution. This indicates that multivariate \( t \)-distribution is more appropriate to depict the characteristics of the data during this sample period.
glance but it does suggest a wider spread between futures and spot prices at contract rollovers, reflecting the consequence of an ‘artificial’ jump in the futures price when switching from one contract to another. As analysed above, while ‘artificial’ jumps should have occurred in the futures price, the observed negative effect of $z_t$ on spot price changes may actually reflect a transmission of futures price jumps at contract rollovers to the spot market.

The coefficient on the lagged basis in the spot equation is found to be positively significant, but only very weakly significant at the 10% level (during the first sub-sample, the lagged basis in the futures equation is strongly significant at the 5% level, see Table 5.4). This implies that compared to obvious evidence of divergence from long-run equilibrium during the first sub-sample, there is very weak evidence of divergence from long-run equilibrium in the second period. Another implication is that the two markets have been more closely linked with each other during the second sub-sample than during the first. Unlike the linear VECM analysis, the DCC-TGARCH analysis suggests that the time to maturity term is negative and significant in the futures equation, although only at the 10% level. It therefore provides weak support to the conjecture that the basis needs to be adjusted by the time to maturity to induce stationarity for individual futures contracts.

For the variance equation, the conditions of $\psi_{i0} > 0$, $\psi_{i1} + \psi_{i2} > 0$ and $0 < \psi_{i3} < 1$ are all satisfied, ensuring a positive definite conditional variance. Both spot and futures variances are highly persistent. The long-run persistence in variance, measured by $\psi_{i1} + \psi_{i2} / 2 + \psi_{i3}$, is 0.990625 for spot returns and 0.991067 for futures returns. But the condition $\psi_{i1} + \psi_{i2} / 2 + \psi_{i3} < 1$ is satisfied for both return series, ensuring the existence of the second moment of $\epsilon_t$. As anticipated, the parameter of the term measuring the leverage effect is positive and significant for both spot and futures variances, suggesting that the volatilities of both markets react more strongly to bad news than to good news. The coefficients in the covariance equation are highly significant, indicating time-varying correlation between the two markets. The persistence in the conditional correlation, measured by $\theta_1 + \theta_2$, is 0.960361, suggesting high persistence in the conditional correlation.
Table 5.5 Results of DCC-TGARCH(1, 1), 27/10/1997 - 30/12/2005

\[
\Delta_k s_t = \alpha_0 + \sum_{j=1}^{M_1} \alpha_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{M_1} \alpha_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{M_2} \alpha_{3j} k_{t-j} + \sum_{j=0}^{M_2} \alpha_{4j} z_{t-j} + \\
\alpha_5 b_{t-k} + \alpha_6 (T_k - t + k_i) + \epsilon_{1t}.
\]

\[
\Delta_k f_t = \beta_0 + \sum_{j=1}^{N_1} \beta_{1j} \Delta_k s_{t-j} + \sum_{j=1}^{N_1} \beta_{2j} \Delta_k f_{t-j} + \sum_{j=0}^{N_2} \beta_{3j} k_{t-j} + \sum_{j=0}^{N_2} \beta_{4j} z_{t-j} + \\
\beta_5 b_{t-k} + \beta_6 (T_k - t + k_i) + \epsilon_{2t}.
\]

<table>
<thead>
<tr>
<th>variable</th>
<th>Spot Return Equation (5.1a)</th>
<th>Futures Return Equation (5.1b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.000737 (-0.51)</td>
<td>-0.000902 (-0.60)</td>
</tr>
<tr>
<td>(\Delta_k s_{t-1})</td>
<td>-0.180083 (-1.93)</td>
<td>0.434034 4.65</td>
</tr>
<tr>
<td>(\Delta_k s_{t-2})</td>
<td>-0.242736 (-2.38)</td>
<td>0.176173 1.72</td>
</tr>
<tr>
<td>(\Delta_k s_{t-3})</td>
<td>-0.012389 (-0.12)</td>
<td>0.261806 2.55</td>
</tr>
<tr>
<td>(\Delta_k s_{t-4})</td>
<td>0.058810 0.61</td>
<td>0.188131 1.96</td>
</tr>
<tr>
<td>(\Delta_k s_{t-5})</td>
<td>0.044969 0.57</td>
<td>0.078775 0.96</td>
</tr>
<tr>
<td>(\Delta_k f_{t-1})</td>
<td>0.172409 1.89</td>
<td>-0.433061 -4.72</td>
</tr>
<tr>
<td>(\Delta_k f_{t-2})</td>
<td>0.196469 1.98</td>
<td>-0.221423 -2.20</td>
</tr>
<tr>
<td>(\Delta_k f_{t-3})</td>
<td>-0.033108 -0.33</td>
<td>-0.303997 -3.00</td>
</tr>
<tr>
<td>(\Delta_k f_{t-4})</td>
<td>-0.049506 -0.53</td>
<td>-0.179961 -1.89</td>
</tr>
<tr>
<td>(\Delta_k f_{t-5})</td>
<td>-0.066974 -0.87</td>
<td>-0.104971 -1.31</td>
</tr>
<tr>
<td>(k_t)</td>
<td>0.000425 1.66</td>
<td>0.000402 1.57</td>
</tr>
<tr>
<td>(k_{t-1})</td>
<td>0.000104 0.44</td>
<td>0.000100 0.41</td>
</tr>
<tr>
<td>(k_{t-2})</td>
<td>-0.000260 -1.03</td>
<td>-0.000214 -0.83</td>
</tr>
<tr>
<td>(k_{t-3})</td>
<td>0.000485 1.89</td>
<td>0.000502 1.89</td>
</tr>
<tr>
<td>(k_{t-4})</td>
<td>0.000258 1.07</td>
<td>0.000277 1.11</td>
</tr>
<tr>
<td>(k_{t-5})</td>
<td>0.000022 0.09</td>
<td>0.000080 0.31</td>
</tr>
<tr>
<td>(z_t)</td>
<td>-0.000037 -2.01</td>
<td>0.000006 0.34</td>
</tr>
<tr>
<td>(z_{t-1})</td>
<td>-0.000013 -0.72</td>
<td>0.000020 1.11</td>
</tr>
<tr>
<td>(z_{t-2})</td>
<td>-0.000010 -0.69</td>
<td>0.000011 0.74</td>
</tr>
<tr>
<td>(z_{t-3})</td>
<td>-0.000015 -0.95</td>
<td>0.000001 0.04</td>
</tr>
<tr>
<td>(z_{t-4})</td>
<td>0.000008 0.51</td>
<td>0.000014 0.92</td>
</tr>
<tr>
<td>(z_{t-5})</td>
<td>0.000009 0.71</td>
<td>0.000009 0.73</td>
</tr>
<tr>
<td>(b_{t-k})</td>
<td>0.082300 1.72</td>
<td>0.015680 0.33</td>
</tr>
<tr>
<td>(T_k - t + k_i)</td>
<td>-0.000011 -1.53</td>
<td>-0.000014 -1.90</td>
</tr>
</tbody>
</table>

Notes:

a significant at the 1% level; b significant at the 5% level; c significant at the 10% level.
Table 5.5 Results of DCC-TGARCH(1, 1), 27/10/1997 – 30/12/2005 (continued)

Conditional variance equation: $\sigma_{it}^2 = \psi_{i0} + \psi_{i1}e_{i,j-k}^2 + \psi_{i2}e_{i,j-k}^2I_i + \psi_{i3}\sigma_{i,j-k}^2$

Conditional covariance equation: $Q_i = (1 - \theta_1 - \theta_2)\bar{Q} + \theta_1\Xi_{i-k}\Xi_{i-k} + \theta_2Q_{i-k}^{-1}$

Conditional correlation equation: $R_i = Q_i^{-1}Q_{i-k}Q_{i-k}^{-1}$

<table>
<thead>
<tr>
<th>coefficient</th>
<th>variable</th>
<th>Spot Variance Equation (5.4)</th>
<th>Futures Variance Equation (5.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{i0}$</td>
<td>Intercept</td>
<td>0.000001 3.45 $^a$</td>
<td>0.000001 3.61 $^a$</td>
</tr>
<tr>
<td>$\psi_{i1}$</td>
<td>$e_{i,j-k}^2$</td>
<td>0.033346 3.00 $^a$</td>
<td>0.030429 2.77 $^b$</td>
</tr>
<tr>
<td>$\psi_{i2}$</td>
<td>$e_{i,j-k}^2I_i$</td>
<td>0.067243 4.84 $^a$</td>
<td>0.077126 5.30 $^a$</td>
</tr>
<tr>
<td>$\psi_{i3}$</td>
<td>$\sigma_{i,j-k}^2$</td>
<td>0.923657 105.07 $^a$</td>
<td>0.922075 104.87 $^a$</td>
</tr>
<tr>
<td>$\psi_{i1} + \psi_{i2}$</td>
<td></td>
<td>0.100589</td>
<td>0.107555</td>
</tr>
<tr>
<td>$\psi_{i1} + \psi_{i2} / 2 + \psi_{i3}$</td>
<td></td>
<td>0.990625</td>
<td>0.991067</td>
</tr>
</tbody>
</table>

Conditional Covariance Equation (5.5)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Xi_{i-k}\Xi_{i-k}$</td>
<td>0.110125 7.91 $^a$</td>
</tr>
<tr>
<td>$Q_{i-k}$</td>
<td>0.850236 40.30 $^a$</td>
</tr>
<tr>
<td>$\theta_1 + \theta_2$</td>
<td>0.960361</td>
</tr>
</tbody>
</table>

Note:

$^a$ significant at the 1% level.
To see how the variance of price changes relates to price changes themselves, spot conditional variance is plotted against spot price and futures conditional variance is plotted against future price in both sub-samples (see Figure 5.7, Figure 5.8, Figure 5.9 and Figure 5.10). For both spot and futures in both sub-samples, in general there is a clear pattern of the conditional variance that tends to fall when the price rises and to rise when the price falls. This can be seen most clearly from the 1987 crash period (in the first sub-sample) and the three years recession period (2000–2002) following the collapse of the dot-com bubble (in the second sub-sample), when both spot and futures markets were remarkably volatile. On the other hand, during 1993–1997 (in the first sub-sample) and during 2003–2005 (in the second sub-sample) when the market has been rising steadily, both spot and futures markets were very stable with low volatility. This provides strong support to the hypothesis that the variance of asset price changes tends to be negatively correlated with the asset price changes (Wong and Vlaar, 2003).

It can also be seen from the time plot of futures conditional variance versus spot conditional variance (Figure 5.11 for the first sub-sample and Figure 5.12 for the second sub-sample) that the futures variance is in general higher than the spot variance (especially during the first sub-sample), consistent with evidence widely reported for many markets in the literature (see Sutcliffe (2006) for a summary). The dynamic conditional correlations between spot and futures are plotted in Figure 5.13 and Figure 5.14 respectively for the two sub-samples. It can be seen that the conditional correlation between spot and futures markets has been varying over time. Although the two markets were in general very highly positively correlated during both periods they sometimes drifted far apart from each other, as indicated by several obvious spikes in the time plot of their conditional correlation. This suggests that index arbitrageurs may not be continuously active (for reasons such as lack of capital, high arbitrage risk and limitations on trading), resulting in a loose link between the two markets. The conditional correlation displays an overall rising trend during the first period (28/10/1986–17/10/1997) with a higher level maintained during the second period (27/10/1997–30/12/2005), suggesting a closer link between the two markets induced possibly by more efficient arbitrage activities during the more recent history.
Figure 5.7 FTSE100 spot conditional variance versus spot price, 03/11/1986 – 17/10/1997
Figure 5.8 FTSE100 futures conditional variance versus future price, 03/11/1986 – 17/10/1997
Figure 5.11 Futures conditional variance versus spot conditional variance, 03/11/1986 – 17/10/1997
Figure 5.12 Futures conditional variance versus spot conditional variance, 03/11/1997 – 30/12/2005
Figure 5.13 Correlation between spot and futures markets, 03/11/1986 – 17/10/1997
Figure 5.14 Correlation between spot and futures markets, 03/11/1997 – 30/12/2005
5.4.3 CCF test results on causality-in-variance

It can be seen from Figure 5.13 and Figure 5.14 that over both periods, FTSE100 spot and futures markets are closely linked to each other (the contemporaneous conditional correlation between the two markets is close to unity, especially during the second period). But is this the whole story? Given strong evidence of bidirectional causality-in-mean between the two markets in both sample periods, it is natural to ask whether there is also evidence of causality-in-variance. Cross correlation functions for up to 5 leads and lags of squared-standardized residuals estimated from the DCC-TGARCH (1,1) model were used to examine causality-in-variance. Significant \( CCF^2(k) \) at any leads or lags would suggest causality-in-variance between the two markets.

There are 2776 observations in the period 28/10/1986–17/10/1997 and 2051 observations in the period 27/10/1997–30/12/2005. Therefore, for the first period, \( CCF^2(k) \) is classed as significant at the 1% level if it is outside the band \([-0.0486, +0.0486]\), significant at the 5% level if it is outside the band \([-0.0372, +0.0372]\) and significant at the 10% level if it is outside the band \([-0.0311, +0.0311]\). For the second period, \( CCF^2(k) \) is classed as significant at the 1% level if it is outside the band \([-0.0565, +0.0565]\), significant at the 5% level if it is outside the band \([-0.0433, +0.0433]\) and significant at the 10% level if it is outside the band \([-0.0362, +0.0362]\). The CCF test results are reported in Table 5.6. It can be seen that \( CCF^2(0) \), the contemporaneous correlation between the two markets, is highly significant at the 1% level (0.7924 during the first period and 0.9506 during the second period). Therefore, consistent with what has been suggested by the dynamic conditional correlation (DCC) estimated from the DCC-TGARCH model, the CCF test also suggests that the two markets are more closely contemporaneously correlated during the more recent period. There seems to be no strong evidence of causality-in-variance between the two markets, except that there is a fairly weak evidence of causality-in-variance from spot to futures in the first period, indicated by \( CCF^2(-4) \) that is significant at the 10% level. A possible explanation is that since it is the variance of price changes, not simple price changes, that is more related to the rate of information flow to the market (Chan et al., 1991), and since most information is impounded into spot and futures markets simultaneously (as indicated by high contemporaneous correlation between the two markets, especially during the
second period), there should be no strong evidence of lead-lag relationship between spot and futures variances in response to information shocks, though lead-lag relationships do exist between spot and futures price changes themselves.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>CCF²(k)</td>
<td>CCF²(k)</td>
</tr>
<tr>
<td>-5</td>
<td>-0.0188</td>
<td>0.0148</td>
</tr>
<tr>
<td>-4</td>
<td>0.0359 *</td>
<td>-0.0103</td>
</tr>
<tr>
<td>-3</td>
<td>-0.0169</td>
<td>-0.0025</td>
</tr>
<tr>
<td>-2</td>
<td>-0.0017</td>
<td>0.0082</td>
</tr>
<tr>
<td>-1</td>
<td>0.0084</td>
<td>-0.0004</td>
</tr>
<tr>
<td>0</td>
<td>0.7924 *</td>
<td>0.9506 *</td>
</tr>
<tr>
<td>1</td>
<td>-0.0213</td>
<td>-0.0078</td>
</tr>
<tr>
<td>2</td>
<td>0.0124</td>
<td>0.0090</td>
</tr>
<tr>
<td>3</td>
<td>-0.0226</td>
<td>0.0050</td>
</tr>
<tr>
<td>4</td>
<td>-0.0044</td>
<td>-0.0182</td>
</tr>
<tr>
<td>5</td>
<td>-0.0258</td>
<td>0.0165</td>
</tr>
</tbody>
</table>

Notes:

CCF²(k) represents correlation between ε₁ₙ² and ε₂ₙ₋ₖ, indicating causality-in-variance from spot to futures when k < 0 and from futures to spot when k > 0.

* significant at the 1% level; * significant at the 10% level.

5.5 Discussion and Conclusions

A comprehensive re-examination of the dynamics between the FTSE100 spot index and futures price using the DCC-TGARCH analysis and the two-step CCF test for causality-in-variance in this chapter offer several new and important findings. First, while strong evidence of bidirectional causality-in-mean between FTSE100 spot and futures markets is documented over both sub-samples, there is no strong evidence of
causality-in-variance between the two markets. Causality from spot to futures could arise because the FTSE100 index is a quote-based index, which can reflect information more quickly than the transactions-based futures price. Causality from futures to spot could arise because of lower costs of trading in the futures market than in the spot market. There is evidence that the price relationship between the two markets differs from their variance relationship. One possible explanation is that the price changes and the variance of price changes are related to different aspects of investors' assessment of new information. Price changes themselves may be more related to investors' interpretations of new information while the variance of these price changes may be more related to investors' confidence in their beliefs about the new information. As a result, depending on how participants in the two markets interpret and assess new information, the price relationship between the two markets and the variance relationship between them may not share the same pattern. Bidirectional causality-in-mean implies that beliefs in one market might be transmitted to participants in the other market slightly later, and vice versa. This is possible if different people use spot and futures markets for different trading purposes. Special trading purposes could result in slightly lagged reaction of participants in one market to the same potential information compared to the reaction of participants in the other market. Alternatively, the occurrence together of causality-in-mean and non-causality-in-variance could simply suggest that investors in both markets interpret the same information at the same speed and hold it with equal confidence, but that there are microstructural mechanisms and equilibrium links between the two markets that ensure lagged price adjustments even after the direct price impact of information has been fully understood by investors in each market separately.

The finding that there is no obvious evidence of causality-in-variance (especially during the second sub-sample) between the two markets does not mean constant variance. In fact, conditional variances in both spot and futures markets display strong evidence of predictability – both spot and futures volatilities tend to be highly persistent. This implies that the linear VECM (which ignores time-varying volatilities) captures only part of the dynamics in the spot and futures prices and therefore may offer only inconclusive inference. The evidence on conditional correlation between FTSE100 spot and futures markets indicates that most of the time they are closely linked by arbitragers, but that the link between them also becomes occasionally loose, possibly due to
unfavourable conditions for arbitrage. This also means that fund managers using stock index futures to protect their portfolios against market risks need to watch closely both the volatility pattern and the dynamic correlation between the index futures market and the underlying spot market, hence to update the hedging ratio accordingly. The evidence of no obvious causality-in-variance is consistent with the high level of conditional correlation between the two markets, both suggesting that most information is impounded into the two markets simultaneously, especially during the more recent sample. In fact, there is no clear evidence of divergence from equilibrium during the second sample, further suggesting a closer link between the two markets. It may be that the new order-driven trading system (SETS) employed by the LSE and the new electronic platform (LIFFE CONNECT) employed by LIFFE introduce more efficient environments for market participants in both markets. In addition, the introduction by Barclays Global Investors (BGI) of the exchange traded fund (ETF) in April 2000 to track the FTSE 100 index allows purchase and sale of FTSE100 index portfolio in spot market more quickly, easily and cheaply than before.

Since the DCC-TGARCH model by itself does not allow for inter-market dependence between the volatility series, it could fail to capture some aspects of the volatility dynamics in the individual series. Furthermore, while it has the advantage of being easy to implement, it suffers from the problem that two-step estimation may cause loss of efficiency, because the optimal parameters estimated in the first step of the estimation procedure are not necessarily optimal in the second step (Wong and Vlaar, 2003). Similarly, the two-step CCF test for causality-in-variance based on the first-step estimation of the DCC-TGARCH may also be inefficient. Therefore, the results of the DCC-TGARCH and the inference of no obvious causality-in-variance between FTSE100 spot and futures markets need to be interpreted with caution. As the measure of volatility used is a critical factor for examining volatility, further research using different measures of volatility, including other multivariate GARCH models, may offer additional insights into the persistent feature of variance-covariance processes as well as causality-in-variance for FTSE100 spot and futures prices.
Chapter 6 Index Arbitrage and Nonlinear Dynamics between FTSE100 Spot and Futures Prices – Threshold Cointegration Analysis

6.1 Introduction

In previous chapters, the dynamics between FTSE100 spot and futures prices were studied within a VECM framework, as has been routine in the literature. An assumption implied by the standard error correction model of Engle and Granger (1987) is that the market is always frictionless, involving no transaction cost and no risk, and hence arbitrage trading is possible in every time period, creating continuous movement towards equilibrium. However, it has long been recognized that the presence of transaction costs, capital constraints, execution risk, interest rate and dividend risks could allow futures prices to fluctuate within a band around the theoretical value defined by the cost of carry model (the 'no-arbitrage window') without triggering profitable arbitrage (Mackinlay and Ramaswamy, 1988; Yadav and Pope, 1994). Profitable arbitrage opportunities arise only if divergence from equilibrium, or mispricing, is sufficiently large to compensate for transaction costs, execution risks and other constraints (see Figure 6.1 for a stylised representation). The existence of transaction costs for example implies that adjustment towards equilibrium driven by arbitrageurs, or the error correction effect, is discontinuous — adjustment will occur only outside the no-arbitrage window, where arbitrage revenues exceed the transaction costs of arbitrage operations. Threshold cointegration, first introduced by Balke and Fomby (1997), combines cointegration with a nonlinear error correction effect. The arbitrage-induced error correction effect is allowed to be inactive inside the no-arbitrage window, only becoming active when the system moves sufficiently far from equilibrium. The threshold error correction model of Balke and Fomby (1997) therefore allows different levels of past mispricing to have different effects on current price dynamics.
Given that arbitrage transactions affect market dynamics, the model for spot and futures prices should vary over time, depending on the presence or absence of arbitrage activities. A multivariate threshold error correction model was therefore used in this chapter to study regime-dependent price dynamics between FTSE100 spot and futures prices. If transaction costs and other factors allow the futures price to fluctuate within a band around its theoretical value without triggering profitable arbitrage, it is of particular interest to examine exactly how far the futures contract can be mispriced without inducing profitable arbitrage opportunities. The width of the no-arbitrage band can be estimated directly using substantive information about transaction costs. For example, Yadav and Pope (1990, 1994) and Butterworth and Holmes (2000) have studied the mispricing of index futures contracts and the implied arbitrage profitability in the UK market by comparing the mispricing with round-trip transaction costs faced by different categories of index arbitrageurs. However, these subjectively-estimated transaction costs may not be representative of the true average transaction costs for most arbitrageurs. Therefore, percentage mispricing, first introduced by Mackinlay and Ramaswamy (1988), is used in this study as the threshold variable for the threshold error correction model. The threshold values yielded from the estimation should reflect...
the average transaction costs faced by most arbitrageurs. We would claim that the transaction costs implied by such estimation are more reliable and fair than subjective estimation of average transaction costs for most arbitrageurs because the transaction costs implied by the threshold values are based on an observed market pattern — that is, the price dynamics switch between regimes depending on the presence or absence of arbitrage activities. Furthermore, substantive information about transaction costs is only used to specify a reasonable range over which to search for the threshold values.

This rest of this chapter is organized as follows: Section 6.2 introduces a new variable, the percentage mispricing of futures contracts, and explains why percentage mispricing is used as the threshold variable. It also discusses briefly the selection of two sub-samples in the context of threshold cointegration analysis. In section 6.3, the Tsay test (arranged autoregression test) (Tsay, 1989) is used to detect threshold nonlinearity in the FTSE100 basis, which is used in this study (as in many others) as the error correction term. Given the evidence of threshold nonlinearity in the FTSE100 basis, a three-regime threshold VECM is used to examine nonlinear cointegrating relationships and regime-dependent price dynamics between the FTSE100 spot and futures markets. In section 6.4, the least squares estimates of the threshold VECM for both sub-samples are reported and analysed. Section 6.5 draws conclusions.

6.2 The threshold variable and sub-samples

Whether or not an observed deviation from equilibrium, or futures 'mispricing', represents a profitable arbitrage opportunity is determined by the costs involved in arbitrage transactions. The larger the magnitude of the mispricing, the more likely it is to represent a profitable arbitrage opportunity, after accounting for transaction costs. Mispricing of futures contracts is therefore a reasonable threshold variable to be used in a threshold cointegration analysis of spot and futures price dynamics. In the literature, the lagged mispricing error (in logarithms) $y_{t-1} = \ln F_{t-1} - \ln S_{t-1} - (r - d)_{t-1}(T - t + 1)$ has normally been used as the threshold variable (Tsay, 1998; Martens et al., 1998). (Here $r$ is the risk-free interest rate and $d$ is the index dividend yield.) It is important to specify a reasonable a range for searching the threshold values. For estimation, Tsay (1998) assumes the true threshold values to lie within a range based on the empirical range of $y_{t-1}$. The limitation of this method is that the threshold is not chosen
endogenously by the model but demands an initial input and 'plausibility' checks by the researcher, essentially a search procedure. It appears to be *ad hoc* and lacking in firm theoretical support. In this chapter, we used instead the 'percentage mispricing' of the futures contract as the threshold variable. The percentage mispricing of the futures contract, first introduced by Mackinlay and Ramaswamy (1988), is defined to be:

\[ X_t = \frac{(F_t - F_t^*)}{S_t}, \]

where \( F_t \) is the actual futures price at time \( t \) of a contract maturing at time \( T \) and \( F_t^* \) is the theoretically 'fair' futures price defined by the cost of carry model \( F_t^* = S_t e^{(r-d)(T-t)} \). Percentage mispricing is normalized by the value of the index, \( S_t \), and is directly comparable to transaction costs because the costs associated with index arbitrage are normally defined as a proportion of the index value. Therefore, using percentage mispricing as the threshold variable has strong advantages. It makes it possible to exploit substantive information about the transaction costs associated with arbitrage and thus to specify less arbitrarily a more reasonable range over which to search for the true threshold values. Furthermore, the estimated threshold values give immediate information about the average transaction costs faced by arbitrageurs active in the market, which in turn provides a valuable reference for other researchers as well as for practitioners.

To calculate percentage mispricing, the FTSE100 index and futures price are needed, as well as the risk free interest rate and forecasted dividend yield of the index. As analysed in Chapter 4, the time at which the daily settlement price for FTSE100 futures is fixed at LIFFE is not necessarily the same as the daily closing time on LSE for trading FTSE100 component shares. The possible asynchronicity between the prices of FTSE100 futures and the underlying shares might produce noise in mispricing estimates, though it should not lead to *systematic* errors (Yadav and Pope, 1990). Also, with no access to a complete term structure of interest rates that would allow matching of the maturity of each loan with the time-dependent maturity of the futures contract, LDMON (London Discount Market Overnight Rate, \( r' \)) was used in this study to proxy for the interest rate for a loan with maturity \( (T-t) \). Finally, the realised *ex post* dividend yield for the FTSE100 index was used as a proxy for the forecasted dividend yield. These could also introduce bias in estimating the fair futures price and the true futures mispricing. However, it seems that these possible biases are unavoidable and shared by all research on this issue. Given these biases, we argue that using percentage
mispricing as the threshold variable, which allows us to define a reasonable range over which to search for the true threshold values based on substantive information about arbitrage transaction costs, is much better than the search procedure and subjective definition of the range by Tsay (1998).

As analysed before, the linkage between futures and spot markets is maintained by arbitrage, but this is affected to a great extent by the costs of trading in the two markets. The usual hypothesis is that the presence of transaction costs in the two markets has the effect of allowing the futures price to fluctuate within a band around its theoretical value (or a no-arbitrage window) without representing a profitable arbitrage opportunity. The width of the no-arbitrage window is determined by transaction costs of the most favourably situated arbitrageurs in the market. Outside the no-arbitrage window, arbitrage revenue exceeds the costs associated with arbitrage transactions, hence arbitrage is active and determines spot-futures price dynamics; inside the no-arbitrage window, however, arbitrage revenue is less than transaction costs involved in arbitrage, hence arbitrage is inactive and spot-futures price dynamics are not affected by arbitrage. The development of threshold cointegration analysis is indeed based on the hypothesis that arbitrage is not always active, due to the existence of transaction costs. Therefore, nonlinear spot-futures price dynamics switches from one regime to another, depending on the presence or absence of arbitrage with different regimes separated by thresholds.

From the above analysis we can see that the to-be-estimated thresholds in a threshold cointegration analysis are actually the upper and lower bounds of the no-arbitrage window. Given that the no-arbitrage window is basically determined by transaction costs, the location of the thresholds should also rely primarily on transaction costs. The lower are the transactions costs, the smaller is the degree of futures mispricing allowed by arbitrage, and the closer are the threshold values to zero, *ceteris paribus*. It is therefore especially important to perform a threshold cointegration analysis over a time period with a relatively stable transactions cost structure so as to permit identification of a stable no-arbitrage window and stable threshold values. In this sense, the two sub-samples defined earlier are plausible and also reasonable for this study: during the first sub-sample 28/10/1986–17/10/1997, FTSE100 securities were traded through a dealership system and FTSE100 futures were traded through 'open outcry'; during the second sub-sample 27/10/1997–30/12/2005, FTSE100 securities were traded through
an order-driven system and FTSE100 futures were traded on an electronic platform (only after 30/11/1998). It is important to note that transaction costs were influenced by the move to the new trading systems, and are lower in the second sub-sample than that in the first in both markets. Indeed, our findings in Chapter 4 provide empirical evidence of the effect of such microstructure transformations: during the first sub-sample an arbitrage-induced cost of carry relationship is invalid (indicating less active arbitrage) due probably to high transaction costs in this period; during the second sub-sample an arbitrage-induced cost of carry relationship tends to be valid (indicating more active arbitrage) as a result of lower transaction costs in this period. There may also have been other changes in transactions costs that could have affected the estimated thresholds so that the latter can only be interpreted as an average of transaction costs faced by arbitrageurs during each sub-period.

It is also necessary to report here that the dataset used in this study was adjusted by removing the effect of the 1987 stock market crash, as in Chapter 5. That is, based on the SUR estimation of equations (4.11a) and (4.11b) of Chapter 4, the spot return series were adjusted by subtracting the estimated coefficients $\alpha_{ij}$ from the original spot return values on relevant dates, and the futures return series were adjusted by subtracting the estimated coefficients $\beta_{ij}$ from the original futures return values on relevant dates. Since the dummy variable (DM) and 4 lags of it conform to 5 days from 19/10/1987 to 23/10/1987, the adjusted spot and futures return series contain exactly the same data as the original spot and futures return series on all other days. The reason for the adjustment was to correct zero standard errors that were estimated for some parameters in the threshold VECM using the original dataset.

6.3 Methodology

6.3.1 Tsay test for threshold nonlinearity in the basis

The analysis in Chapter 4 suggests that the basis is a no-arbitrage link between spot and futures markets. It determines whether arbitrage opportunities are available. If the basis is too large or too small, as a result of serious divergence from the long-run equilibrium

\[ \text{\footnotesize 26 The threshold VECM was estimated for the period 30/11/1998 to 30/12/2005 (LIFFE replaced the old open outcry system by the new electronic platform on 30/11/1998). However, the estimation results show only a very small number of observations (32) in the upper regime, indicating that any inference regarding the upper regime and the upper threshold would be susceptible to small sample bias.} \]
between the two markets, arbitrage transactions should occur immediately to drive the price relation between the two markets back to equilibrium. In econometric terminology, the basis is the cointegrating series for futures and spot prices, and the basis should be a stationary process. Our preliminary data analysis shows that the basis for FTSE100 spot and futures prices is indeed stationary. However, as analysed before, the arbitrage-induced cointegrating process between spot and futures prices may not be continuous, depending on the presence or absence of arbitrage. Arbitrageurs may only enter into the market when divergence from equilibrium relationship between spot and futures markets is sufficiently large to offset transaction costs and potential arbitrage risks (Balke and Fomby, 1997). If the divergence from equilibrium is not large enough to induce profitable arbitrage, the dynamic relationship between futures and spot prices would be free from arbitrage effects, which means that futures and spot prices are free from the arbitrage induced cointegration constraint.

The above analysis suggests that if the arbitrage-induced cointegrating process between spot and futures prices is indeed discontinuous, it should be reflected in the dynamics of the cointegrating series (the basis), which should display a nonlinear process. As analysed in Chapter 3, here we assume homogeneity of traders, with identical transaction costs, who agree on the fair price of the futures contract and act simultaneously in a uniform manner in the presence of arbitrage opportunities (Tse, 2001; McMillan and Speight, 2006). Based on this, we assume a threshold nonlinear process for the basis defined by several threshold values if the basis is indeed nonlinear. A threshold nonlinear process for the basis would suggest that it is necessary to model FTSE100 spot-futures price dynamics using threshold cointegration analysis. It is therefore useful to check first whether the basis for FTSE100 spot and futures prices is a linear or a nonlinear process. The linear dynamics of the basis can be specified as a simple autoregressive process:

\[ b_t = \varphi_0 + \sum_{i=1}^{p} \varphi_i b_{t-i} + \mu_t \]  

(6.1)

Using lagged percentage mispricing, \( X_{t-d} \), as the threshold variable, the threshold nonlinear autoregressive process of the basis can be written as:
\[ b_t = \varphi_0^{(j)} + \sum_{i=1}^{p} \varphi_i^{(j)} b_{t-i} + \mu_t^{(j)} \]  

(6.2)

\[ C_{j-1} < X_{t-d} < C_j \]

Here \( d \) is a positive integer denoting the average time taken to execute an arbitrage trade; \( j = 1, \ldots, n \) indicates different regimes and \( -\infty = C_0 < C_1 < \ldots < C_n = \infty \) are the threshold values separating one regime from another (\( j = 1 \) indicates that the basis is a linear process as described in equation (6.1)). \( \mu_t^{(j)} \) are i.i.d. random error terms with zero mean and constant variance \( \sigma^2(j) \).

Threshold nonlinearity (equation (6.2)) versus linearity (equation (6.1)) of the basis was tested using the Tsay arranged autoregression test (Tsay, 1989). The arranged autoregression (equation (6.1)) involves rearranging the data according to the increasing order of the threshold variable, \( X_{t-d} \), and then estimating using recursive least squares (RLS). If the basis is linear, the predictive residuals of the RLS estimation of the arranged autoregression are white noise and uncorrelated with the regressors of regression (6.1). If however, the basis follows a nonlinear process, then the predictive residuals are no longer white noise, because the least squares estimator is biased. In this case, the predictive residuals are correlated with the regressors (Tsay, 1989, 1998). Threshold nonlinearity can therefore be tested by regressing the predictive residuals on the same regressors in regression (6.1) and calculating the \( F \) statistic of the resulting regression. If the \( F \) statistic exceeds the critical value of the \( F \) distribution, it implies threshold nonlinearity in the basis.

The lag order \( p \) of regression (6.1) was selected on the basis of the \( p \)th partial autocorrelation function of \( b_t \), represented by \( \varphi_p \). That is, \( \varphi_p \) was included in the regression if it was significant at the 10% level. 4 lags were selected for 28/10/1986–17/10/1997 and 5 lags for 27/10/1997–30/12/2005. Possible threshold lags of 1, 2, 3, 4 or 5 were used for both sub-samples. The optimal threshold lag \( d \) was selected such as to maximize the \( F \)-statistic, since the test is most powerful when the threshold lag \( d \) is correctly specified (Tsay, 1989, 1998). The results of the Tsay test for \( d = \{1, 2, 3, 4, 5\} \) for both sub-samples are reported in Table 6.1, where it can be seen that the \( p \) values of
the $F$-statistics are highly significant, rejecting the null of linearity in favour of threshold non-linearity at up to 5 lags. This suggests strongly that the basis is a threshold nonlinear process during both sub-samples. The test also indicates that for both sub-samples, the $F$-statistic has a maximum at $d = 1$, suggesting that this is the optimal threshold lag.

Table 6.1 Tsay test for threshold nonlinearity in the basis

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$F$-stat.</td>
<td>11.82924</td>
<td>3.21696</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.00000</td>
<td>0.00671</td>
</tr>
<tr>
<td></td>
<td>12.78686</td>
<td>4.55848</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.00000</td>
<td>0.00013</td>
</tr>
</tbody>
</table>

Notes: $d$ is the threshold lag; percentage mispricing $X$, is used as the threshold variable.

6.3.2 A three-regime threshold VECM for spot and futures returns

The above test shows that the basis is a threshold nonlinear process, which is, as analysed before, caused by discontinuous cointegrating process between spot and futures prices. An implication is that arbitrage has been switching between active and inactive status, depending on the regime in which the (percentage) mispricing lies. Given that arbitrage affects market dynamics, a threshold VECM should therefore be used to model the nonlinear price dynamics between spot and futures markets, with the percentage mispricing being a reasonable threshold variable. The next step is to consider the number of thresholds. This is difficult, since there exists no formal test using the classical approach (Martens et al., 1998). Tsay (1998) points out that, in some applications, past experience and substantive information may provide useful information on the choice of $k$. In others, the computational complexity and the data
may restrict $k$ to a small number such as 2 or 3. The primary interest here is to find the band around the no-arbitrage value within which arbitrageurs will not enter the market, and also for the sake of computational ease, three regimes are considered, where arbitrageurs react to large enough negative or positive mispricing in a previous period, i.e. $X_{t-d} \leq C_1$, or $X_{t-d} > C_2$, but not to mispricing in the inner regime when $C_1 < X_{t-d} \leq C_2$. Based on the VECM model developed in chapter 4, a three-regime threshold VECM (TVECM) as follows can be used to model the regime dependent spot-futures price dynamics:

if $X_{t-1} \leq C_1$

\[
\Delta_k s_t = \alpha_0^{(1)} + \sum_{j=1}^{M_1} \alpha_{1j}^{(1)} \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j}^{(1)} \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_{3j}^{(1)} k_{t-j} \\
+ \sum_{j=0}^{M_3} \alpha_{4j}^{(1)} z_{t-j} + \alpha_5^{(1)} b_{t-k} + \alpha_6^{(1)} (T_g - t + k_t) + e_t^{(1)} \tag{6.3.1a}
\]

if $C_1 < X_{t-1} \leq C_2$

\[
\Delta_k s_t = \alpha_0^{(2)} + \sum_{j=1}^{M_1} \alpha_{1j}^{(2)} \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j}^{(2)} \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_{3j}^{(2)} k_{t-j} \\
+ \sum_{j=0}^{M_3} \alpha_{4j}^{(2)} z_{t-j} + \alpha_5^{(2)} b_{t-k} + \alpha_6^{(2)} (T_g - t + k_t) + e_t^{(2)} \tag{6.3.2a}
\]
\[ \Delta_k f_t = \beta_0^{(2)} + \sum_{j=1}^{N_1} \beta_{1j}^{(2)} \Delta_k s_{t-j} + \sum_{j=1}^{N_2} \beta_{2j}^{(2)} \Delta_k f_{t-j} + \sum_{j=0}^{N_3} \beta_{3j}^{(2)} k_{t-j} \]
\[ + \sum_{j=0}^{N_4} \beta_{4j}^{(2)} z_{t-j} + \beta_{5}^{(2)} b_{t-k} + \beta_{6}^{(2)} (T_g - t + k_t) + e_{2t}^{(2)} \]  

(6.3.2b)

if \( X_{t-1} > C_2 \)

\[ \Delta_k s_t = \alpha_0^{(3)} + \sum_{j=1}^{M_1} \alpha_{1j}^{(3)} \Delta_k s_{t-j} + \sum_{j=1}^{M_2} \alpha_{2j}^{(3)} \Delta_k f_{t-j} + \sum_{j=0}^{M_3} \alpha_{3j}^{(3)} k_{t-j} \]
\[ + \sum_{j=0}^{M_4} \alpha_{4j}^{(3)} z_{t-j} + \alpha_{5}^{(3)} b_{t-k} + \alpha_{6}^{(3)} (T_g - t + k_t) + e_{1t}^{(3)} \]  

(6.3.3a)

\[ \Delta_k f_t = \beta_0^{(3)} + \sum_{j=1}^{N_1} \beta_{1j}^{(3)} \Delta_k s_{t-j} + \sum_{j=1}^{N_2} \beta_{2j}^{(3)} \Delta_k f_{t-j} + \sum_{j=0}^{N_3} \beta_{3j}^{(3)} k_{t-j} \]
\[ + \sum_{j=0}^{N_4} \beta_{4j}^{(3)} z_{t-j} + \beta_{5}^{(3)} b_{t-k} + \beta_{6}^{(3)} (T_g - t + k_t) + e_{2t}^{(3)} \]  

(6.3.3b)

Short run price dynamics between the two markets and long run error correction effects are allowed to be different in different regimes. In particular, the error correction term would be expected to have a much smaller effect in the inner regime than in the upper and lower regimes since arbitrage is expected to be inactive in the inner regime when mispricing is relatively small. Furthermore, there could be differences in arbitrage activities in the upper and lower regimes, since arbitrage in the lower regime involves short-selling stocks in the spot market, which is restricted for most market participants. In the LSE, only market makers have the privilege of borrowing stocks and are therefore free from short sale restrictions in the spot market. The three-regime TVECM for spot and futures returns was estimated and the thresholds selected with a grid search method using the criterion of the least sum of squared errors.
6.4 Empirical results

Based on the ACFs and PACFs of spot and futures returns (see Table 5.1) and to compare with previous chapters, 4 lags were included in the TVECM for the first subsample and 5 lags in the TVECM for the second (Ljung-Box test would be performed later on residuals to check for model adequacy). Estimating the threshold value requires two ranges to be specified, within which the lower and upper thresholds are located. The specification of these ranges is based on past experience and substantive information. Arbitrage transaction costs in London include round-trip cash and futures market trading costs, transaction taxes (0.5% of the index value after the 'Big Bang'), and the costs of borrowing fixed interest capital and index stocks. However, market makers who recycle stocks within seven days are exempt from transaction taxes for stocks in which they make a market. Thus, for the period 28/04/1986 – 23/03/1990, Yadav and Pope (1994) suggest two levels of transaction costs, 0.25% and 0.75%, corresponding to two broad categories of potential arbitrageurs. Category A arbitrageurs face transaction costs of about 0.25% because they are not subject to transaction taxes for arbitrage dealings. Examples of Category A arbitrageurs include market makers, those who are otherwise committed to enter or exit the market (due to e.g. portfolio insurance or tactical price based strategies) and use the futures market only as an intermediary, and those with existing arbitrage positions who seek to profitably rollover their position or to profitably unwind early. Category B arbitrageurs face transaction costs of 0.75% because they have to pay transaction taxes in their dealings. It is normally argued that the arbitrage window should depend on arbitrageurs with the lowest transaction costs (Yadav and Pope, 1990). Therefore, for the period 28/10/1986–17/10/1997, a lower threshold of $C_1 \in [-1.0, -0.05]$ and an upper threshold $C_2 \in [0.05, 1.0]$ were first assumed, which were considered wide enough to account for most active arbitrageurs in the market. A grid search with 300 points on each of the two intervals was used to select the lower and the upper thresholds. The basic principle for selecting the thresholds is to search for two critical values of percentage mispricing that define a most stable price dynamics within each of the three regimes separated by the two critical (threshold) values compared to that yielded from other values. This method selected $\hat{C}_1 = -0.965$ and $\hat{C}_2 = 0.291$. However, an error correction effect was found to be significant in the futures return equation in the estimated inner regime of
−0.965 < X_{t-1} ≤ 0.291, suggesting active arbitrage and hence implying that the inner regime was incorrectly estimated. These estimated threshold values and regimes cannot be explained by arbitrage and must be due to other noisy reasons that have the same effect of causing regime-different spot-futures price dynamics.

Given that the error correction term is found to be significant in the estimated inner regime, the true threshold values for \( C_1 \) and \( C_2 \) defined by arbitrage should be closer to zero than these first estimates. One reason for failing to find the true threshold values defined by arbitrage may be that the specified ranges for searching \( C_1 \) and \( C_2 \) were too wide. The search was therefore refined by gradually narrowing down the ranges from their outer ends. Using the criterion that no error correction effect should be found in the inner regime, reasonable ranges for searching the threshold values were found to be \( C_1 \in [-0.75, -0.05] \) and \( C_2 \in [0.05, 0.75] \), which should still be wide enough to account for most active arbitrageurs in the market. A grid search with 200 points on each of the two intervals was used to select the lower and the upper thresholds. The refined search selects \( \hat{C}_1 = -0.617 \) and \( \hat{C}_2 = 0.2845 \). The magnitude of the estimated lower threshold \( (\hat{C}_1) \) is larger than the magnitude of the estimated upper threshold \( (\hat{C}_2) \), implying that the transaction cost associated with arbitrage in the lower regime is higher than the transaction cost associated with arbitrage in the upper regime. Arbitrage in the lower regime is created by underpriced futures contracts and the arbitrage strategy involves buying futures and short selling spot. The higher transaction costs associated with arbitrage in the lower regime therefore reflects extra difficulties and costs imposed on arbitrage by short sale restrictions in the spot market, which is consistent with expectations.

We would expect arbitrage transaction costs to be lower in the second period than in the first because of the introduction of SETS (automated matching of buy and sell orders) at the LSE and the promotion of LIFFE CONNECT (the fully electronic derivatives trading platform) by LIFFE. Furthermore, as analysed before, exchange traded funds (ETFs) for the FTSE100 were introduced in April 2000, making it much easier and cheaper to trade the FTSE100 index in the spot market. Thus the no-arbitrage thresholds should theoretically be closer to zero for the second sample period than for the first sample period. The two intervals \( C_1 \in [-0.75, -0.05] \) and \( C_2 \in [0.05, 0.75] \)
used for the first sample period were therefore considered to be wide enough to search the lower and upper thresholds for the second sample period. Again, a grid search with 200 points on each of the two intervals was used to search for the threshold values. This method selects $\tilde{C}_1 = -0.554$ and $\tilde{C}_2 = 0.2635$ for the second sub-sample. As before, the magnitude of the estimated lower threshold ($\tilde{C}_1$) is larger than the magnitude of the estimated upper threshold ($\tilde{C}_2$), reflecting the higher transactions costs associated with arbitrage in the lower regime. Furthermore, both thresholds are smaller in the second sub-sample period than in the first, as conjectured, suggesting that arbitrage transactions costs were in general lower during the more recent period. This provides further evidence that the modifications to the trading systems of both the LSE and the LIFFE have improved its efficiency, and specifically here its transactions cost efficiency.

A Ljung-Box test was performed on the residuals from the selected TVECM to check model adequacy for both sample periods. No evidence of serial correlation in the residuals was found for either sub-sample (see Table 6.2). A TVECM was also estimated using lower lag orders, less than 4 lags for the first period and less than 5 lags for the second. However, the estimated residuals using these lower lag orders were found to be serially correlated, suggesting that to adequately handle FTSE100 spot and futures price dynamics, at least 4 lags must be included for the first period and 5 lags must be included for the second.
<table>
<thead>
<tr>
<th>lag</th>
<th>Spot residuals $e_{1t}$</th>
<th>Futures residuals $e_{2t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q-statistic</td>
<td>p-value</td>
</tr>
<tr>
<td>-----</td>
<td>--------------</td>
<td>----------</td>
</tr>
<tr>
<td><strong>28/10/1986 - 17/10/1997</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.00519</td>
<td>0.9426</td>
</tr>
<tr>
<td>2</td>
<td>0.00773</td>
<td>0.9961</td>
</tr>
<tr>
<td>3</td>
<td>0.01495</td>
<td>0.9995</td>
</tr>
<tr>
<td>4</td>
<td>0.03599</td>
<td>0.9998</td>
</tr>
<tr>
<td>5</td>
<td>0.622</td>
<td>0.9870</td>
</tr>
<tr>
<td><strong>27/10/1997 - 30/12/2005</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.01129</td>
<td>0.9154</td>
</tr>
<tr>
<td>2</td>
<td>0.01168</td>
<td>0.9942</td>
</tr>
<tr>
<td>3</td>
<td>0.01659</td>
<td>0.9994</td>
</tr>
<tr>
<td>4</td>
<td>0.02654</td>
<td>0.9999</td>
</tr>
<tr>
<td>5</td>
<td>0.08194</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

**Note:** The reported Ljung-Box $Q$-statistics and their $p$-values provide a joint test of serial correlation in the residuals up to the specified lag.

The estimation results of the threshold VECM for the first sub-sample are reported in Table 6.3. It can be seen that in the lower regime, when futures contracts are underpriced ($X_{t-1} \leq C_1$), the divergence from equilibrium is mainly corrected in the futures market, indicated by a negative and significant error correction term ($b_{t-k}$) in the futures equation. A negative effect of the lagged basis on the current futures price change indicates that if the basis was too large in a previous period, the current futures price would have fallen to drive it back to equilibrium (and *vice versa*). The lagged time to maturity ($T_{g-t+k}$), part of the error correction term, is weakly significant (at the 10% level) in the spot equation, possibly suggesting a weak error correction effect in the spot market. The arbitrage opportunity in this regime involves buying futures and short selling spot. The findings therefore imply that arbitrageurs can actually avoid
short sale restrictions in the spot market. In fact, the most active arbitrageurs in the market should be market makers or institutional traders: market makers because they can short sell stocks; institutional investors because they are often long in stocks and therefore can easily sell stocks directly from their long positions (Neal, 1996). As anticipated, in the inner regime when the mispricing of the futures contract is relatively small ($C_1 < X_{t-1} \leq C_2$), there is no error correction effect in either market.

Contrary to expectation, however, in the upper regime when futures contracts are overpriced ($X_{t-1} > C_2$), the error correction terms ($b_{t-k}$ and $T_{g-t+k}$) are insignificant in both spot and futures equations. The upper regime involves arbitrage opportunities that should easily be exploited by buying in the spot market and selling in the futures market. The historical development of the LSE may offer an explanation for this seemingly puzzling result. During the period 28/10/1986–17/10/1997, the LSE experienced a long-lasting bull market, especially in 1995-1997 during the dot-com bubble (see Figure 4.1 and Figure 4.2 in Chapter 4). Futures contracts may be overpriced in a bull market because of the possibility of increased risk associated with the arbitrage strategy of buying spot and selling futures. This is because the sell-order for index futures could be executed relatively quickly, but the buy-order in the spot market is typically slower. If the market is rising, the buy order may be executed at a higher price than the one at which the order was initially placed. In this case, the arbitrage strategy of buying spot and selling futures may turn out to be loss-making and stock market players may choose to forgo arbitrage opportunities because they are not risk-free. Another possible explanation is that in bull market conditions the cost of carry relationship based on a simple no-arbitrage argument may not completely reflect the price process of stock index futures. Considering jointly the factors (including arbitrage, hedging, speculation and other investment strategies) determining the supply and demand equilibration in the futures market and between spot and futures markets, arbitrage may be a weaker factor affecting the futures price process in bull market conditions: futures contracts could be ‘overpriced’ compared to the ‘fair’ value defined by the cost of carry because stock market players in general have optimistic attitudes about the current and future state of the economy and their personal financial situations, giving them confidence to make large investments. In such a case there could be more buyers than sellers in the futures market. Indeed, Hsu and Wang (2004) argue that the
arbitrage mechanism cannot be complete (particularly for index arbitrage) since capital markets are not perfect or frictionless, and that price expectation and risk aversion may play an important role in determining index futures prices in the real world.

Turning next to short-run price dynamics, it is found that in the lower and upper regimes, there is unidirectional positive causal effect from spot to futures, whereas in the inner regime there is a bidirectional positive causal effect between spot and futures. The evidence of stronger causality from spot to futures suggests that in most cases the quote-based index can reflect information more quickly than the transaction-based futures price, though lower transaction costs in the futures market compared to the spot market could have resulted in causality from futures to spot. Also, there is clear evidence of mean-reversion in futures price changes, which is less obvious for the spot price changes, implying either the existence of a bid-ask bounce effect in futures prices but not in spot prices, or possibly a stronger tendency of the futures price to overreact to certain new information, which is corrected subsequently. In all three regimes, there is evidence of significant $k_t$ or its lags in both spot and futures equations, suggesting that the passage of time itself also has some impact on price changes. There is also evidence of significant $z_t$ or its lags in both futures and spot equations, suggesting that the artificial price jumps at contract rollovers need to be explicitly accounted for in order to properly model spot-futures price dynamics. As anticipated, those $z_{t-j}$ that are significant in the futures equation are all positive while the joint effect of $z_{t-j}$ on the current futures price changes minus the joint effect of $z_{t-j}$ on the current spot price changes is also positive, reflecting a wider spread between futures and spot prices at contract rollovers.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>t-ratio</th>
<th>Coeff.</th>
<th>t-ratio</th>
<th>Coeff.</th>
<th>t-ratio</th>
<th>Coeff.</th>
<th>t-ratio</th>
<th>Coeff.</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>-0.00241</td>
<td>-0.39</td>
<td>-0.01168</td>
<td>-1.65&lt;sup&gt;e&lt;/sup&gt;</td>
<td>-0.00332</td>
<td>-1.74&lt;sup&gt;e&lt;/sup&gt;</td>
<td>-0.00456</td>
<td>-2.05&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.00161</td>
<td>-0.46</td>
</tr>
<tr>
<td>$\Delta_t s_t$</td>
<td>0.22824</td>
<td>1.88&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.58262</td>
<td>4.16&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.2285</td>
<td>-2.80&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.10983</td>
<td>1.15</td>
<td>-0.21162</td>
<td>-1.55</td>
</tr>
<tr>
<td>$\Delta_t f_t$</td>
<td>0.02356</td>
<td>0.19</td>
<td>0.18483</td>
<td>1.29</td>
<td>0.0949</td>
<td>1.11</td>
<td>0.29575</td>
<td>2.97&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.10250</td>
<td>0.73</td>
</tr>
<tr>
<td>$\Delta_t s_{t-1}$</td>
<td>-0.03869</td>
<td>-0.31</td>
<td>-0.04145</td>
<td>-0.29</td>
<td>0.1005</td>
<td>1.25</td>
<td>0.25830</td>
<td>2.75&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.15908</td>
<td>1.13</td>
</tr>
<tr>
<td>$\Delta_t s_{t-2}$</td>
<td>0.12515</td>
<td>1.05</td>
<td>0.08111</td>
<td>0.59</td>
<td>0.0271</td>
<td>0.39</td>
<td>0.11193</td>
<td>1.37</td>
<td>0.12252</td>
<td>1.06</td>
</tr>
<tr>
<td>$\Delta_t s_{t-3}$</td>
<td>-0.13576</td>
<td>-1.30</td>
<td>-0.50804</td>
<td>-4.23&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.2788</td>
<td>3.85&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.02889</td>
<td>-0.34</td>
<td>0.18443</td>
<td>1.56</td>
</tr>
<tr>
<td>$\Delta_t s_{t-4}$</td>
<td>-0.14313</td>
<td>-1.29</td>
<td>-0.25329</td>
<td>-1.98&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.0683</td>
<td>-0.91</td>
<td>-0.28630</td>
<td>-3.26&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.06918</td>
<td>-0.55</td>
</tr>
<tr>
<td>$\Delta_t f_{t-1}$</td>
<td>-0.04090</td>
<td>-0.38</td>
<td>-0.03588</td>
<td>-0.29</td>
<td>-0.0802</td>
<td>-1.13</td>
<td>-0.23395</td>
<td>-2.82&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.16013</td>
<td>-1.28</td>
</tr>
<tr>
<td>$\Delta_t f_{t-2}$</td>
<td>-0.08202</td>
<td>-0.79</td>
<td>-0.08321</td>
<td>-0.70</td>
<td>-0.0262</td>
<td>-0.43</td>
<td>-0.10662</td>
<td>-1.49</td>
<td>-0.09973</td>
<td>-0.95</td>
</tr>
<tr>
<td>$\Delta_t f_{t-3}$</td>
<td>0.071</td>
<td>0.92</td>
<td>-0.22073</td>
<td>-2.48&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.0143</td>
<td>0.30</td>
<td>-0.05698</td>
<td>-1.02</td>
<td>0.00125</td>
<td>0.02</td>
</tr>
<tr>
<td>$\Delta_t f_{t-4}$</td>
<td>0.00005</td>
<td>-1.78&lt;sup&gt;e&lt;/sup&gt;</td>
<td>-0.00002</td>
<td>-0.59</td>
<td>0.00001</td>
<td>1.27</td>
<td>0.00002</td>
<td>1.40</td>
<td>-0.00002</td>
<td>-1.08</td>
</tr>
<tr>
<td>$T_s - t + k_t$</td>
<td>1.17&lt;sup&gt;e&lt;/sup&gt;</td>
<td>-0.59</td>
<td>1.27</td>
<td>1.40</td>
<td>-0.00002</td>
<td>-1.08</td>
<td>-0.00003</td>
<td>-0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>( k_t )</td>
<td>-0.00107</td>
<td>-1.20</td>
<td>-0.00002</td>
<td>-0.02</td>
<td>0.00002</td>
<td>0.05</td>
<td>0.00021</td>
<td>0.59</td>
<td>0.00025</td>
<td>0.49</td>
</tr>
<tr>
<td>( k_{t-1} )</td>
<td>0.00118</td>
<td>1.21</td>
<td>0.00283</td>
<td>2.53^b</td>
<td>0.0006</td>
<td>1.82^c</td>
<td>0.00077</td>
<td>1.99^b</td>
<td>0.00047</td>
<td>0.82</td>
</tr>
<tr>
<td>( k_{t-2} )</td>
<td>0.00059</td>
<td>0.55</td>
<td>0.00209</td>
<td>1.70^c</td>
<td>0.00065</td>
<td>2.00^b</td>
<td>0.00097</td>
<td>2.54^b</td>
<td>-0.00012</td>
<td>-0.19</td>
</tr>
<tr>
<td>( k_{t-3} )</td>
<td>0.00143</td>
<td>1.36</td>
<td>0.00241</td>
<td>1.99^b</td>
<td>0.00065</td>
<td>2.01^b</td>
<td>0.00084</td>
<td>2.23^b</td>
<td>0.00048</td>
<td>0.82</td>
</tr>
<tr>
<td>( k_{t-4} )</td>
<td>0.00134</td>
<td>1.60</td>
<td>0.00219</td>
<td>2.27^b</td>
<td>0.00025</td>
<td>0.85</td>
<td>0.00025</td>
<td>0.72</td>
<td>0.00134</td>
<td>2.43^b</td>
</tr>
<tr>
<td>( z_t )</td>
<td>0.00002</td>
<td>0.03</td>
<td>0.00022</td>
<td>3.50^a</td>
<td>0.00003</td>
<td>0.15</td>
<td>0.00017</td>
<td>7.53^a</td>
<td>-0.00005</td>
<td>-1.74^c</td>
</tr>
<tr>
<td>( z_{t-1} )</td>
<td>0.00008</td>
<td>2.03^b</td>
<td>0.00009</td>
<td>1.87^c</td>
<td>-0.00002</td>
<td>-0.90</td>
<td>0.00001</td>
<td>0.46</td>
<td>-0.00002</td>
<td>-0.64</td>
</tr>
<tr>
<td>( z_{t-2} )</td>
<td>0.00001</td>
<td>0.27</td>
<td>0.00001</td>
<td>0.21</td>
<td>-0.00001</td>
<td>-0.40</td>
<td>0.00001</td>
<td>0.31</td>
<td>0.00004</td>
<td>1.01</td>
</tr>
<tr>
<td>( z_{t-3} )</td>
<td>0.00003</td>
<td>0.61</td>
<td>0.00004</td>
<td>0.85</td>
<td>0.00002</td>
<td>1.28</td>
<td>0.00004</td>
<td>1.60</td>
<td>0.00002</td>
<td>0.50</td>
</tr>
<tr>
<td>( z_{t-4} )</td>
<td>0.00003</td>
<td>0.83</td>
<td>0.00004</td>
<td>0.91</td>
<td>-0.00004</td>
<td>-1.93^c</td>
<td>-0.00003</td>
<td>-1.35</td>
<td>0.00007</td>
<td>2.06^b</td>
</tr>
</tbody>
</table>

**Notes:**

* significant at the 1% level

^b* significant at the 5% level

^c* significant at the 10% level.
The estimation results of the threshold VECM for the second sub-sample are reported in Table 6.4. The empirical findings for this period provide an interesting contrast to those of the first period. In the inner regime, when the mispricing of the futures contract is relatively small \((C_1 < X_{t-1} \leq C_2)\), there is no significant error correction effect in either market. In the upper regime when futures contracts are overpriced \((X_{t-1} > C_2)\), the error correction term \((b_{t-k})\) is negative and significant in both spot and futures equations. However, it is more negative in the futures equation than that in the spot equation, suggesting that if the lagged basis was too large/small both futures and spot prices may have fallen/risen, but that the futures price may have fallen/risen by more than the spot price to recover the equilibrium relationship between the two markets. The error correction term \((b_{t-k})\), however, is found to be insignificant in both spot and futures equations in the lower regime when futures contracts are underpriced \((X_{t-1} \leq C_1)\), though the lagged time to maturity, \(T_{g-t+k_i}\), is significant in both equations. Since \(T_{g-t+k_i}\) is part of the error correction term, its significance may suggest some error correction effect in both markets, but this is not very convincing when \(b_{t-k}\) itself is insignificant. The arbitrage opportunity in this regime involves buying futures and short selling spot, which should theoretically be less difficult for arbitrageurs if they have already been able to make it in the earlier period. 

Again, an explanation for this puzzle may come from the historical development of the LSE. It can be seen from Figure 4.1 and Figure 4.2 in Chapter 4 that spot and futures prices were mainly declining over 1998-2003 and that both experienced several sharp falls. In particular, the market fell continuously over the three years from the beginning of 2000. The arbitrage strategy of buying futures and short selling spot would be difficult to execute in such markets because it would be more difficult to find a counterparty either for short sales or outright sales. Therefore, in a bear market it is likely that arbitrageurs may have to pass up profitable arbitrage opportunities created by underpriced futures contracts. It could also be that in bear market conditions, the cost of carry relationship based on a simple no-arbitrage argument cannot completely reflect the price process of stock index futures. In a bear market, arbitrage may become a very weak factor affecting the futures price process compared to the joint factors determining the supply and demand equilibration in the futures market.
contracts would tend to be ‘underpriced’ compared to the ‘fair’ value defined by the cost of carry when stock market players have pessimistic attitudes about the current and future state of the economy and their personal financial situations, and when they are reluctant to make investments, and hence there should be more sellers than buyers in the futures market. Again, this empirical evidence provides support to the proposal of Hsu and Wang (2004), that price expectation and risk aversion may play an important role in determining index futures prices in real futures markets.

Regarding short-run price dynamics, there is strong evidence of a positive causal effect from spot to futures in both the lower and the inner regimes. However, there is also weak evidence of a negative causal effect from futures to spot in the lower and the inner regimes. In the upper regime, there is evidence of a positive causal effect from futures to spot. The evidence of a stronger positive causal effect from spot to futures further indicates that in most cases the quote-based index can reflect information more quickly than the transaction-based futures price, although lower transaction costs in the futures market compared to the spot market may have also resulted in causality from futures to spot. Again, there is clear evidence of mean-reversion in the futures price changes, but only in the lower and the inner regimes, again implying either the existence of a bid-ask bounce effect in futures prices or overreaction of the futures price to certain new information.

The evidence of significant $k_t$, $z_t$ or their lags is not as strong in the second period as in the first, suggesting that compared to other factors, the passage of time and artificial futures price ‘jumps’ at contract rollovers are not so important in explaining spot-futures price dynamics during the second sub-sample. Indeed, some of the $k_{t-j}$ are negative and significant in the upper regime, although only at the 10% level. In the lower regime, $z_{t-3}$ and $z_{t-4}$ are positive and significant, but only at the 10% level, in both spot and futures equations. The joint effect of $z_{t-3}$ and $z_{t-4}$ on the current futures price change minus the joint effect of $z_{t-3}$ and $z_{t-4}$ on the current spot price change is also positive, again reflecting a wider spread between futures and spot prices at contract rollovers.
Table 6.4 Least squares estimates and their t-ratios of the threshold VECM, 27/10/1997 – 30/12/2005

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regime 1 (157 observations)</th>
<th>Regime 2 (1727 observations)</th>
<th>Regime 3 (162 observations)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \Delta_k s_t )</td>
<td>( \Delta_k f_t )</td>
<td>( \Delta_k s_t )</td>
</tr>
<tr>
<td>intercept</td>
<td>Coeff.</td>
<td>( t )-ratio</td>
<td>Coeff.</td>
</tr>
<tr>
<td></td>
<td>0.00886</td>
<td>0.80</td>
<td>0.0103</td>
</tr>
<tr>
<td>( \Delta_k s_{t-1} )</td>
<td>-0.41681</td>
<td>-0.77</td>
<td>0.59106</td>
</tr>
<tr>
<td>( \Delta_k s_{t-2} )</td>
<td>0.88175</td>
<td>1.17</td>
<td>1.52532</td>
</tr>
<tr>
<td>( \Delta_k s_{t-3} )</td>
<td>1.10186</td>
<td>1.41</td>
<td>1.61558</td>
</tr>
<tr>
<td>( \Delta_k s_{t-4} )</td>
<td>1.28098</td>
<td>1.71&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.43773</td>
</tr>
<tr>
<td>( \Delta_k s_{t-5} )</td>
<td>0.33668</td>
<td>0.53</td>
<td>0.3206</td>
</tr>
<tr>
<td>( \Delta_k f_{t-1} )</td>
<td>0.33854</td>
<td>0.65</td>
<td>-0.65911</td>
</tr>
<tr>
<td>( \Delta_k f_{t-2} )</td>
<td>-0.84025</td>
<td>-1.14</td>
<td>-1.46035</td>
</tr>
<tr>
<td>( \Delta_k f_{t-3} )</td>
<td>-1.40597</td>
<td>-1.82&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-1.90254</td>
</tr>
<tr>
<td>( \Delta_k f_{t-4} )</td>
<td>-1.24099</td>
<td>-1.69&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-1.41872</td>
</tr>
<tr>
<td>( \Delta_k f_{t-5} )</td>
<td>-0.5147</td>
<td>-0.82</td>
<td>-0.51638</td>
</tr>
<tr>
<td>( b_{t-k} )</td>
<td>-0.06784</td>
<td>-0.15</td>
<td>-0.04849</td>
</tr>
<tr>
<td>( T_g-t+k_t )</td>
<td>-0.0002</td>
<td>-2.31&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.00024</td>
</tr>
<tr>
<td>Variable</td>
<td>Regime 1 (157 observations)</td>
<td>Regime 2 (1727 observations)</td>
<td>Regime 3 (162 observations)</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td></td>
<td>(\Delta_k s_t)</td>
<td>(\Delta_k f_t)</td>
<td>(\Delta_k s_t)</td>
</tr>
<tr>
<td>(k_t)</td>
<td>0.00127 0.52</td>
<td>0.00105 0.43</td>
<td>0.00083 1.80&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>(k_{t-1})</td>
<td>-0.0015 -0.72</td>
<td>-0.00167 -0.80</td>
<td>0.0006 1.40</td>
</tr>
<tr>
<td>(k_{t-2})</td>
<td>-0.0016 -0.75</td>
<td>-0.00166 -0.77</td>
<td>0.00015 0.33</td>
</tr>
<tr>
<td>(k_{t-3})</td>
<td>0.00183 0.97</td>
<td>0.00192 1.02</td>
<td>0.00104 2.27&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>(k_{t-4})</td>
<td>0.0003 0.16</td>
<td>0.00058 0.30</td>
<td>0.00068 1.56</td>
</tr>
<tr>
<td>(k_{t-5})</td>
<td>-0.00113 -0.39</td>
<td>-0.00071 -0.25</td>
<td>0.00004 0.08</td>
</tr>
<tr>
<td>(z_t)</td>
<td>0.00008 0.64</td>
<td>0.00012 0.96</td>
<td>-0.00003 -0.94</td>
</tr>
<tr>
<td>(z_{t-1})</td>
<td>0.00019 1.01</td>
<td>0.00016 0.88</td>
<td>-0.0000003 -0.01</td>
</tr>
<tr>
<td>(z_{t-2})</td>
<td>0.00003 0.20</td>
<td>0.00005 0.34</td>
<td>-0.000004 -1.55</td>
</tr>
<tr>
<td>(z_{t-3})</td>
<td>0.00022 1.72&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.00022 1.79&lt;sup&gt;e&lt;/sup&gt;</td>
<td>-0.00002 -0.81</td>
</tr>
<tr>
<td>(z_{t-4})</td>
<td>0.00020 1.67&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.00032 1.84&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.00003 1.27</td>
</tr>
<tr>
<td>(z_{t-5})</td>
<td>-0.00001 -0.09</td>
<td>-0.00002 -0.13</td>
<td>0.00002 0.57</td>
</tr>
</tbody>
</table>

**Notes:**
- <sup>a</sup>significant at the 1% level
- <sup>b</sup>significant at the 5% level
- <sup>c</sup>significant at the 10% level.
In this chapter, the price relationship between the FTSE100 spot and futures markets is re-examined using a three-regime threshold VECM. At least two contributions are made to the literature. First, a threshold cointegration analysis is used to provide insights into the regime-dependent (rather than constant) price dynamics between FTSE100 spot and futures markets, providing an alternative way to test the cost of carry model. A basic finding in this study is that the cost of carry model may not be suitable for pricing stock index futures in certain 'abnormal' market conditions. The second contribution is about the selection of the threshold variable in the threshold cointegration analysis. In this study, the threshold variable for the threshold VECM is carefully selected for two concerns. First, we believe that an accurate estimate of the threshold values relies on specifying a reasonable range within which to search the threshold values. The percentage mispricing of stock index futures proposed by Mackinlay and Ramaswamy (1988) can be directly compared to actual transaction costs because both are expressed as a proportion of the underlying index value. Therefore, substantive information about transaction costs associated with arbitrage allows us to specify less arbitrarily a more reasonable range for estimating the threshold values. We believe that our approach is more reliable than the search procedure employed by Tsay (1998). Second, given that the percentage mispricing is directly comparable to actual transaction costs, the estimated threshold values would provide immediate information about average transaction costs faced by most active arbitrageurs in the market, which is to the great interest of both researchers and practitioners.

In both sample periods, the estimated lower threshold ($\hat{C}_1$) is larger than the estimated upper threshold ($\hat{C}_2$), implying that the transaction cost associated with arbitrage in the lower regime is higher than that associated with arbitrage in the upper regime. This supports the argument that short sale restrictions in the spot market may represent a hurdle for arbitrage. Since arbitrage opportunities in the lower regime are created by underpriced futures contracts and the arbitrage strategy involves buying futures and short selling spot, the empirical findings imply that while some arbitrageurs could have avoided short sale restrictions in the spot market, the average transaction cost is higher for arbitrage involving short sales in the spot market. Both lower and upper thresholds are smaller during the second period than during the first, suggesting that transaction
costs for arbitrage have become in general lower in more recent times. This provides support to the contention that the order-driven trading system (SETS) launched by the LSE in 1997 and the electronic platform (LIFFE CONNECT) started by LIFFE in 1998 have successfully reduced transaction costs for market participants.

As suggested by the theory of threshold cointegration, both the short run price dynamics and the long run error correction effect are found to be different in different regimes. There are also some puzzling empirical findings: in the first period the error correction term is insignificant in the upper regime (when futures contracts are overpriced) while in the second period the error correction term is insignificant in the lower regime (when futures contracts are underpriced). These seemingly puzzling results have potentially important implications. One is that under some market conditions, arbitrage may not be practicable, either because it is difficult to find the counterparties for the arbitrage transaction, or because there is significant risk associated with the arbitrage. It is also possible that the observed mispricing of futures means that the cost of carry model is not always suitable for pricing stock index futures (Hemler and Longstaff, 1991). First, no-arbitrage models that effectively assume the stock market to be exogenous might fail to capture dynamic interactions between spot and futures markets. This is important, given the overwhelming evidence of interdependence between stock index spot and futures prices documented in the literature as well as in this study. Second, the cost of carry model is intended to price forward contracts, and futures and forward prices may not be equal if interest rates are stochastic. Third, Hemler and Longstaff (1991) find that stock index futures prices are related to both market volatility and stochastic interest rates, which are not accounted for by the cost of carry model. Fourth, arbitrageurs are only one type of traders in futures markets. In some special market conditions, arbitrage may become a very weak factor determining the supply and demand equilibration in the futures market and between spot and futures markets. Therefore, the evidence of so many ‘overlooked’ arbitrage opportunities may simply be a statistical illusion created by the wrong estimation of ‘mispricing’ – the ‘fair’ price implied by the cost of carry model may be not fair!

A limitation of this study is that all arbitrageurs are assumed to have the same transaction costs and therefore only three regimes are considered. In reality,
arbitrageurs are heterogeneous and arbitrage transaction costs are different for different market participants. As a result, there may not be a single valid no-arbitrage band that applies to all individuals. Yadav and Pope (1994) suggest two levels of transaction costs for two broad categories of potential arbitrageurs. Therefore, a possible extension to this study is to estimate a five-regime TVECM, with two pairs of thresholds corresponding to the two levels of transaction costs faced by two broad categories of arbitrageurs. One drawback to proceeding in this way is the computational complexity of estimating a five-regime TVECM, as well as the argument that arbitrageurs with higher transaction costs should be driven out by those with lower transactions costs.
Chapter 7 Concluding Remarks

The major aim of this thesis was to investigate daily price dynamics between the FTSE100 spot and futures markets and to examine the validity of the cost of carry model for pricing FTSE100 futures contracts. This aim was achieved by using a VECM as the basic empirical framework (Chapter 4). The basic VECM was then extended with a DCC-TGARCH to model conditional variance-covariance processes (Chapter 5) and extended to a three-regime threshold VECM to model arbitrage-induced regime-dependent price dynamics (Chapter 6). The empirical studies were performed in the following framework:

1. The single time series of FTSE100 futures price was created artificially by splicing together the price series of individual futures contracts.
2. The daily settlement price for futures contracts was assumed to be synchronous with daily closing price for the index.
3. The LDMON (London Discount Market Overnight Rate) was used as a proxy for the risk free interest rate for a loan matching the changeable maturity of those futures contracts.
4. The realized ex post dividend yield was used as a proxy for the forecasted dividend yield for FTSE100 index.
5. The price dynamics between FTSE100 spot and futures markets were assumed to be stable within two sample periods: 28/10/1986 – 17/10/1997 and 27/10/1997 – 30/12/2005.
6. For the basic linear VECM analysis and the DCC-TGARCH analysis, the adjustment toward equilibrium was assumed to be both continuous and of constant speed, regardless of the size of the deviation from equilibrium. This assumption was relaxed in the threshold VECM analysis. However, in the threshold VECM analysis, all arbitrageurs were assumed to be homogeneous and have identical transaction costs, identical margin requirements and position limits.
Regarding the VECM structure, in addition to standard lagged own and cross-market returns and the lagged error correction term, three ‘extra’ variables were included to account for the effect of the passage of time, artificial price ‘jumps’ at contract rollovers and the trend effect in the basis for each individual futures contract. In particular, a dummy variable first proposed by Green and Joujon (2000) was carefully redefined in this thesis to account for artificial ‘jumps’ in futures prices without changing the real transaction prices. These ‘extra’ variables are in most cases found to have significant, though quantitatively small, explanatory power for futures and/or spot price movements, suggesting that it is necessary to include them in order to accurately model the dynamic processes for a pair of spot and futures prices.

The selection of the sample period and two sub-samples in this research was based on the principle that since a linkage between spot and futures markets should be maintained by arbitrage, which is affected to a great extent by the costs of trading in the two markets, stable price dynamics between FTSE100 spot and futures markets and a stable no-arbitrage window should only exist over a time period with a relatively stable transactions cost structure. The LSE claimed that their new order-driven trading system (SETS) on 20/10/1997 should ‘bring greater speed and efficiency to the market’. LIFFE moved to an all electronic trading platform (LIFFE CONNECT) on 30/11/1998 in order to ‘have a lower cost base’. The empirical relationship between FTSE100 spot and futures prices was therefore analysed in two sub-samples: the first sub-sample covers 28/10/1986 – 17/10/1997, during which the LSE employed a dealership system (SEAQ) and LIFFE used ‘open outcry’; the second sub-sample covers 27/10/1997 – 30/12/2005, when the LSE moved to an order-driven system (SETS) and LIFFE moved to an electronic platform.

The empirical results indicate that the principle for selecting the overall sample period and the two sub-samples in this research is plausible. The findings in Chapter 4 and Chapter 5 suggest that during the first sub-sample, the cost of carry relationship is violated for FTSE100 spot and futures prices and the divergence from equilibrium is corrected in the futures market; during the second sub-sample, the cost of carry relationship tends to be valid and there is no obvious evidence of divergence from equilibrium, suggesting that any divergence has been corrected by short-run price

27 LIFFE moved to the electronic platform on 30/11/1998. The second sub-sample was started from 27/10/1997 to avoid estimation problems (see Chapter 1 and Chapter 4 for details).
adjustments rather than by the actions of arbitrageurs. In Chapter 6, the threshold values yielded by estimating the threshold VECM suggest that the average transactions costs associated with arbitrage are lower in the second sub-sample than in the first. These findings are encouraging – the modifications to the trading systems of the LSE and LIFFE have improved their transactions cost efficiency, as reflected in a systematic decline in the profitability of arbitrage between FTSE100 spot and futures markets. They also have important regulatory implications for both the UK and elsewhere, because lower transactions costs and technological advances can contribute to more effective and efficient functioning of financial markets, presumably through increased competitiveness in arbitrage. Perhaps the UK government should also consider lowering or eliminating the stamp duty (0.5% since 1986) on purchases of shares of British companies (London is the only one of the three major financial centers with such a tax – New York and Tokyo go without).

Some important findings of this research are summarized below.

First, regarding short-run dynamics between FTSE100 spot and futures markets, a general result is that there is bi-directional causality-in-mean between spot and futures prices, with stronger evidence of causality from spot to futures than the reverse. This seems to contradict the general results reported in the literature (where most evidence is documented for the US). The results found here may be specific to the UK market because UK indices are quote-based, while indices in other markets are usually transaction-based. Transaction-based indices are computed using the last available transaction price for each of the constituent stocks. Quote-based indices should in principle reflect more up-to-date information than transaction-based indices because market makers (or market participants) are free to adjust their quotes for the constituent shares in immediate response to new information. Therefore, UK indices are less susceptible to the effects of stale transaction prices associated with the transaction-based indices used in other markets. Since the FTSE100 futures price can reflect new information only if transactions on the futures contract have taken place, it is not surprising to find that the FTSE100 cash index usually leads the futures price in response to new information. Nonetheless, since trading in the futures market is cheaper than trading in the spot market, FTSE100 futures price changes sometimes do lead spot price changes. These findings have important implications. While futures markets are
usually believed to serve as a 'price discovery' vehicle for the underlying spot markets, this could be a statistical illusion resulting from the stale price effects contained in transaction-based indices (Yadav and Pope, 1994). Although quote-based indices should reflect more up-to-date information than transaction-based indices, the quotes are not necessarily real transaction prices, so seemingly apparent arbitrage opportunities between FTSE100 index and futures could also sometimes be illusory.

Another general result is that both FTSE100 spot and futures returns tend to be mean-reverting, although the futures return displays much stronger evidence of mean-reversion than the spot return. As analysed before, mean reversion may imply that both prices sometimes overreact to new information, so that subsequent corrections in the prices are driven by the true value contained in the information. In this regard, mean-reversion is necessary for the market to be efficient in the long run. Compared to the index itself, the futures price seems to overreact more, possibly due to the transaction advantages of trading in futures markets, where there are lower transaction costs, higher liquidity and leverage opportunities (futures positions require only a fraction of the contract value for initial margin). Furthermore, futures transactions can be effected more quickly than spot transactions because the purchase of one futures contract represents a claim on the whole index portfolio. Another common explanation for mean-reversion in futures returns is the bid/ask bounce effect, which would not affect the spot index. Overall, therefore, futures returns tend to show stronger evidence of mean-reversion than spot returns.

Where there is evidence of deviation from equilibrium (especially in the first subsample), the deviation is mainly corrected in the futures market, further indicating either a stronger lead of the spot price over the futures price or a stronger tendency of the futures price to overreact to new information. The stronger lead of the spot price over the futures price suggests a generally lagged response of the futures price to new information, which may cause divergence from equilibrium between the two markets. Subsequent adjustment in the futures market is then necessary to drive both the futures price and the diverged relationship between the two markets to a new equilibrium. On the other hand, due to ease of trading in futures markets, the futures price may sometimes overreact to new information, causing divergence from equilibrium of both the futures price and the relationship between the two markets. Subsequent adjustment
in the futures market would therefore be necessary to correct the overreaction of the futures price and the divergence from equilibrium.

Extending the basic linear VECM to DCC-TGARCH to model conditional variance-covariance processes provides a more complete insight into both the first and the second moment dynamics of FTSE100 spot and futures prices. In general, the DCC-TGARCH successfully captures the main dynamics in the second moment. Both the conditional variances of the two price series and the conditional correlation between them are found to be highly persistent and predictable. This finding has important implications for those (fund managers) using stock index futures in their daily risk management. Although high correlation between stock index futures and the underlying spot market makes index futures the most widely used instrument for hedging against market risks, both the conditional variance and the conditional correlation between futures and spot prices vary significantly over time. This implies that to acquire efficient hedging, the hedge ratio needs to be duly updated according to market movements. The DCC-TGARCH model was chosen in this research because it is an easy-to-implement multivariate GARCH with parsimonious parametric models for both the conditional variance and the conditional correlation. However, since the DCC-TGARCH model does not allow for inter-market dependence between volatility series, it may ignore some aspects of the volatility dynamics in each individual series. Two-step estimation of DCC-TGARCH can be combined with the two-step CCF test for causality-in-variance, though this may suffer from a loss of efficiency because the optimal parameters estimated in the first step are not necessarily optimal in the second step of the estimation procedure. The DCC-TGARCH estimation results therefore need to be interpreted with caution, especially with respect to the inference of no obvious evidence of causality-in-variance between FTSE100 spot and futures (only weak evidence of causality-in-variance from spot to futures during the first period) compared to obvious evidence of bidirectional causality-in-mean between them.

Threshold cointegration analysis offers a more complex picture of regime-dependent price dynamics between FTSE100 spot and futures markets that is consistent with the argument that arbitrageurs only enter into the market when the divergence from equilibrium is sufficiently large to offset transaction costs and potential arbitrage risks. As a result, price dynamics between spot and futures markets might be threshold
nonlinear, depending on the presence or absence of arbitrage. The results show that transactions costs indeed affect the stochastic process governing asset prices because they create a no-arbitrage band within which the futures price is free to fluctuate without triggering profitable arbitrage. On the other hand, even in the presence of transaction costs and risks, arbitrage opportunities arise outside the no-arbitrage window, generating much trading activity that exploits mispricing and helps to drive asset prices toward their equilibrium levels.

As anticipated, with the carefully selected threshold variable (percentage mispricing), the estimated threshold values provide immediate information about average arbitrage transaction costs that is consistent with economic and financial intuition. In both sub-samples the arbitrage transaction cost in the lower regime is higher than in the upper regime, consistent with the argument that short sale restrictions in the spot market might represent a hurdle for arbitrage. Both lower and upper thresholds are smaller during the second sub-sample than during the first, consistent with the expectation that transactions costs associated with arbitrage should have been reduced by technological advances in more recent times.

An important inference from the threshold cointegration analysis is that the cost of carry model may not always be suitable for pricing stock index futures, especially in certain 'abnormal' market conditions. The reason is that factors (such as volatility, stochastic interest rate and price expectation) not included in the cost of carry model could also be the significant drivers of price change in the markets. In circumstances unfavorable to arbitrage the no-arbitrage relationship on which the cost of carry model is based could break down, so that the cost of carry model is doomed to fail. Alternatives to the cost of carry model that account for real world frictions and other relevant factors might reflect more accurately the price process of stock index futures.

Overall, we would claim that this thesis adds to our knowledge of stock index futures markets and their pricing relationship to the underlying indices, especially of FTSE100 index and London market. These contributions to knowledge are mainly due to the usage of unique dataset and analysis of the same basic issue under different yet comparable model frameworks:
First, since the FTSE100 index is a quote-based index, it provides a unique test for lead-lag relationships between a stock index and its futures price that is less subject to infrequent trading problems (e.g., transaction-based US indices widely analysed in the literature are potentially subject to infrequent trading problems). Interestingly, based on comparable studies under different model frameworks, we find invariably that there is a stronger tendency for FTSE100 spot price to lead its futures price than the reverse, which imposes a challenge to the normal belief that futures price should lead the underlying spot price. To the existent knowledge of factors that could affect lead-lag relationships between index cash and futures prices, therefore, we may add another consideration: a quote-based index that is less subject to infrequent trading problems should be able to reflect information more quickly than a transaction-based index and the transaction-based futures price.

Second, because of transformations of trading systems at both LSE and LIFFE at about the same time (with one year gap: LSE transformed from a dealership system (SEAQ) to an order-driven trading system (SETS) on 20/10/1997; LIFFE replaced the old 'open outcry' system with a new electronic platform on 30/11/1998), the 20 years daily data covering these transformations used in this research provides a natural experiment of the impact of transaction costs and other market frictions on arbitrage and hence on spot-futures price dynamics. Empirical evidence of different behaviours of FTSE100 spot-futures price dynamics under different trading systems, in turn, provides a signal of increased market efficiency due to improvement of trading environment in both markets: the cost of carry relationship between FTSE100 spot and futures markets is rejected in the first period, but tends to be accepted in the second period. Our findings have important implications for future research in efficient market hypothesis (EMH). It is well known that tests of EMH are always joint tests of market efficiency and the underlying asset pricing model: if the EMH is rejected, one is not sure whether it indicates inefficiency of the market or invalidity of the underlying asset pricing model. In this regard, a direct comparison of the validity of the same asset pricing model in different time periods or in different market states for the same market is helpful. In the best case, provided that the model is fairly good, it would provide useful information about whether the efficiency of the market has been improved over time, a useful signal to both regulators and market participants. In the worst case that the model is not good, comparing different degrees of rejecting the model under different market situations
would allow us to analyse possible reasons for the failure of the model and to improve the model. In our case, for example, after analyzing the market, we conclude that the invalidity of the cost of carry model under ‘abnormal’ market is due to failure to take into account factors such as volatility, stochastic interest rate and price expectation, which should play more important roles in defining futures prices under ‘abnormal’ market conditions than under stable market conditions.

Further possible research directions implied from the findings of this thesis are summarized as follows. First, a possible limitation of this thesis is that the empirical studies are based on daily data for FTSE100 spot and futures prices. Since arbitrage in the real world should normally occur immediately in the presence of any large enough mispricing, empirical studies based on intraday data would be able to provide a better and more accurate insight into arbitrage-induced price dynamics and the relevant error correction effect. Second, this thesis suggests that the FTSE100 cash price tends to lead the futures price. The no-arbitrage cost of carry relationship seems to hold for FTSE100 spot and futures markets over the second sample period, 27/10/1997 – 30/12/2005. Since the FTSE100 index is composed of the 100 most highly capitalized and also most actively traded shares listed on the LSE, it may be that the findings of this research are specific to FTSE100 index and futures and cannot be generalized to the whole UK market. Shares listed on the LSE but not included in the FTSE100 index are normally less frequently traded so that the price quotes for such shares should accordingly be less frequently updated. It would therefore be interesting to examine price dynamics between FTSE250 index and futures prices within the framework used here. This would provide a complementary insight into the performance of UK stock index futures.

With regard to methodology, although the basic VECM and its extensions highlight important features of the dynamic relationship between FTSE100 spot and futures markets, there is a lot of room for improvement. In particular, to overcome the loss of efficiency associated with the two-step estimation of the DCC-TGARCH model, other versions of multivariate GARCH models can be used to capture conditional variance and covariance processes of spot and futures prices. In a market with heterogeneous agents who face different levels of transactions costs (or different margin requirements or position limits), agents essentially face no-arbitrage bands of different size. That is, there may be no valid single no-arbitrage band applying to all individuals. Thus a
possible extension to the three-regime threshold VECM is to estimate a multiple-regime threshold VECM, with multiple pairs of thresholds corresponding to different levels of transaction costs faced by different categories of arbitrageurs (Martens et al., 1998). Further insights into the nonlinear adjustment process could be gained by developing a threshold autoregressive (TAR) model for the basis, possibly also involving other variables (such as $k_i$, $z$, and $T-t+k$, introduced in this research) capable of affecting both the equilibrium level of the basis and its dynamic adjustment towards equilibrium. In response to the widely held view that 'arbitrage is like gravity', that is, relatively larger deviations of asset prices from their fair values (larger mispricing) induce relatively faster adjustment of asset prices toward their equilibrium values, it may be possible to develop a smooth transition autoregressive (STAR) model for the basis. This would allow smooth rather than discrete nonlinear adjustment toward equilibrium, with the speed of adjustment allowed to vary with the degree of divergence from equilibrium (Monoyios and Sarno, 2002). These proposals remain on the agenda for future research.
Bibliography


Antoniou, A. & Garrett, I. 1989, "Long-Run Equilibrium and Short-Run Dynamics: the Case of the FTSE 100 Share Index and Futures Contract", Discussion Papers in Economics No. 8906, Brunel, the University of West London, Uxbridge Middlesex UB8 3PH, pp. 1-32.


