Modelling the Behaviour of Arbitragers and Speculators in the Crude Oil Futures Market

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Obaid Anwar Awan

B. Bus&IT (Hons), M. Bus&IT, M. Fin, GDip Acct

School of Economics, Finance, and Marketing

College of Business

RMIT University

December 2015
DECLARATION

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Obaid Anwar awan

15th December 2015
ACKNOWLEDGEMENTS

This dissertation has indeed been an emotional yet a memorable and rewarding life experience. A number of people played an important role in this journey and I would like to state my appreciation to them.

First and foremost, I am grateful to my senior supervisor Professor Imad Moosa for his guidance and insightful comments on preliminary drafts which helped immensely in refining this thesis. I am forever indebted to him for always finding time to assist me notwithstanding his busy work schedule. I would also like to thank my associate supervisor Dr. George Tawadros for his tireless support throughout my PhD journey. In addition, I am grateful to Elite Editing for providing the editorial work.

A special thanks goes to Professor Terry Hallahan, and late Professor Tony Naughton for providing me the opportunity to undertake my PhD. Also, I am grateful to Greg Maynes, Dr Michael Gangemi, Dr Monica Tan and Dr Marie-Anne Cam for their advice and help in managing my studies and teaching commitments. Besides, I would like to acknowledge my fellow PhD students, especially Guillaume Galanos, Peter Le and Bin Liu for their encouragement and motivation that kept me going through difficult times.

Finally, I would like to thank my father, Muhammad Anwar Awan and my mother, Imtiaz Anwar for their unconditional support, affection and belief in my abilities. Words alone cannot describe my gratitude for all they have done for me. I dedicate this dissertation to my parents.
# TABLE OF CONTENTS

Declaration.............................................................................................................................................. ii  
Acknowledgements................................................................................................................................. iii  
Table of Contents .................................................................................................................................... iv  
List of Tables ........................................................................................................................................ viii  
List of Figures .......................................................................................................................................... x  
Abstract................................................................................................................................................ xii  

Chapter 1 — Introduction ...................................................................................................................... 1  
  1.1 Background .................................................................................................................................. 1  
  1.2 What is Financialisation? ............................................................................................................. 2  
  1.3 How Financialisation Impacts the Futures Market ..................................................................... 3  
  1.4 Motivation for the Study ............................................................................................................. 8  
  1.5 Structure of the Thesis ................................................................................................................. 9  

Chapter 2 — Impact of Financial Trading on Crude Oil Prices: Theory and Evidence ............. 13  
  2.1 Introduction ................................................................................................................................ 13  
  2.2 Evidence of the Impact of Financialisation ................................................................................. 14  
    2.2.1 Crude Oil Prices and Inventory Dynamics ........................................................................ 15  
    2.2.2 The Role of Open Interest ................................................................................................. 19  
    2.2.3 The Role of Index Investments in Driving Crude Oil Price ............................................. 23  
  2.3 A Behavioural Finance Approach to Modelling Crude Oil Prices ......................................... 26  
  2.4 Summary and Conclusion .......................................................................................................... 33  

Chapter 3 — A Theoretical Framework .............................................................................................. 36  
  3.1 Introduction ................................................................................................................................ 36  
  3.2 The Role of Arbitrages ................................................................................................................. 37  
  3.3 The Role of Speculators .............................................................................................................. 41  
  3.4 A Diagrammatic Exposition ....................................................................................................... 43  
  3.5 Model Specification ..................................................................................................................... 47  
  3.6 A Testable Model for Futures Prices ......................................................................................... 49  
  3.7 Summary .................................................................................................................................... 49
6.4.1 Oil Demand
6.4.2 Oil Supply
6.5 Summary and Conclusion

Chapter 7 — Empirical Results on the Role of Expectations
7.1 Introduction
7.2 Data
7.3 Modelling the Equilibrium Price of Oil
7.4 Modelling Expectations
7.5 Equilibrium Price Estimation Results
7.6 Oil Price Forecast Evaluation
7.7 Cointegration Results
7.7.1 Extrapolative Expectations
7.7.2 Adaptive Expectations
7.7.3 The Random Walk
7.7.4 Regressive Expectations
7.8 Summary and Conclusion

Chapter 8 — A Heterogeneous Agent model of Speculation
8.1 Introduction
8.2 Heterogeneous Agents Models
8.3 The Model
8.3.1 Fundamentalists
8.3.2 Chartists
8.3.3 Learning Mechanism
8.4 Methodology
8.5 Data Description and Empirical Results
8.5.1 Crude Oil Price Forecast Evaluation
8.5.2 The Heterogeneous Agent Model Estimation Results
8.5.3 Cointegration Testing
8.6 Summary and Conclusion

Chapter 9 — Conclusion
9.1 Recapitulation
9.2 Contribution
LIST OF TABLES

Table 5.1: Descriptive Statistics at Levels from January 1988 to February 1999 ............... 75
Table 5.2: Descriptive Statistics at Levels from March 1999 to December 2003 ............... 77
Table 5.3: Descriptive Statistics at Levels from January 2004 to April 2013 ..................... 77
Table 5.4: Unit Root Tests at Log Levels ............................................................................ 88
Note: All variables are expressed natural logs ................................................................... 88
Table 5.5: Unit Root Tests at First Difference ................................................................... 88
Table 5.6: Cointegration Regression—Rational Expectations ........................................... 90
Table 5.7: Expected Profit from Speculation ..................................................................... 91
Table 5.8: Cointegration Regression with Zero Convenience Yield .................................. 92
Table 5.9: Test of Coefficient Equality .............................................................................. 93
Table 5.10: Impact of Estimated Convenience Yield on West Texas Intermediate Futures
Price ........................................................................................................................ 94
Table 5.11: Cointegration Regression, January 1986 to December 1991 ............................ 95
Table 6.1: Statistical Significance of the Drift Term ......................................................... 113
Table 7.1: Data Description ............................................................................................. 120
Table 7.2: Correlation Matrix for Oil Price Fundamentals in Log Levels ....................... 121
Table 7.3: Unit Root Test at Levels .................................................................................. 133
Table 7.4: Unit Root Test at First Difference ................................................................... 133
Table 7.5: Cointegration Regression for Equilibrium Oil Price ....................................... 135
Table 7.6: Root Mean Squared Error of Extrapolative Expectations ............................... 141
Table 7.7: Root Mean Squared Error of Adaptive Expectations ..................................... 142
Table 7.8: Root Mean Squared Error of Regressive Expectations ................................... 143
Table 7.9: Directional Accuracy of the Forecasts ............................................................. 144
Table 7.10: Directional Accuracy of the Forecasts using the Pesaran–Timmermann (PT)
Test ......................................................................................................................... 145
Table 7.11: Unit Root Test for Expected Spot Price and Futures Price at Log Levels ........ 149
Table 7.12: Unit Root Test for Expected Spot Price and Futures Price at First Difference .. 151
Table 7.13: Cointegration Regression—Extrapolative Expectations ................................. 154
Table 7.14: Summary Statistics for Expected Profit—Extrapolative Expectations ............ 154
Table 7.15: Cointegration Regression—Adaptive Expectations ....................................... 158
Table 7.16: Summary Statistics for Expected Profit—Adaptive Expectations ................. 158
Table 7.17: Cointegration Regression—Random Walk .................................................... 160
Table 7.18: Summary Statistics for Expected Profit—Random Walk ........................................160
Table 7.19: Cointegration Regression—Regressive Expectations ........................................162
Table 7.20: Summary Statistics for Expected Profit—Regressive Expectations ..................163
Table 8.1: Root Mean Squared Errors for the Heterogeneous Agent Model with Time-varying Weights ..........................................................................................................................185
Table 8.2: Root Mean Squared Errors for the Heterogeneous Agent Model with Constant Weights ............................................................................................................................................186
Table 8.3: Directional Accuracy of the Forecasts ................................................................187
Table 8.4: Directional Accuracy of the Forecasts Using the Pesaran–Timmermann (PT) Test ............................................................................................................................................187
Table 8.5: Heterogeneous Agent Model ..................................................................................189
Table 8.6: Unit Root Test for Expected Spot Price at Log Levels ....................................191
Table 8.7: Unit Root Test for Expected Spot Price at First Difference ..............................191
Table 8.8: Cointegration Regression—Heterogeneous Agent Model ................................193
Table 8.9: Summary Statistics for Expected Profit—Heterogeneous Agent Model ..........193
Table 8.10: Test of Coefficient Equality .............................................................................195
Table 8.11: Cointegration regression for the Heterogeneous Agent Model (January 2003–April 2013) .............................................................................................................................................196
LIST OF FIGURES

Figure 3.1: Equilibrium Futures Price under Varying Elasticities of Speculation Schedule...45
Figure 3.2: Equilibrium Futures Price under Varying Elasticities of Arbitrage Schedule ......46
Figure 5.1: West Texas Intermediate Spot and Futures Prices ............................................78
Figure 5.2: Crude Oil Price Decomposition to Trend and Cyclical Components ...............81
Figure 5.3: Trend and Cyclical of West Texas Intermediate Crude Oil Price at Log Levels ...82
Figure 5.4: Comparison between Basis and Cost of Carry .................................................83
Figure 7.1: Baltic Dry Index .................................................................................................121
Figure 7.2: OPEC Crude Oil Production .............................................................................123
Figure 7.3: Time Series Graph of US Dollar Trade-weighted Index ....................................124
Figure 7.4: OECD Crude Oil Inventory ..............................................................................125
Figure 7.5: OPEC Spare Capacity ......................................................................................126
Figure 7.6: Oil Price Forecasts—1 and 3-month Horizons..................................................138
Figure 7.7: Oil Price Forecasts—6 and 9-month Horizons..................................................140
Figure 7.8: Recursive Root Mean Squared Errors ...............................................................147
Figure 7.9: Expected Profit—Extrapolative Expectations ..................................................155
Figure 7.10: Futures Price Differentials ..............................................................................156
Figure 7.11: Expected Profit—Adaptive Expectations .......................................................159
Figure 7.12: Expected Profit—Random Walk ....................................................................161
Figure 7.13: Expected Profit—Regressive Expectations .....................................................164
Figure 8.1: Oil Price Forecasts —Heterogeneous Agent Model ..........................................183
Figure 8.2: Expected Profit—Switching Heterogeneous Agent Model.................................194
ABSTRACT

Since the crude oil futures price peaked at $147 per barrel, the role of speculators has come under tremendous scrutiny. The rise in oil price, along with the increased participation of financial traders, led to claims that speculators are responsible for exacerbating crude oil price movements. This study attempts to assess these claims regarding speculative influences on crude oil futures price. Specifically, the study explains the determination of crude oil futures price in terms of arbitrage and speculation. For this purpose a theoretical framework is formulated in which the observed futures price is outlined as a function of the arbitrage price and expected spot price. The arbitrage price, which reflects the influence of arbitrage, is obtained by adjusting the spot price by factors that account for the cost of carry and the convenience yield. The expected spot price, which represents the role of speculation, is estimated by using various expectation formation mechanisms. To determine the relation between the observed futures price, the arbitrage price and the expected spot price of crude oil, cointegration analysis is employed.

The empirical results reveal that although both arbitrage and speculation have a significant influence on futures price, arbitrage plays a dominant role in price determination. Moreover, the convenience yield effect is absent from the arbitrage process. The impact of speculation, on the other hand, varies according to the expectation formation mechanism employed and the length of the futures contract. Additionally, speculators form their expectations heterogeneously and switch between different expectation formation mechanisms on the basis of past forecasting performance. The results also show that in the build-up of oil prices after 2003, speculators remained focused on the expected profitability of their trading strategies. Overall, the findings firmly establish that financial traders impact the price dynamics of crude oil futures.
1.1 Background

Throughout the 1990s, the price of crude oil exhibited relatively stable behaviour while fluctuating within a narrow range. However, from 2001 the crude oil price began to rise sharply and reached $70 per barrel by 2006. It continued to rise and reached its peak of $143 per barrel in July 2008. By the end of the same year it had plummeted to $33 per barrel and then started to rise again in 2011. Interestingly the sharp fluctuations in the price of crude oil occurred alongside a large influx of financial capital into commodity markets. According to one report from the Institute of International Finance (Institute of International Finance, 2011), investment into commodity-linked derivatives increased from $10 billion in 2000 to $450 billion in 2011. Over this period, the number of open contracts in commodity exchanges almost doubled, and hedge fund trades tripled between 2004 and 2007 (Domanski and Heath, 2007).

The increased participation of commodity investors led some researchers and policymakers to assert that the rapid growth in commodity-linked investments is responsible for the unwarranted changes in the price of crude oil. They argue that financial traders view commodities as an alternative asset class for diversifying traditional stock and bond portfolios (Büyükşahin and Robe, 2012; Domanski and Heath, 2007; Gorton and Rouwenhorst, 2006; Singleton, 2012). Speculative buying by these financial traders has caused the price of crude

---

1 See, for example, Masters (2008)
oil to rise above the level warranted by supply and demand fundamentals.

This assertion has prompted considerable research on the issue of determining whether the participation of financial traders is responsible for the run-up in crude oil price. This issue is important as attempts to limit trading in crude oil futures and commodity-linked derivatives have been based on the claim that “financialization” of commodities has distorted the price of crude oil.

1.2 What is Financialisation?

To assess the role it plays in driving crude oil price, it is necessary first to understand the concept of financialisation. Broadly, the term is used to refer to the increasing dominance of capital market financial systems in the total economic activity (Dore, 2000). However, a more popular definition is the one provided by Epstein (2005): ‘financialisation means the increasing role of financial motives, financial markets, financial actors and financial institutions in the operation of the domestic and international economies’.

In the past, commodity markets had been viewed as being largely segmented from financial markets due to having little or no co-movements with equity price indices (Gorton and Rouwenhorst, 2006) as well as the price indices of different sectors (Erb and Harvey, 2006). However, this relationship changed in the early 2000s as the commodity markets attracted a large inflow of new capital. According to Gorton and Rouwenhorst (2006), the inflow of funds resulted from the collapse of the equity market at that time, which helped investors discover a negative correlation between equity and commodity returns. An inverse correlation with conventional assets and a positive co-movement with inflation encouraged investors to choose commodities as an alternative to traditional asset markets. As a result, various commodity-linked instruments and indices attracted billions of dollars of investment from
institutional and individual investors (Hughes, 2006). The increased appetite for commodity
derivatives by these financial investors initiated the financialisation process, which
(according to some commentators) exacerbated the rise and fall of crude oil price.

Market practitioners argue that the increasing participation of financial investors in the crude
oil futures market diminished the role of fundamentals in the crude oil price dynamics and
casted the price to be more volatile (Singleton, 2012). Accordingly, the focus has shifted
towards the financialisation process and the channels through which the demand for financial
assets affected the price of the crude oil.

1.3 How Financialisation Impacts the Futures Market
One of the main functions of the futures market is to assist the activity of hedging commodity
price risk. According to Keynes’ (1923) hedging pressure theory, hedgers must provide
speculators with a risk premium to hedge their spot price exposure. By involving speculators,
financialisation mitigates hedging pressure and allows for risk sharing as proposed by Tang
and Xiong (2012). However, Cheng et al. (2012) and Acharya et al. (2013) note that the risk
premium may be time varying due to capital constraints and the risk appetites of financial
investors. For example, speculators may have to unwind their positions and reduce risk when
facing large capital losses in other markets. The liquidation of their positions can lead to
outside risk being transmitted into the futures market. It follows that financialisation affects
the futures market through speculation in two ways: (i) through provision of liquidity by
speculators to accommodate hedging and (ii) through consumption of liquidity for
speculative gains.

Financialisation can also influence the information-discovery function of the futures market.
Information discovery refers to the incorporation of new information into commodity prices
(Garbade and Silber, 1983). Commodity prices are widely used to aggregate information regarding the strength of economic activity and to guide production decisions. Singleton (2012) argues that in the presence of information asymmetries, heterogeneous expectations of investors can cause futures prices to drift away from fundamental values. Similarly, Socking and Xiong (2013) show that informational noise introduced by heterogeneous traders can interfere with the expectations of producers and influence demand for the commodity. Therefore, the influx of speculative trading can distort the signalling mechanism of the futures market, leading to confusion regarding the strength of the economic activity.

Several commentators argue that if speculation drives the price of crude oil then changes in crude oil price must be accompanied by a significant accumulation of crude oil inventories (Hamilton, 2009; Juvenal and Petrella, 2012; Knittel and Pindyck, 2013; Kilian and Murphy, 2014). A critical shortcoming of this view is that it assumes that investors distinguish successfully between price changes induced by speculative trading and those resulting from shift in inventory demand. Moreover, even if speculative activity can be measured reliably by inventory changes, some speculators may only speculate on futures contracts with varying times to maturity without taking any positions in the physical commodity (Moosa, 2000).

Yet another set of studies focuses on the relationship between oil price changes and open interest in the futures market. Open interest is typically determined by using US Commodity Futures Trading Commission’s (CFTC) data on commercial and non-commercial traders (Gilbert, 2009, 2010; Singleton, 2011; Brunetti and Büyükşahin, 2009; Sanders et al., 2010; Stoll and Whaley, 2010; Sanders and Irwin 2011a, 2011b; Tang and Xiong, 2012). Commercial traders generally include hedgers whereas non-commercial traders are speculators.
However, the validity of this approach has been the subject of considerable debate due to the limitations associated with classifying open interest into commercial and non-commercial categories. Ederington and Lee (2002) and Sanders et al. (2004) argue that these classifications do not provide a sufficient account of the extent of speculative activity, particularly when trading motives vary from the underlying classification. For example, speculators may classify their trading activity as commercial hedging to circumvent position limits. In such instances, these designations fail to provide reliable estimates of the overall demand for and supply of commodity futures.

Other explanations for the behaviour of futures price are based on Working’s (1949) theory of storage (Garbade and Silber, 1983; Crowder and Hamed, 1993; Brenner and Kroner, 1995; Heaney, 1998, 2002b). According to the theory of storage, the futures price is determined by the net cost of carrying the commodity until the delivery date. This implies that futures price can be explained on the basis of arbitrage. Weymar (1968), however, rejects this hypothesis and argues that Working’s theory would be applicable only for shorter time intervals between spot and futures prices. At long time periods, expectations play a role. The importance of expectations is further stressed by Samuelson (1965), who posited that futures prices for contracts nearing maturity exhibit greater volatility than those for contracts with longer maturities. This line of reasoning suggests that futures prices can be explained in terms of arbitrage and speculation. Based on this idea, Moosa and Al-Loughani (1995) show that both speculation and arbitrage have significant influence on the price of crude oil.

While the ongoing research on financialisation has placed considerable emphasis on the relation between non-commercial activity in future markets and subsequent price changes, the theory of storage and the role expectations has received very little attention. This research addresses this gap by using the framework of Moosa and Al-Loughani (1995). For this
purpose, two important extensions are outlined. First, the role of convenience yield is incorporated to understand the influence of arbitrage activity. The early works of Kaldor (1939) and Working (1948, 1949) define convenience yield as an inverse carrying charge to explain the phenomenon of backwardation in the futures market. According to this definition, convenience yield represents the benefits that accrue to the commodity holder from the ability to smooth production, particularly during periods of limited availability.

Both earlier studies (Brennan, 1958; Telser, 1958) and recent research (Fama and French, 1987; Heinkel et al., 1990; Milonas and Thomadakis, 1997a, b; Heaney, 2002a; Hochradl and Rammerstorfe, 2011) on a number of commodities provide strong support for the presence of convenience yield. Hence, the current study proposes to update the estimation of arbitrage futures price by incorporating recent theoretical and empirical approaches that view convenience yield as having characteristics similar to those of an option contract. Also, recent theory suggests that price variability should influence the level of commodity held and consequently the convenience yield.\(^2\) This emphasises the role of price volatility in explaining the inter-temporal price spreads used in estimating storage costs. This study aims to incorporate these considerations into the estimation of convenience yield.

Second, this study extends the empirical work of Moosa and Al-Loughani (1995) by employing different expectation-formation mechanisms, in addition to rational expectations, to approximate oil price expectations. Although the rational expectation hypothesis (REH) has been used extensively in previous studies, empirical tests of expectation formation show that the expectations of individuals across different markets are on average inconsistent with the REH (Lovell, 1986; Frankel and Froot, 1987; Ito, 1990; Macdonald and Taylor, 1993; Pilbeam, 1995; MacDonald and Marsh, 1993; Ellen and Zwinkels, 2010; Prat and Uctum, 2010). See Khoury and Martel (1989) for an example.
Instead, the findings indicate that market participants tend to shift from one expectation formation mechanism to another. In other words, the expectations are formed heterogeneously.

To summarise, although risk sharing is the primary function of the futures market, the distinction between hedgers and speculators in practice is less prominent. If speculation is characterised as trading for profit by anticipating price changes then all groups of traders can be termed as speculators. An analysis on the basis of trading motives can therefore be economically more relevant for classifying trading activity than classifications based on trader status (Cheng and Xiong, 2013). Moreover, the role of expectations has become increasingly important in understanding the influence of speculation based on trading motives (Moosa and Al-Loughani, 1995; Moosa, 2000; Ellen and Zwinkels, 2010; Reitz et al., 2012). Investors can hold different expectations regarding future economic events, which can give rise to rich patterns in trading volume, asset returns and price volatility (e.g., Cao and Ou-Yang, 2009; Banerjee and Kremer, 2010). Nimark (2009) and Xiong and Yan (2010) show that trader groups with different expectations will eventually engage in speculative activity with each other. Allen et al. (2006) demonstrate that differential information between investors can lead to a bias towards public information, which in turn can exacerbate market volatility. Given this evidence of dispersion in investor beliefs and informational frictions, the question that arises and is particularly pertinent to this study is how heterogeneity in the expectations of speculators influences the price of crude oil. This study attempts to address this question by extending the framework of Moosa and Al-Loughani and allowing expectations to be formed heterogeneously. Also, an effort is made to determine if non-linear models (for example, see Brock and Hommes, 1997; Jong et al., 2009; Ellen and Zwinkels, 2010; Singleton, 2011) provide a more precise approximation of the expectation formation mechanism of traders.
1.4 Motivation for the Study

The aim of this study is to formulate a model for determining price of crude oil futures with the ultimate objective of analysing the role of financial trading in the process. The motivation for the inquiry is based on the following considerations:

1. Despite the abundance of literature analysing the effect of financial trading on the price of crude oil, only a handful of studies have examined the link between observed changes in the crude oil futures price and the behaviour of financial traders. In addition, these studies are limited to traders implementing technical trading rules only, with little consideration towards market fundamentals (Moosa and Al-Loughani, 1995; Moosa, 2000; Wang, 2004; Ellen and Zwinkels, 2010).

2. Sophisticated econometric techniques and a rich repository of data have allowed for a wide body of research into crude oil price. Despite this, a robust analysis of the heterogeneity in the behaviour of crude oil futures traders is yet to be performed.

3. The empirical evidence shows that the changes in net positions of speculators in the futures market are consistent with contrarian and momentum trading strategies. This evidence makes it necessary to analyse the extent to which the profitability of these trading strategies explains movements in the price of crude oil futures. The analysis can also benefit by establishing inferences as to whether a particular trader type may have a destabilising influence on futures prices.

4. A few studies have examined the explicit role of expectations in determining the price of crude oil futures. However, the role of predictive accuracy of the expectations formation mechanisms with respect to crude oil futures price remains largely unaddressed.

5. Futures prices have been explained in terms of arbitrage and speculation in some past studies. The evidence documents a significantly dominant role of arbitrage (relative to speculation) in determining futures prices (Moosa and Al-Loughani, 1995; Moosa,
However, the expectations hypothesis stipulates that the role of speculation becomes increasingly important as price uncertainty increases (Stein, 1979). The increase in variability of crude oil prices in the last decade provides a valuable setting for examining the relative importance of arbitrage and speculation in explaining the prices of crude oil futures.

This study takes a step towards addressing these gaps in the futures pricing literature. It adopts the framework proposed by Moosa and Al-Loughani (1995) to examine the effectiveness of arbitrage and speculation in explaining the price of crude oil futures. The objective of this empirical analysis is to answer the following key research questions: (i) what is the impact of arbitrage activity on crude oil futures price and does the inclusion of convenience yield in future price estimation improve the effectiveness of arbitrage?; (ii) does speculation play a significant role in determining the price of crude oil futures and, if so, is the impact of speculation sensitive to the underlying expectation formation mechanism?; (iii) do speculators in the crude oil futures market behave heterogeneously in forming expectations regarding future movement of prices; and (iv) is the predictive accuracy of expectation formation mechanisms an important factor in explaining the impact of speculation on crude oil futures price?

1.5 Structure of the Thesis

This thesis is structured as follows: Chapter 2 presents a survey of the relevant literature on the impact of the financial trading on commodity future prices. To identify gaps in the literature, a comprehensive review of different approaches for measuring speculative activity is conducted. For the studies reviewed, emphasis is placed on the limitations of different empirical approaches and measures of speculative activity.
Chapter 3 establishes a theoretical framework for evaluating the effectiveness of arbitrage and speculation in explaining the price of crude oil futures, and for this purpose it draws upon the study of Moosa and Al-Loughani (1995). The framework identifies two types of traders in the futures market: arbitragers and speculators. The futures price is shown to be a positive linear function of the excess demand for futures contracts by arbitragers and speculators. The demand for futures contracts by arbitragers is represented as a positive function of the difference between the actual and theoretical futures prices. Speculative demand for futures, on the other hand, is indicated as a positive function of the difference between the futures price and expected spot price at maturity of the futures contract. Using the respective demand functions I am able to show that equilibrium futures price can be determined by the theoretical futures price and expected spot price.

Chapter 4 provides a discussion on the efficiency of the futures market in the context of the theory of storage. It describes the cost of carry model and components of net cost of storage. In particular, a detailed review of developments in the theoretical and empirical literature on convenience yield estimation is provided. Also, the cointegration theory and the findings of the previous literature, evaluating the efficiency of the futures market by using the cointegration approach are discussed.

Chapter 5 examines the role of arbitrage and speculation in determining the price of West Texas Intermediate (WTI) crude oil futures. This role of arbitrage is assessed on the basis of theoretical price (sometimes called the cost of carry price) of futures. In this regard, the assumption of zero convenience yield on crude oil stocks is relaxed. The option pricing framework outlined by Heaney (2002a) is employed to approximate the convenience yield arising from crude oil stocks. The role of speculation, on the other hand, is outlined as a function of the expected spot price at contract maturity. The rational expectations mechanism
is used to proxy the expected spot price. Subsequently, a long-run relationship between the observed and theoretical futures price and expected spot price is estimated using the Engle and Granger (1987) (EG) test approach. Finally, several coefficient restriction tests are conducted to assess the validity of the results.

Chapter 6 departs from the assumption of the rational expectation hypothesis (REH) to model oil price expectations. In this regard a survey of the literature investigating the role of trading behaviour in driving market prices is conducted, wherein the chapter focuses on the theory of bounded rationality and heterogeneous expectations. A systematic review of numerous studies is conducted and several different expectation formation mechanisms are outlined for generating the crude oil price expectations. The expectations process relies on historical price trends and market fundamentals for predicting oil price movements.

Chapter 7 extends the analysis performed in Chapter 5 in two ways. First, it focuses on the empirical validity of the heterogeneous expectations framework in explaining the influence of speculation on crude oil price. For this purpose a sophisticated set of oil price expectations of speculators is generated using the expectation formation mechanism outlined in Chapter 6. Second, it evaluates the predictive accuracy of the expectation formation mechanisms in terms of error magnitude and directional accuracy to determine if speculative influences on crude oil price are driven by the predictive performance of the expectation formation mechanism.

In Chapter 8, a behavioural model of oil price dynamics is estimated using the framework proposed by Ellen and Zwinkels (2010). In this framework, speculators are classified as chartists and fundamentalists. Speculators are allowed to switch between the technical and fundamentalist expectation formation mechanism based on the past predictive performance of
the underlying mechanism. The model enables determination of which group of traders (chartists or fundamentalists) exerts more influence on the price. To determine whether heterogeneous speculators are more influential (relative to purely fundamentalist traders) in the determination of the crude oil price, we test the hypothesis that the impact of speculation based on heterogeneous expectations is greater than that based on chartist and fundamentalist expectations. Chapter 9 provides a conclusion by reviewing the main results and the research objectives. Also discussed are the limitations of the analysis and potential avenues for future research.
2.1 Introduction

During the past decade the amount of investment funds flowing into the commodity markets has increased tremendously. According to the Commodity Futures Trading Staff Report, institutional holdings increased from $15 billion in 2003 to over $200 billion in 2008 (CFTC, 2008). Unsurprisingly, trading in crude oil futures and options contracts has also increased considerably over this period. The number of contracts outstanding, also known as ‘open interest’, experienced a more than three-fold increase and the number of traders almost doubled (Interagency Task Force on Commodity Markets, 2008). Strong growth in open interest has primarily been led by non-commercial traders, also called ‘speculators’ (Domanski and Heath, 2007; Tang and Xiong, 2010). The increased presence of speculators has led some practitioners to hypothesise that the unwarranted increase in the oil price has been driven by the mark-up of speculative activity in the futures market. This phenomenon has also been referred to as “financialization”.3

Debate on the impact of speculation on the prices of crude oil has intensified over the past few years as oil prices have become more volatile. To explain the stronger investor interest in commodities, some academics have suggested that increased variation in prices, together with

---

3 At the London Energy Ministers’ Meeting of December 2008, A.S. El-Badri, OPEC Secretary General, asserted that “in the summer, prices were driven to record highs by unlimited speculation, as the dollar weakened and investors sought cover in commodity market”. (http://www.opec.org/opecina/Speeches/2008/LondonEnergy.htm)
the growth in market liquidity and financial innovation, has attracted new investors in search for positive returns. For instance, Gorton and Rouwenhorst (2006) find that commodity futures can on average yield an excess return of 5% over Treasury bills. Gorton and Rouwenhorst further report that the risk premium on commodity futures has been about equal to that of stocks, and exceeds the risk premium on bonds. These findings lead Gorton and Rouwenhorst to classify commodity futures as an attractive asset class for diversifying traditional stock and bond portfolios. Büyükşahin and Robe (2014) investigate substitutability between commodity and equity returns by examining the extent to which equity–commodity market linkages can be predicted by the presence of financial investors. They find that speculative trading, particularly by hedge funds, does help predict the fluctuations in the magnitude of commodity and equity co-movements. In another study, Domanski and Heath (2007) note a positive association between the share of net long positions taken by speculators and returns in the oil and gas markets. Domanski and Heath conclude that the return considerations in the commodity markets have become more important for financial traders since 2001.

2.2 Evidence of the Impact of Financialisation

Given the mounting evidence for a positive association between commodity markets returns and the presence of financial investors, it is not surprising that many observers have been investigating the link between financial trading and the recent surge in the level and volatility of oil prices.\(^4\) It may be intuitively appealing to link the inflow of funds into the futures market to the rise in crude oil price above the level warranted by market fundamentals. The _prima facie_ evidence also appears to support this notion as the increase in price level

---

\(^4\) for example, see Gilbert, (2008), Büyükşahin and Harris (2011), Singleton (2012); Knittel and Pindyck (2013) and Beidas-Strom and Pescatori, (2014)
occurred in parallel to the rise in financial activity in the futures market. However, the impact of financialisation is less clear.

2.2.1 Crude Oil Prices and Inventory Dynamics

Several methods are used in the literature to examine the role of financial activity in driving oil prices. One strand of literature examining the impact of financialisation is motivated by the theory of storage. It is based on the proposition that if speculative activity is responsible for the surge in crude oil futures price then it must be accompanied by a significant inventory response. According to this view, the term structure of futures price provides information regarding change in inventories to the market participants. Expectation of higher oil prices would thus be viewed as a signal to boost inventory holdings.

Relying on this premise, Hamilton (2009) examines the role of speculation by analysing supply and demand imbalances in the crude oil inventory. Hamilton argues that in order to rationalise the speculation motive for the alleged oil price ‘bubble’ it is necessary to account for changes in the supply of and demand for crude oil inventories. Using a representative agent approach, Hamilton shows that an increase in oil price induced by speculative trading should result in the accumulation of crude oil inventories, provided that the short-run price elasticity of demand is very low. According to Hamilton, the same condition (low oil price elasticity of demand) must be valid to attribute the surge in crude oil price to fundamentals alone.

Like Hamilton, Krugman (2008) argues that the only way speculative trading can impact oil prices is through a build-up of crude oil inventories. Krugman states that Organization for Economic Cooperation and Development (OECD) inventory data do not indicate any

---

5 See Singleton (2012) and Cheng and Xiong (2013) for a brief review of the methodologies
accumulation of inventories during the alleged ‘oil bubble’ period, which suggests that the rise in oil prices is not a consequence of an influx of financial trading, but is due to a decline in supply and the higher growth experienced by emerging economies.

Knittel and Pindyck (2013) assess the role of speculation by including data on consumption, production and futures prices in addition to inventory changes. Additionally, comparisons are drawn between a counterfactual series of convenience yield (obtained in the absence of speculative activity) and the actual convenience yield potential (arising from speculative inventory accumulation). The study finds that the influence of speculation on crude oil prices cannot be rejected. However, the study rules out any role for speculation in exacerbating price cycles since 2004. Knittel and Pindyck (2013) suggest that unless the elasticities of oil supply and demand are close to zero (as initially proposed by Hamilton, 2009) the changes in the level of inventories are inconsistent with the financialisation view.

Kilian and Murphy (2014) also examine the role of financial trading in driving crude oil prices by employing a structural vector autoregressive (VAR) model. Their study attempts to address a key issue of the debate on financialisation—that of separating the impact of speculating on oil prices from that of fundamentals. The impact of financial activity is captured by speculative demand shocks, which are identified as the shifts in demand for above-ground oil inventories arising specifically from variations in expectations of oil prices. Kilian and Murphy find that although changes in speculative demand play an important role in the oil price shocks of 1979, 1986, 1990 and 2002, they are unable to explain the surge in oil prices after 2003.

Juvenal and Petrella (2012) employ a wider set of structural shocks to examine the impact of speculative trading on oil prices. Consistent with previous studies, Juvenal and Petrella find
that the real price of oil is primarily driven by changes in global demand. However, the study reports that speculation contributed to the increase in oil price from 2004 to 2008, which is a period involving significant flow of investment capital into commodity markets.

Kilian and Lee (2014) argue that the quality of inventory data plays a significant role in correctly identifying speculative demand shocks using structural models. They propose a new proxy for oil inventory that uses above-ground US crude oil stocks scaled by the ratio of OECD petroleum inventories over US petroleum inventories. The study finds that while speculative activity does influence the oil price it does not play a significant role in explaining the surge in oil price between 2003 and 2008. Kilian and Lee report that the speculative demand explains only 4–18% of the surge in oil price and that the bulk of the increase results from shifts in demand from emerging Asian economies and OECD countries. The study also finds that speculation does add a premium to the real price of crude oil during various oil price shocks. These include the Iranian Revolution, the Kuwait Invasion, the Iraq War and the Libyan crisis. According to Kilian and Lee, the speculative pressure surrounding these events arises out of concerns regarding the stability of oil supplies.

Einloth (2009) examines the role of speculation by analysing the joint behaviour futures price and the convenience yields arising from crude oil inventory. The empirical approach is motivated by the theory of storage, which posits that a decline in inventory level should result in an increase in crude oil price, causing an increase in marginal convenience yield. Einloth uses this relationship to identify speculative commodity holdings. The study finds that convenience yield increased for the majority of the rise in crude oil futures price up to $100 per barrel. However, the convenience yield declined as the price rose above $100 per barrel, in conflict with the theory of storage. Einloth associates the decline in convenience yield at higher price level to the increase in crude oil inventory holdings by speculators.
According to Cheng and Xiong (2013), one of the assumptions implicit in studies examining inventory response to oil price changes is that traders in equilibrium are able to distinguish between changes in oil price induced by speculative activity, and those induced by changes in supply and demand fundamentals. However, as pointed out by Singleton (2011), in the presence of informational frictions in the spot market this assumption may not be realistic. Cheng and Xiong (2013) argue instead that speculative activity in the futures market can distort the price discovery role, causing a price surge accompanied by a shift in demand that mistakes the rise in futures price as a sign of positive growth without an underlying inventory response.

Beidas-Strom and Pescatori (2014) also criticise studies examining inventory response for failing to identify whether or not speculative activity is related to market fundamentals. They assert that in defining speculation, these studies do not account for the varying motivations of speculators and vastly different time horizons. Moreover, Moosa (2000) notes that speculators may only speculate on future contracts with varying times to maturity, without taking any position in the physical commodity.

On the other hand, Singleton (2011) argues that representative agent models and simplified forms of demand and supply imbalances do not account for learning under uncertainty and heterogeneity of expectations as well as agency conflicts that could limit arbitrage and lead to price bubbles. According to Singleton, the speculative trading dynamics that rely on the conditional distributions of futures and spot prices are more complex than the static models pertaining to inventory demand by homogeneous agents.
2.2.2 The Role of Open Interest

A second strand of the literature adopts a different framework for identifying speculative activity and focuses on the direct linkages between oil price changes and the volume of open interest of traders in the futures market. The Commodity Futures Trading Commission (CFTC) maintains historical reports on open interest of traders in the futures market, which identify traders as either ‘commercial’ or ‘non-commercial’ participants in the futures market. These classifications are further divided into sub-categories. The commercial trader sub-categories include producers, commercial dealers, commercial manufacturers, and swap dealers. The sub-categories for the non-commercial traders include hedge funds, traders and floor brokers. These categories are responsible for approximately 80% of the total open interest in the crude oil futures market (Interagency Task Force on Commodity Markets, 2008). However, according to Vansteenkiste (2011), CFTC classifies the positions on the basis of business activity (commercial v. non-commercial) and not by trading purpose (speculation v. hedging).

Studies examining the link between trader activities and price changes tend to use non-commercial positions to signify the extent of speculative activity in the futures markets. These studies rely on the premise that contemporaneous futures returns must be correlated with the open interest. A test of financialisation is formulated on the assumption that observed changes in the position result entirely from the shifts in the demand for futures by financial traders. The market clears when total changes in the positions of financial traders match the shifts in the demand curves of financial and other traders. A formal test of the relationship between daily price changes and position changes involves an examination of

---

whether changes in the positions of different trader groups take place in advance of the gains from price changes.

According to Sanders et al. (2004), the amount of speculation is typically measured by using Working’s speculative index (also known as Working’s T). Developed by Working (1960), this index is a ratio of the positions held by the speculators to the positions held by hedgers. A high value of the index is generally interpreted as an indication of excessive speculative activity. Most of the evidence gathered by studies employing this approach suggests that changes in open interest have either weak or no impact on commodity prices. One such study conducted by the US Interagency Task Force on Commodity Markets examines whether changes in positions of various groups of financial traders cause changes in the price of crude oil at the New York Mercantile Exchange (Interagency Task Force on Commodity Markets, 2008). The study finds little evidence for the proposition that changes in net non-commercial positions systematically precede oil price changes. Conversely, the study finds that price changes in commercial positions lead to adjustment in the positions of many trader groups over the sample period.

Likewise, Gilbert (2008) finds that the impact of investor activity on metal prices is not significant. However, a significant influence of speculative positions on futures prices is reported for agricultural futures markets. On the other hand, the findings obtained by Sanders et al. (2010) raise doubts regarding the assertion that speculative trading is responsible for the unprecedented increase in the prices of agricultural futures prices. Similarly, using data from the commitment of traders (COT) report, Till (2009) rules out the role of excessive speculation in driving the prices of WTI crude oil futures.
Using alternative proxies for non-commercial positions, Büyükşahin and Harris (2011) investigate the statistical link between the position of non-commercial speculators and changes in crude oil futures prices. The study reports correlations between the changes in futures price and in speculative index of –0.007 for nearby contracts and –0.018 for all contracts. A Granger test of causality is also conducted on a bivariate VAR involving the price and position variables. The study finds no evidence of causality (unidirectional or bidirectional) running from position changes to changes in price for any trader group from 1 to 5 days.

Some studies report evidence in support of the financialisation view. For instance, Robles et al. (2009) study the relation between the lagged ratio of non-commercial to total reportable positions and the prices of several agricultural commodities. The Granger test of causality indicates that speculative positions in the futures market may have been responsible for driving the commodity price boom of 2007/08. Irwin and Holt (2004) find a significant but weakly positive association between the net long positions of large hedge and futures price volatility in 13 different markets.

Hong and Yogo (2012) show that under the assumption of limited risk absorption by speculators in the futures market, changes in open interest act as a more reliable predictor of changes in economic activity than do the prices of oil futures. Hong and Yogo (2012) test the hypothesis that changes in open interest can help predict changes in commodity returns. In their model, the open interest is mainly influenced by expectations of higher economic activity. Their empirical results show that changes in open interest are significantly and positive related to the changes in oil spot and futures prices.
It has been suggested that the practice of employing open interest faces several limitations, the first of which is that it ignores the incentives of hedgers to trade futures contracts. Cheng and Xiong (2013) show that changes in the positions of hedgers are more responsive to commodity price changes than to revisions in output forecasts. Sanders and Baker (2012) assert that the positive association between hedging demand and price changes could suggest willingness to capture price gains rather than hedging against price uncertainty. Using data on the aggregate positions of traders in 26 commodity markets, Kang et al. (2014) show that hedgers exhibit contrarian behaviour by selling commodities that experience relatively poor performance and buying commodities that provide higher gains. Hartzmark (1987) finds that commercial hedgers earn significantly positive returns on their short positions. These findings suggest that hedgers have various motives for participating in the futures market in addition to risk sharing.

Another limitation of using net position of traders as a measure of financial activity arises due to the treatment of hedging positions as exogenous when examining the relation between position changes and movement in prices. According to Cheng et al. (2012), the empirical approach must distinguish trades initiated by financial traders from those undertaken to accommodate other traders (e.g., hedging). The trades initiated by financial traders should be positively related to price changes, whereas trades stemming from the need to accommodate other traders should be negatively correlated with price changes. The assumption that all positions are exogenous may introduce the classic simultaneity bias. Board, Sandmann, and Sutcliffe (2001) show that simultaneity bias arises because trading activity cannot be assumed as weakly exogenous since both trading activity and price are jointly determined with the arrival of new information. The resulting simultaneity can lead to a downwardly biased estimation of the price impact (Cheng et al., 2013). These limitations have also manifested in most of the studies on the role of commodity index speculation discussed in the next section.
2.2.3 The Role of Index Investments in Driving Crude Oil Price

Studies examining financialisation have traditionally viewed commercial and non-commercial traders as the only two broad categories of participants in commodity markets. However, over the past few years commodity index investors have gained considerable attention. These investors seek exposure to commodity prices by undertaking commodity index investments, which is a popular investment strategy that allows investors to diversify risk by investing in a basket of commodities in a given commodity index. The commodity index derives its value from the basket of futures contracts, which helps in avoiding the cost of holding the commodity. Following maturity of the first month’s futures contract underlying the commodity index, investors transfer their positions to the nearest-to-maturity contract in a process referred to as a ‘rollover’. In this way, the index provides returns that are comparable to those obtained on long positions in listed futures contracts.

The role of commodity index investment in driving commodity prices became prominent after the testimony of Michael W. Masters before the US Senate (Masters, 2008). The testimony made the simple assertion (often referred to as ‘Master’s hypothesis’) that the surge in commodity futures prices was caused by the participation of index investors. Masters argued that the massive long positions of index investors caused a bubble and pushed commodity prices above the level warranted by the underlying fundamentals.

Gilbert (2010a) provides another explanation for commodity price bubbles by suggesting that excessive buy-side interest by index investors can result in an increase in the risk premium required by counterparties to take offsetting short exposure. Consequently, expectations for further price increases due to index buying may limit the ability of potential counterparties to take up short positions. In this way the impact of index investment may be observationally equivalent to that of traders’ extrapolative past price trends. The resulting illiquidity may
neutralise information considerations leading to speculative price bubbles.

Studies examining the so-called Master’s hypothesis typically employ a Granger causality test between commodity index investment position changes and changes in futures prices.\(^7\) The results obtained by these studies yield mixed evidence with regard to the impact of speculative activity and price bubbles. For example, Gilbert (2010a) reports evidence of index investment having a significant impact on crude oil, copper and aluminium prices. Moreover, index investors are found to rely on passive strategies involving the extrapolation of past price trends. In a subsequent study, Gilbert (2010b) finds a significant relationship between the prices of agricultural commodities and index investments.

Singleton (2011) finds that after controlling for stock returns, open interest and lagged futures returns, growth in commodity index position has a significant impact on the price of crude oil futures. The results of empirical investigation also show an increase of 0.272% in the WTI crude oil price for every one million contract increase in index position in the previous 13 weeks. Both Gilbert (2010a) and Singleton (2011) construct measures of index positions in the energy market by imputing agricultural commodities positions.

Tang and Xiong (2010) examine the role of index investments in the financialisation of commodity prices across a range of commodities. Using the net long positions of investors, Tang and Xiong measure the changes in return after 2004 for indexed and non-indexed commodities. The study finds that an increase in net long positions of index investors results in a greater increase in return correlations with crude oil than non-indexed commodities after 2004. Index investments are also found to contribute to the spill-over of price volatility on commodity markets. Lending support to the financialisation hypothesis, Tang and Xiong

\(^7\) Notable studies include Brunetti and Büyükşahin (2009), Büyükşahin and Harris (2011), Irwin and Sanders (2012), Irwin et al. (2009), Sanders and Irwin (2011a, 2011b) and Stoll and Whaley (2010).
assert that in addition to the supply and demand fundamentals, the risk appetite for financial assets and the behaviour of financial investors also play a significant role in the determination of commodity futures prices.

Brunetti and Büyükşahin (2009), on the other hand, use swap dealer positions to proxy commodity index fund positions. Granger causality tests are conducted between trader activities and returns for a number of futures markets. The study fails to find a significant link between the positions of swap dealers and subsequent returns on crude oil, corn and natural gas futures. Stoll and Whaley (2010) carry out regressions of inflows from index investors (obtained from COT reports) on a variety of agricultural futures returns. Granger tests are also performed to determine if changes in the positions of investors precede changes in returns. They find that long position changes of index investors do not predict the returns for agricultural futures contracts. Similar findings are reported by Sanders and Irwin (2010, 2011a, 2011b) for energy and agricultural futures markets.

However, there are problems with index investments data. Singleton (2011) argues that studies examining inflows of funds into index positions often rely on COT data to gather information on different trader types. Prior to 2009, however, COT reports separated traders into two broad categories: ‘commercial’ and ‘non-commercial’. According to Ederington and Lee (2002), such legacy designations do not provide enough information to classify traders as either hedgers or speculators. For example, some speculators may classify their trading activity as commercial hedging to exceed position limits in some markets. Using ratios of total long and short positions of different types of traders, Ederington and Lee find significant misalignments between long and short positions and the business activities of trader groups. They conclude that the classification of traders as commercial or non-commercial does not adequately encompass different motives of traders.
The use of commercial and non-commercial categories received further criticism after the emergence of commodity index trading (CFTC, 2006a, 2006b), which typically involves gaining long-side exposures to a wide range of commodity prices through standardised commodity indices. In this manner, commodity index trading differs from speculation as it is not based on the current or expected price of a single commodity; neither does it resemble traditional hedging activity, as it does not convey any information about spot market activity. As a result, the risk associated with an underlying futures position does not correspond to a position in the physical market. Hence, classifying commodity index traders as either commercial or non-commercial may yield a misleading assessment of trading activity. Any analysis using COT data could therefore provide an unreliable estimate of the impact of index speculation on commodity prices.

The above studies mainly focus on the relationship between trader activity and changes in prices. The findings presented so far do not provide clear evidence on the role of financialisation in distorting oil prices. Some studies have also raised concerns regarding the limitations of empirical design and quality of data. Hence, the central question as to whether the speculative activity (either by index investors or other types of speculators) caused the unprecedented movements in oil price still remains. Singleton (2011) asserts that in the presence of informational frictions, using observed trader activity is not sufficient to justify or rule out a speculative influence on the commodity prices. According to Singleton, the empirical approach to understanding the role of speculative activity in the futures market should instead place emphasis on the behaviour of market participants.

2.3 A Behavioural Finance Approach to Modelling Crude Oil Prices

Another strand of research has examined the role of financialisation using behavioural models. This strand of literature emerged mostly in response to the traditional 'efficient
market’ paradigm. In broad terms, it argues that market participants are often irrational and exhibit certain predictable biases; for example, overconfidence (Barber and Odean, 2001; Gervais and Odean, 2001), overreaction (DeBondt and Thaler, 1986), loss aversion (Kahneman and Tversky, 1979; Shefrin and Statman, 1985; Odean, 1998), herding (Huberman and Regev, 2001) and psychological accounting (Tversky and Kahneman, 1981). According to Lo (2004), such behavioural biases are central to understanding investment decision making under uncertainty. Hirshleifer (2001) also suggests that these behavioural anomalies have clear implications for asset pricing and are rapidly becoming part of the academic mainstream.

According to Moosa (2000), the role of behavioural aspects of market participants in determining futures prices can be explained in terms of three main hypotheses. The first hypothesis, which is based on Working’s (1949) theory of storage, posits that futures price is determined by the net cost of carrying the commodity until the delivery date. This implies that arbitrage is the only determinant of the futures price. Weymar (1968), however, rejects this hypothesis and argues that Working’s theory would be applicable only when the time interval between the spot and futures price or between various futures prices is not long enough. With longer time intervals, market expectations also play a role in determining futures price.

The second hypothesis, referred to as the ‘Samuelson effect’, stipulates that futures prices for contracts nearing maturity exhibit greater volatility than those for contracts with longer maturity periods (Samuelson, 1965). This hypothesis thereby underlines the importance of expectations in the formation of futures prices. The third hypothesis builds upon the Samuelson effect and states that the variability of futures prices is systematically higher.
during periods of high uncertainty (e.g., Stein, 1979). Once more, the role of market expectations is stressed.

Together these hypotheses suggest that futures prices can be explained by the joint impact of arbitrage and speculation. Based on this idea, Moosa and Al-Loughani (1995) model the crude oil futures price in terms of arbitrage and speculation. In their study, the influence of arbitrage and speculation is transmitted to the futures price through the supply of and demand for futures contracts. They identify two main types of participants in the futures market: arbitragers and speculators. The equilibrium price is then shown to be determined by the activities of these two types of traders.

The demand of the futures contracts by the arbitragers is represented as a positive function of the difference between the actual and theoretical futures prices, where the latter is estimated using the cost of carry model. The model postulates that the futures price of a commodity equals the cash price and net cost of storing the commodity until the delivery date. On the other hand, the demand for futures by speculators is expressed as a positive function of the difference between the actual futures price and the expected spot price at the maturity of the futures contract. Moosa and Al-Loughani (1995) use the rational expectation hypothesis (REH) to proxy the expectations of speculators. The findings of the study show that although arbitrage plays a dominant role, the influence of speculative activity cannot be rejected in explaining the movement of futures prices. Avoiding the limitations of the approaches discussed earlier, the study of Moosa and Al-Loughani (1995) provides a valuable insight into the role of financial trading in driving crude oil prices.

Two important aspects of the study of Moosa and Al-Loughani (1995) warrant careful deliberation. The first pertains to the assumption of zero convenience yield in estimating the
cost of carry price. They argue that because crude oil inventory is held primarily for speculative motive and is not used as an input in a production process, convenience yield is unlikely to be realised. The convenience yield refers to the benefits that are unique to possessing the commodity, particularly during periods of limited availability. The theory of storage posits that convenience yield is inversely related to the level of physical commodity held. According to Gibson and Schwatrz (1990), this relationship implies that a constant convenience yield may arise only under very restrictive assumptions. Refuting this assumption, Gibson and Schwatrz (1990) suggest a stochastic representation of convenience yield with constant variance.

Heaney (2002a) relaxes the assumption of constant variance and proposes a model in which convenience yield arises from the available option of trading the commodity in the intermediate market. The value of the option is equal to the economic profit obtained from selling the commodity at the maximum price over the contract maturity and then buying it back before the expiry of the contract.8 Heaney finds that the option price-based measure of convenience yield results in a significant reduction in the difference between the observed and theoretical futures prices. Milonas and Thomadakis (1997a, 1997b) also show that the decision to hold crude oil inventories provides convenience yield that is similar to the payoff structure of an option contract. Similarly, using and option-based approach, Heinkel et al. (1990) and Hochradl and Rammerstorfe (2011) report significant convenience yields that are inversely related to the inventory level of the underlying commodities. The importance of convenience yield is also highlighted by Brennan (1958), who argues it is necessary in explaining the backwards futures market (a situation where futures price falls below the spot price). Hence, in the case of physical commodities, it is necessary to extend the simple cost of

---

8 This estimated profit is conditional on a positive probability of high demand for commodity in the intermediate period and positive autocorrelation in the spot prices.
carry to include the impact of convenience yield. However, the question then arises regarding the extent to which the convenience yield estimate based on anticipating price rises (see Heaney, 2002a) explains the amount of mispricing and therefore the role of arbitrage.

The second important aspect of the study involves specifying the expectation formation mechanism of speculators. One of the dominant paradigms regarding expectation formation by market participants is the REH. First formulated by Muth (1961), the REH posits that expectations of the market regarding the future value of an economic variable are on average equal to the mathematical conditional expectations of the variable. In other words, the market expectations, which essentially represent the predictions of future events, are the same as the predictions of the underlying economic model. Under the REH, investors use all available information, both past and present, efficiently in forming their expectations. In this way, expectations are informationally optimal and only the arrival of unanticipated news can cause the market expectations to deviate from the actual outcome.

Although the concept of rational expectations (RE) may be appealing in theory, its usefulness in forecasting economic variables has received tremendous criticism. Perhaps the most common concern raised in this regard pertains to the assumption that rational agents use all the available information optimally and avoid any systematic errors in forming their expectations. According to Shaw (1984), this assumption imposes an inferential process requiring agents to learn from past forecasting errors and adjust their predictive method accordingly.

However, Gertchev (2007) argues that the learning requirement does not specify a reasonable period for individuals to learn from their errors or period before the error is realised and the cause is identified. He further asserts that even if learning occurs, an individual can still make
the same error again due to changes in conditions in a manner that makes the error unavoidable despite new knowledge. Evans and Ramey (1998) extend this argument and show that the expectations of agents approximate RE equilibrium only when the forecasting horizons are sufficiently long and the intensity of expectations calculation is high. As a result, at a given point some expectations will be rational whereas others will be non-rational. Direct tests of the error learning process also show that agents commit systematic errors, and indicate the presence of psychological biases in forming expectations regarding simple economic events (e.g., Tversky and Kahneman, 1974; Grether and Plott, 1979; Tversky and Thaler, 1990; Kahneman et al., 1991).

Another commonly cited criticism of the REH is that it requires agents to process all available information optimally. For expectations to be rational, agents must know the true probability distribution of the mathematical process governing the evolution of the underlying variable. According to Sargent (1993) and Evans and Honkapohja (2001), this requirement entails a complicated learning process whereby agents must anticipate the expectations of other agents. Consequently, the expectation formation mechanism resembles a game–theoretic process in which agents attempt to revise their expectations in response to the behaviour of other agents; but adjusting their expectations may in turn alter the expectations of other agents. Eventually the learning process converges to a Nash equilibrium under the assumption that agents have prior information regarding their opponent’s equilibrium strategies.

Although the above explanation may be intellectually satisfying, it imposes an unrealistic requirement of extremely strong computational abilities on market participants (Branch, 2004). Nachbar (1997, 2001,) and Foster and Young (2001) criticise this approach as being excessively demanding, and Grauwe and Grimaldib (2006) equate the necessary
computational and processing requirements to ‘God-like’ abilities. Similarly, Sargent (1993) states that “when implemented numerically or econometrically, rational expectations models impute much more knowledge to the agents within the model (who use the equilibrium probability distributions in evaluating their Euler equations) than is possessed by an econometrician, who faces estimation and inference problems that the agents in the model have somehow solved”. Evans and Ramey (1998, 2006) show that agents do not respond in an optimal manner to the behaviour of other agents as represented by a Nash equilibrium. Subsequently, Norman (2011) highlights that market participants attempting to optimise their forecasts of the underlying variable spend almost the entire forecasting horizon in approximating the RE equilibrium.

Empirical tests of expectation formation mechanisms also provide little support for the REH. For instance, using survey data for inflation and other variables, Lovell (1986) finds the REH does not explain market expectations. Thomas (1999) and Makiw et al. (2003) reject the REH in modelling inflation expectations. Dominguez (1986), Frankel and Froot (1987), Wakita (1989), Ito (1990), Macdonald and Taylor (1993), Pilbeam (1995) and De Drauwe and Markiewicz (2013) show that the behaviour of exchange rates is inconsistent with the assumptions of the REH. Instead, they find that market participants behave heterogeneously (i.e. they choose from a wide variety of forecasting methods in forming exchange rate expectations). Similarly, MacDonald and Marsh (1993), Ellen and Zwinkels (2010) and Prat and Uctum (2011) report evidence of heterogeneous behaviour among oil price forecasters. They also find that the choice of expectations formation mechanism is driven by its past performance.

The evidence of heterogeneous behaviour also extends to the futures market. For instance, Wang (2004) finds that speculators within the major futures markets exhibit contrarian
behaviour whereas hedgers tend to engage in positive feedback trading. He further reports that the choice of trading strategy is driven by relative profitability of the trading strategy. These empirical findings indicate that the REH provides little value in understanding the expectation formation process of market participants. Instead, the evidence suggests that traders in different markets adhere to different expectation formation mechanisms. In other words, they form their expectations heterogeneously.

2.4 Summary and Conclusion

The key issue addressed in this review of the literature is whether there is clear evidence of speculation having systematically driven recent fluctuations in the price of crude oil. The academic literature in this regard has adopted a number of empirical approaches leading to several different conclusions. Some studies highlight the lack of inventory response to rising crude oil price as a compelling reason to doubt the financialisation hypothesis. However, several commentators argue that in the presence of informational frictions faced by market participants, using observed changes in inventory demand is not sufficient to quantify the change in price attributable to speculative activity (Cheng and Xiong, 2013).

The other strand of literature adopts a different framework by examining the relationship between oil price changes and open interest in the futures market. According to Gorton et al. (2013) it has become customary for studies employing open interest to view commercial traders as hedgers and non-commercial ones as financial traders. A critical limitation of this practice is that it fails to identify the extent of speculative activity in instances when trading motives differ from the classification of the traders (Sanders et al., 2004). Similarly, Fattouh et al. (2012) argue that such classifications are not informative about the extent to which financial speculation in oil futures markets has been held responsible for distorting of oil prices. The conclusion emerging from the review of these studies is that there is a lack of
consensus on how speculation is defined. If speculation is characterised as trading for profit by anticipating price changes, then all groups of traders can be termed as speculators. According to Kilian and Murphy (2014), a common aspect in all speculative buying is that the buyer anticipates futures increases in oil prices. In this case, an oil company that is stockpiling crude oil due to potential disruption of oil supplies is anticipating a rise in prices. Given these arguments, an analysis on the basis of trading behaviour can be economically more relevant for classifying trading activity than classifications based on trader status (commercial or non-commercial).

However, the empirical work on behavioural explanations for crude oil price movements is limited with the exception of a notable study by Moosa and Al-Loughani (1995), who find that speculative activity in conjunction with arbitrage has a significant influence on the price of crude oil futures. The behavioural approach of Moosa and Al-Loughani provides an important insight into the role of financial trading in driving crude oil price. Two important extensions to their study are outlined. The first involves relaxing the assumption of zero convenience yield for estimating the theoretical futures price. This is in line with the use of convenience yield to explain the inter-temporal spread between nearby and distant futures prices in the literature.9 This extension is also warranted on account of the evidence of significant convenience yields documented by a number of studies across different markets.

The second extension pertains to modelling the expectations of speculators. Although the REH has been a dominant paradigm of financial market research, alternative approaches to modelling expectations have been gaining popularity. A survey of the studies employing these approaches shows that individuals across different markets form their expectations

---

heterogeneously. Moreover, the expectations of individuals across different markets are on average, inconsistent with the REH.\textsuperscript{10} Given the expanding empirical literature challenging the notion of rationality, it would be of great interest to understand how the expectation formation mechanisms of futures market traders influence crude oil future price. The study of Moosa and Al-Loughani (1995) in this regard provides the relevant framework and avoids the limitations associated with other studies examining the impact of financialization.

3.1 Introduction

According to the theory of storage, the price of a futures contract for a storable commodity is determined by the net cost of carrying the commodity until the delivery date. This indicates that the futures price is determined purely by arbitrage activity. However, Weymar (1968) argues that this relationship only holds if the maturity of the contract is short (implying that expectations come into play over longer durations). Samuelson (1965) also emphasises that market expectations play an important role in determining future prices, particularly during periods of high uncertainty. Roll (1984) later shows that the private expectations of speculators are important in explaining price changes in the futures market.

Given the above arguments, the price of futures contracts can be explained in terms of arbitrage and speculation, which influence the price through changes in the supply of and demand for futures contracts. The current study aims to examine the extent of the influence exerted by both arbitrages and speculators on crude oil futures prices. The theoretical framework adopted for this purpose is based on the model proposed by Moosa and Al-Loughani (1995) for measuring the effectiveness of arbitrage and speculation in the crude oil futures market. It relies on the premise that the sale and purchase of futures contracts by arbitrages and speculators does not arise from any physical use of the commodity but is motivated purely by potential financial gains. Ideally, arbitrages would enter the market whenever the futures contracts are mispriced relative to the commodity’s cost of carry. The demand for futures by speculators, on the other hand, would depend on the magnitude of the
difference between the expected spot price and the current futures price. At equilibrium, the futures price balances the forward commitments of arbitragers and speculators, given their risk and return preferences.

However, market frictions (such as transaction costs, liquidity constraints, risk premiums, etc.) may reduce the profitable opportunities available to the participants and constrain the effectiveness of the risk transfer mechanism. As a result, the futures price and the equilibrium quantity traded would be strongly influenced by the respective demand elasticities of arbitragers and speculators. A detailed discussion is therefore provided to elaborate on the conditions necessary for arbitrage and speculation, and their relative influence on futures prices under varying levels of demand elasticities.

3.2 The Role of Arbitragers

In an efficiently functioning market the price of a futures contract can be expressed as the sum of the spot price and the cost of holding the commodity until a specified delivery date. This mechanism of pricing futures contracts can be traced back to theory of storage, which is centred on the early works of Kaldor (1939), Working (1949, 1953) and Telser (1958). According to the theory of storage, a no-arbitrage relationship between spot and futures price can be stated as:

\[ F_{t,t+1} = S_t + C_{t+1} + r \]  

(3.1)

where \( F_{t,t+1} \) is the price at time \( t \) of a futures contract maturing at \( t + 1 \), \( S_t \) is the spot price of oil, \( C_{t+1} \) is the cost of the storage incurred until the delivery date in period \( t + 1 \) and \( r \) is the risk-free rate of return, which represents the opportunity cost of holding the commodity. The
sum of the storage and opportunity costs is referred to as the cost of carry.\textsuperscript{11}

In addition to storage costs, holding the commodity also entails certain benefits that arise principally from fluctuations in commodity demand and from increase in production utility due to availability during periods of limited supply (Pindyck, 1993; Milonas and Thomadakis, 1997b). The convenience yield approximates the value of these benefits to the commodity holder. It essentially acts as an ‘inverse carrying charge’ that is used to offset the costs arising from storage of the commodity (Brennan, 1958). Under the theory of storage the convenience yield is a decreasing function of the level of inventory held.\textsuperscript{12} An important implication of this relationship is that when inventories are low relative to demand, the marginal value of the convenience yield may exceed the marginal cost of physical storage, driving the spot price higher than the contemporaneous futures prices (‘backwardation’). In contrast, a lower convenience yield associated with sufficiently higher levels of inventory should lead to higher futures prices relative to the spot price (‘contango’). Hence, in addition to the cost of carry, the convenience yield may have strong predictive power for the term structure of futures prices.

However, Moosa and Al-Loughani (1995) argue that because financial traders hold crude oil inventory for speculative motives or for meeting contract obligations, and not for production or processing of the commodity, any production benefits are unlikely to be realised. Moreover, according to Pindyck (1993), convenience yield may be negligible or insignificantly different from zero where large quantities of inventories are held relative to production. Hence, Equation (3.1) encompasses the assumption of zero convenience yield.

\textsuperscript{11} The cost of storage includes storage space rent, handling or loading and unloading charges and insurance, etc. (Brennan, 1958).

However, it could be argued that convenience yield may still arise from the option of selling the commodity before contract maturity if the transaction is economically profitable (Heinkel et al., 1990; Dixit and Pindyck, 1994). Traders may invest the proceeds at the risk-free rate and then buy back the commodity at a lower price. However, this approach has a few limitations. First, to maximise the profit it is assumed that the commodity holder has perfect market timing ability, which may only be possible under very restrictive conditions (Heaney, 2002a). Second, inventory holdings are unlikely to change for risk averse investors, particularly when they are associated with corresponding short positions in the futures market (Moosa and Al-Loughani, 1995). These implications suggest that financial traders obtain zero convenience yield from storage. According to Brennan (1958), traders can be seen as merely possessing titles to the inventories.

Assuming that the futures market is strongly efficient, the price of a crude oil futures contract should differ from the spot price by an amount equal to the cost of carry.\textsuperscript{13} However, such a relationship between spot and futures prices is conditional on the ability of market participants to trade simultaneously in both markets (Bopp and Lady, 1991). Any violation of the cost of carry relationship represented by Equation (3.1) would result in profitable arbitrage opportunities. For example, if the price of futures contract $F_{t,t+1}$ exceeds the sum of the spot price and the cost of carry $(S_t + C_{t+1} + r)$, then a market participant can earn profit by borrowing money equivalent to $S_t + C_{t+1}$ at an interest rate $r$, buying spot oil at time $t$ and simultaneously selling short a futures contract. The resulting arbitrage trading would drive the spot price up and force the futures prices downwards until they differ no more than the carrying cost. Similarly profitable arbitrage opportunities would arise whenever:

\textsuperscript{13} Efficiency in futures market has been generally couched in terms of predictability of expected spot prices using futures prices (Chowdhury, 1991; Crowder and Hamed, 1993; Green and Mork, 1991; Heaney, 2002b; Moosa and Al-Loughani, 1994).
where $C_{t+1}^*$ represents the sum of the cost of storage and the interest expense. In this case, rational arbitragers can earn riskless profit by shortening the commodity in the spot market, investing the proceeds at a rate of return $r$ and going long in the futures market. Profit taking by arbitragers will create upward pressure on the futures price, whereas selling in the spot market drives down spot prices until the price differential is eliminated.

The restoration of equilibrium in Equation (3.1) is based on several key assumptions that include (i) infinitely elastic supply of arbitrage services, (ii) no taxes or transaction costs, (iii) unlimited borrowing, (iv) the term structure of interest rates is stationary, and (v) no short sale constraints. If these conditions hold and the futures prices are purely determined by arbitrage, then the basis $B_t$ is given by:

$$B_t = F_{t,t+1} - S_t = C_{t+1}^*$$

(3.3)

However, in imperfect market settings several of these assumptions are likely to be modified. For example, heterogeneity in the location and grade of the underlying commodity (Pindyck, 2001), storage constraints (Working, 1953) and restrictions on the available arbitrage capital (Shleifer and Vishny, 1997) may cause the spread between spot and futures prices to change and reduce the substitutability within the two markets (Garbade and Silber, 1983). The elasticity of supply of arbitrage services would therefore be less than infinitely elastic.

Moreover, if market participants are averse to fluctuations in commodity prices, (e.g. due to changes in inventory as postulated by Deaton and Laroque, 1992) a time-varying risk premium may be incorporated into the futures prices.\textsuperscript{14} The risk premium embodies the

\textsuperscript{14} The risk premium can be attributed to the theory of normal backwardation (Keynes, 1930), which characterises the premium as a compensation realised by speculators for assuming the risk of deviation of futures prices from expected future spot price (Carter et al., 1983).
uncertainty associated with future behaviour of prices. Equation (3.2) can therefore be stated as:

\[ F_t = S_t + C_{t+1}^* + \rho_t \]  

(3.4)

where \( \rho_t \) is the risk premium such that \( \rho_t > 0 \) or \( \rho_t < 0 \) depending on the degree of uncertainty and the number of market participants. The basis in this case can be given by:

\[ B_t = F_{t,t+1} - S_t = C_{t+1}^* + \rho_t \]  

(3.5)

According to Equation (3.5), the basis may deviate from the cost of carry. Also if \( \rho_t < 0 \) and \(| \rho_t | > C_{t+1}^* \) then \( F_{t,t+1} > S_t \) and \( B_t < 0 \). In this case, the compensation required by other market participants (particularly speculators) for holding offsetting positions in the futures market would essentially limit the availability of profitable arbitrage opportunities. This implies that arbitrage alone may not be adequate for explaining the formation of equilibrium prices in the futures market.

3.3 The Role of Speculators

The role of speculators can be illustrated by expressing the futures price in terms of the expected spot price for the commodity. If the expected spot price equals the futures price, speculators have no incentive to buy or sell the futures contract. This condition can be expressed as:

\[ F_{t,t+1} = E_t S_{t+1} \]  

(3.6)

where \( E_t S_{t+1} \) is the expected spot price of crude oil at time \( t + 1 \) conditional on the information that is available at time \( t \). However, if the expected spot price of crude oil is less than the futures price then it would be profitable to buy and hold the spot commodity. Spot oil purchases would drive up \( S_t \) until profitable opportunities are eliminated. Similarly, if:

\[ F_{t,t+1} < E_t S_{t+1} \]  

(3.7)
then speculators will buy futures contracts and sell the spot commodity until the underlying incentive disappears. The one-period expected return for speculators is $F_{t,t+1} - E_t S_{t+1} = \rho_t$. This argument suggests that futures prices may be strongly influenced by arbitrage and speculation whenever price discrepancies exist. However, another category of traders pertinent to the oil futures market is represented by hedgers who enter into futures contracts essentially to mitigate the risk of adverse oil price movements by fixing in advance the price for a future delivery or receipt. The resulting hedging activity may exert considerable influence on the supply of and demand for futures. Acknowledging the role of hedgers therefore may be necessary for explaining the formation of equilibrium futures prices. However, it can be argued that by taking long and short positions in futures market hedgers act in a manner that is similar to speculators. For instance, the expected cost of hedging is equal to the difference between the futures price and the expected spot price at maturity: $F_{t,t+1} - E_t S_{t+1}$. This is also equal to the amount of loss the hedgers are willing to incur to avoid price risk. The smaller the cost of hedging ($F_{t,t+1} - E_t S_{t+1}$), the higher will be the demand for futures contracts. Hence, excess demand for hedging with long positions is positively related to $F_{t,t+1} - E_t S_{t+1}$.

On the other hand, the expected cost of hedging for short hedgers is $E_t S_{t+1} - F_{t,t+1}$. As short hedgers are net suppliers of futures contracts, their excess demand will increase as $E_t S_{t+1} - F_{t,t+1}$ (expected cost of hedging) increases. Once again, excess demand of hedgers is positively related to $E_t S_{t+1} - F_{t,t+1}$. It must be noted here that for long hedgers the expected cost of hedging, $F_{t,t+1} - E_t S_{t+1}$ is equal to the expected profit of a speculator with a short position in the futures market and a long position in the spot commodity. For short hedgers, the expected cost of hedging, $E_t S_{t+1} - F_{t,t+1}$, is equivalent to the expected profit of a speculator with a long position in the futures market and a short position in the spot market.
As the excess demand for futures contracts by hedgers and speculators is a function of the same variables, it makes sense to treat arbitragers and speculators as the only two trader categories in the proposed model.

This treatment is also consistent with approaches adopted in some other studies. For example, Stoll (1968) interprets a trader’s decision to hedge against the risk of price decline as arbitrage, and consequently the decision not to hedge the commodity as speculation. Similarly, Moosa and Al-Loughani (1995) and Moosa (2000) do not distinguish between hedging and speculation to model the demand for oil futures. Cheng and Xiong (2013) also note the similarity between hedging and speculation, and state that “while risk-sharing is a central function of futures markets, the line between hedging and speculating is blurred in practice. If one takes the notion of speculation as trading for profit from price movements, all trader groups including commercial hedgers appear to be engaged in speculation on the margin”. Given these arguments, we treat arbitragers and speculators are treated as the only actors in the futures market in the current theoretical framework.

3.4 A Diagrammatic Exposition

This section illustrates the relative effectiveness of arbitrage and speculation under varying levels of demand elasticities. The excess demand functions for arbitragers and speculators are assumed to be downward slopping. The demand of arbitragers can be expressed as a linear function of the difference between the futures price determined by the cost of the carry, $\bar{F}$, and the actual futures price $F$. If $\bar{F} = F$, excess demand for futures would be zero as arbitragers have no incentive to buy or sell.

The excess demand of speculators, on the other hand, depends on the difference between the actual futures price and expected spot price, $ES$. Speculators would have no incentive to trade
futures contracts if \( F = ES \), in which case excess demand will be zero. For the diagrammatic illustration it is assumed that \( ES < F < \bar{F} \), which implies that futures contracts are being traded between net long arbitragers and net short speculators. The equilibrium position in the short run is where the net long positions of arbitragers is equal to the net short positions of speculators. Assuming a finite elasticity of excess demand, it can be shown that the relative effectiveness of arbitrage and speculation is proportional to the relative elasticities of the two excess demand schedules.

Figure 3.1 shows an unchanged arbitrage schedule (AA) and multiple speculation schedules (SS) that vary in elasticity around the vertical axis representing prices. Assuming a constant risk tolerance of market participants, the combination of schedules AA and S₁S₁ results in a cost of carry futures price \( F \), an expected spot price \( ES \) and an actual futures price \( F_1 \). In this case, the quantity of futures contracts bought by arbitragers, \( QA_1 \), is equal to the quantity of contracts sold by speculators, \( QS_1 \). The futures price \( F_1 \) is at equal distance between \( ES \) and \( \bar{F} \), which suggests that the futures price is equally responsive to arbitrage and speculation. This result is obtained because AA and S₁S₁ schedules have similar elasticities.

On the other hand, the effectiveness of arbitrage increases (relative to speculation) when the elasticity of the speculation schedule decreases (S₂S₂). An inelastic speculation schedule can be attributed to a high degree of risk aversion among speculators. For a given expected spot price \( ES \), the speculators require a higher absolute risk premium to sell additional futures contracts. The combination of schedules AA and S₂S₂ results in an equilibrium futures price \( F_2 \) which is closer to \( \bar{F} \).
Alternatively, the effectiveness of speculation is enhanced when the elasticity of the speculation schedule increases relative to arbitrage ($S_1 S_3$). In Figure 3.2, the slope of the speculation schedule is kept constant whereas the arbitrage schedule is allowed to vary such that the equilibrium quantities of futures contracts and futures price remain unchanged.
Similar to Figure 3.1, three different arbitrage schedules (A1A1, A2A2 and A3A3) are obtained, producing three different arbitrage prices ($F_1$, $F_2$ and $F_3$). The equilibrium futures price is equally influenced by arbitrage and speculation when the two schedules have the same elasticities. The futures price $F$ in this case is at equal distance between $F_1$ and $ES$. Whenever the arbitrage schedule (A2A2) is flatter than the speculation schedule, arbitrage exerts greater influence on futures price and $F_2$ is closer to $F$ than to $ES$. Finally, when the excess demand by arbitragers is steeper in comparison to demand from speculators (A3A3), $F$ is closer to $ES$ than to $F_3$ which indicates dominance of speculation.
3.5 Model Specification

Based on the excess demand functions of arbitrage and speculation, a testable model for determining their relative effectiveness can now be specified. For arbitragers, the amount of futures contracts demanded can be expressed as follows:

\[ Q^a_t = \alpha_a(\bar{P}_{t,t+n} - F_{t,t+n}) \alpha_a > 0 \quad (3.8) \]

where \( \alpha_a \) is the elasticity of demand by arbitrageurs, \( \bar{P}_{t,t+n} \) represents the cost of carry price of a futures contract maturing at \( t + n \) and \( F_{t,t+n} \) is the actual price of the contract. Similarly, demand by speculators is given as:

\[ Q^s_t = \alpha_s(E_t S_{t+n} - F_{t,t+n}) \alpha_s > 0 \quad (3.9) \]

where \( \alpha_s \) is the elasticity of the excess demand. The expectations operator \( E(.) \) represents the expectation of spot price of crude oil at time \( t + n \) based on the information available at time \( t \). Assuming speculators and arbitragers are the only two actors in the futures market, the equilibrium can be shown to be the point where the net buy-and-sell commitments of arbitragers are equal to the net buy-and-sell commitments of speculators so that the excess demand is zero. Equilibrium occurs when:

\[ Q^s_t = -Q^a_t \quad (3.10) \]

or

\[ Q^s_t + Q^a_t = 0 \quad (3.11) \]

By substituting the values of \( Q^a_t \) and \( Q^s_t \) into Equation (3.11) we get an \( n \)-period equilibrium futures price, which is the weighted sum of the expected spot price and the cost of carry futures price. This is given as:

\[ F_{t,t+n} = \left( \frac{\alpha_a}{\alpha_a + \alpha_s} \right) E_t S_{t+n} + \left( \frac{\alpha_s}{\alpha_a + \alpha_s} \right) F_{t,t+n} \quad (3.12) \]
In a similar analysis of the foreign exchange market, Callier (1980) assumes greater flexibility for speculation by highlighting the implicit option of trading in the intermediate period of the contract. According to Callier (1980), traders may speculate at any time before the contract maturity by engaging in two offsetting positions at two different times but for the same maturity. This implies that in addition to the expected spot exchange rate the expected forward exchange rate over the intermediate period may also influence the forward prices. If Callier’s (1980) argument is extended to the futures market, it would allow for speculation on the futures price in a manner similar to speculating on the expected spot price. For example, the excess demand for a futures contract maturing at time $t + 2$ depends on the futures price expected to prevail at time $t + 1$. This may be expressed as:

$$Q^f_t = \alpha_{21}(F_{t+1,t+2} - E_t F_{t+1,t+2}) \alpha_{21} > 0$$

where $E_t F_{t+1,t+2}$ represents the price (estimated at time $t$) expected to prevail at time $t + 1$ of a futures contract for delivery at time $t + 2$. The coefficient $\alpha_{21}$ accounts for the slope of the excess demand schedule for speculators. In this case, equilibrium is obtained when:

$$Q^a_t + Q^s_t + Q^f_t = 0$$

(3.14)

If we include this new speculative dimension for pricing of futures then Equation (3.12) can be restated as:

$$F_{t,t+2} = \left( \frac{\alpha_a}{\alpha_a + \alpha_s + \alpha_{21}} \right) E_t S_{t+2} + \left( \frac{\alpha_s}{\alpha_a + \alpha_s + \alpha_{21}} \right) F_{t,t+2} + \left( \frac{\alpha_{21}}{\alpha_a + \alpha_s + \alpha_{21}} \right) E_t F_{t,t+1}$$

(3.15)

The model can be extended further to incorporate futures contracts with maturities up to $n$ periods, which then allows for speculation from $t + 1$ to $t + n - 1$. In that case, the equilibrium futures price can be expressed as:
where $\alpha_{nk}$ measures the sensitivity of the futures price to the speculative demand based on expected futures prices between the periods $t$ and $t + n - k$.

### 3.6 A Testable Model for Futures Prices

An empirical model can be specified for the equilibrium futures price obtained in Equation (3.16). The model adopted here has been drawn from the study of Moosa (2000) on the influence of speculation and arbitrage on the prices of crude oil futures. A testable form of the model is as follows:

$$F_{t,t+n} = \beta_0 + \beta_1 E_t S_{t+n} + \beta_2 F_{t,t+n} + \sum_{k=1}^{n-1} \gamma_k E_t F_{t+n-k} + \varepsilon_t$$

where $\beta_1$ and $\beta_2$ measure the sensitivity of futures prices to arbitrage and speculation on expected spot prices, and $\gamma_k$ accounts for the influence of speculation on the expected futures price at $k$. A larger value of the coefficient implies a greater role for the activity associated with that coefficient. A coefficient value close to 1 implies complete dominance of the underlying activity.

### 3.7 Summary

This chapter outlines the theoretical framework for examining the role of arbitrage and speculation in determining crude oil futures prices. If arbitragers and speculators are the only participants in the futures market, the equilibrium futures price can be expressed as a linear function of their excess demands. Excess demand of arbitragers is derived from the discrepancy between the futures price and sum of the physical and financial costs of holding
the commodity. Similarly, speculative demand is expressed as a positive function of the difference between the futures price and expected spot price. Further, allowance is made for speculation on the basis of expected futures prices over the intermediate period of the contract.

The presence of market frictions (such as capital restrictions) and risk premiums results in excess demand functions with finite elasticities. The relative effectiveness of arbitrage and speculation is then shown to be positively related to the relative elasticities of their respective demand schedules. Finally, by using the futures market equilibrium condition of zero net excess demand, a testable model is derived where the futures price is dependent on the weighted sum of the cost of carry price, the expected spot price and the expected futures price.
4.1 Introduction

One of the most important roles of the futures market is that of price discovery. Accurate anticipation of changes in the cash and futures basis is essential for market participants to use futures contracts to minimise risk. To anticipate these changes accurately, it is important to gain some understanding of the behaviour of the basis and the long-run relation between cash and futures prices. According to one view, the temporal spread between cash and futures prices can be explained in terms of the cost of carry model posited by the theory of storage. This theory postulates that the spread between cash and futures prices of a commodity over a certain period should be equal to the net cost of carrying the commodity over that period. An inequality in this relationship constitutes a mispriced futures contract, which could lead to the occurrence of arbitrage opportunities. In this case, the futures market cannot undertake effectively the role of price discovery.

Whether or not futures markets can perform the price discovery role efficiently is of great significance to market practitioners or indeed to any investor attempting to hedge against price risk. If futures markets do not perform efficiently, the futures price becomes a biased predictor of the future spot price. Consequently, participants in the futures market will incur costs in addition to the transaction costs. These costs are essentially of two types. The first type of cost may arise from the riskless return to the arbitragers engaging in the cost of carry arbitrage. The second type of cost arises from the compensation demanded by the speculators.
for assuming the risk of future spot price changes. Evaluation of these types of cost encompasses the issue of market efficiency.

4.2 The Cost of Carry Model

The price of a futures contract is generally expressed in terms of the theory of storage, which dates back to the early works of Kaldor (1939), Working (1949, 1953) and Tesler (1958). The theory asserts that the difference in the prices of a commodity on the same day for two different delivery dates is equivalent to the net cost incurred by storing the commodity until the delivery date. The difference in the prices is a necessary incentive for commodity holders to store the commodity. The theory of storage may be formalised as follows:

\[ F_{t,t+n} = S_t e^{[n(w_t + r_f - cy)]} \]  (4.1)

where \( F_{t,t+n} \) is the time \( t \) price of a futures contract that is maturing at time \( t + n \), \( S_t \) is the spot price of the commodity at time \( t \), \( w_t \) is the cost of storage, \( r_f \) is the financial cost of carrying the commodity, \( cy \) is the convenience yield and \( n \) is the time to the maturity of the futures contract. The right-hand side of Equation (4.1) can also be referred to as the cost of carry or the theoretical price of the futures contract.

The cost of storage includes rent of the storage space, handling and transportation charges and insurance. The financial cost of the commodity pertains to the opportunity cost of the funds invested in the commodity. Equation (4.1) indicates that the price of a futures contract for a commodity is equivalent to the spot price of the commodity and a continuously compounded net cost of carrying the commodity. The model also suggests that the net cost of carry would be positive if the total cost of holding the commodity exceeds the convenience yield. This ensures that the commodity’s price for immediate delivery is lower than the price for future delivery. However, if the convenience yield exceeds the sum of cost of storage and the opportunity cost, then the net cost of carry becomes negative. This in turn would imply
that the spot price of the commodity would be higher than the futures price. The convenience yield therefore has important implications for the term structure of futures prices. The cost of carry must include the effect of the convenience yield to reflect the benefits realised from holding the commodity.

4.3 Convenience Yield and the Cost of Carry

The theory of storage posits that the inter-temporal spread between the distant and nearby futures prices or cash prices of a commodity is determined by the net cost of storing the commodity over the time interval. A negative spread is observed when nearby prices exceed distant prices (backwardation) and a positive spread results when distant prices are higher than the nearby prices (contango). Based on the theory of storage, rational economic agents would hold the commodity only when returns to storage are positive (Working, 1949). However, the evidence of inventory accumulation in backworded markets has led to a debate regarding the causes underlying this irrational behaviour.\footnote{Brennan (1958), Gray and Peck (1981), Telser (1958) and Weymar (1966).} Explanations for holding stocks during negative inter-temporal spreads have mainly centred on the idea of convenience yield introduced by Kaldor (1939).

The convenience yield on a commodity reflects the benefits accruing to the commodity holder. These benefits are represented by increased utility arising from the availability of the commodity during periods of scarce supply and high price volatility (Brennan, 1958). Under such conditions, the possession of inventory allows the firm to respond flexibly to unexpected changes in supply and demand conditions. Holding stocks of the underlying commodity permits the firm to vary the production level at a lower cost than would be incurred in the absence of stocks (Telser, 1958). The benefits of holding the commodity may also be realised in the form of the option of selling the commodity in the high-priced intermediate market.
The resulting convenience yield can therefore be described as an implicit revenue accruing to the commodity holder. This revenue essentially represents an inverse carrying charge that is used to offset the cost of holding the inventory.

For storable commodities such as crude oil the inventories play a critical role in mitigating the cost associated with adjusting production arising from fluctuations in the commodity’s demand. Having the inventory on hand also helps in ensuring timely deliveries and avoiding stockouts (Pindyck, 2001). However, according to Kaldor (1939), as the inventory rises above the level that is considered adequate for reducing the production and marketing costs, the value of the convenience yield on each additional unit declines. This would suggest an inverse relationship between the level of inventory and the convenience yield. Empirical research on a large number of commodities provides strong support for this relationship. For instance, Telser (1958) formulates a theory for the equilibrium stocks held by a firm using the demand for and supply of stocks schedules. He asserts that the stocks held by the firm when prices are expected to fall (negative inter-temporal spread) can be attributed to the convenience yield arising from the stocks. The theory is confirmed by the empirical findings of the study, which indicate an inverse relationship between the wheat and cotton inter-temporal futures price spreads and the respective inventory levels.

Using the supply and demand functions of storage, Brennan (1958) develops a two-period model to describe the storage behaviour of a competitive firm. To maximise profit, the firm will hold the stocks up to the point where the marginal cost of storing is equal to the expected revenue per unit of time. The net marginal cost of storage is defined as the marginal outlay on the physical storage, plus a marginal risk aversion factor, less the marginal convenience

---

16 Kaldor (1939) refers to this level as the ‘useful’ amount of inventory.
yield.\textsuperscript{17} The study finds that the monthly ending stocks for various agricultural commodities are inversely related to the net marginal storage cost.

Weymar (1966) develops a model to explain the spread between the current price and expected price of a commodity at some future time. According to Weymar, the inter-temporal spread between prices is a function of the expected inventory behaviour over the intervening interval. Weymar reports that the estimated increase in the inventory of the African main crops led to a change in the futures term structure from one representing inverse carrying charges to one with positive carrying charges. According to Weymar, Working’s model (which uses the current inventory level) is likely to hold for agricultural commodities with lumped harvest so that the inventory level declines until the next harvest.

Fama and French (1987) test the theory of storage using data for 21 commodities including metals, and agricultural and wood products. The test involves regressing the futures and spot basis against the nominal interest rates and seasonal dummies. Seasonal dummies are used to capture variability in the marginal convenience yield. The proxy is based on the reasoning that seasonal changes in the inventory produce seasonality in the convenience yield and consequently in the commodity basis. They find a one-for-one variation in the basis with the nominal interest. Moreover, the seasonal effects are found in the basis for the seasonally produced agricultural commodities. Both of these results lend support to the theory of storage.

Using the direct test of the cost of carry model, Heaney (1998) finds evidence of a single cointegrating vector consisting of a constant term, the 3-month London Metal Exchange (LME) lead futures prices, lead spot prices, interest rates and the stocks held by LME-

\textsuperscript{17} The risk arises principally from an unexpected decrease in the price at which stocks can be sold. The risk aversion factor is a positive function of the amount of stocks held.
approved warehouses. The stocks are used to proxy the stock level effects, which include the convenience yield and risk premium. The stock level parameter is negative, which is consistent with the cost of carry model specified by Working (1949) and Brennan (1958).

Sørensen (2002) investigates the behaviour of commodity prices using the seasonal component in the term structure of futures prices. The underlying model is estimated from weekly observations of the futures prices of corn, soybean and wheat on the Chicago Board of Trade. Consistent with the Kaldor–Working hypothesis, the study finds a negative relationship between the convenience yield and the level of inventory normalised by US production.

Heinkel et al. (1990) argue that convenience yield may also arise from the option of selling the commodity before the contract maturity if economically profitable. They develop a three-date model in which uncertain demand for commodity is realised. Spot commodities are assumed to trade at time 0 whereas futures contracts trade at times 0 and 2. The option value of the convenience yield is obtained from the increase in spot price (without a concurrent increase in the futures price) due to unexpected high demand at time 1. Consistent with earlier findings, the study finds that the level of inventory is inversely related to the option value of the convenience yield.

In a subsequent study, Milonas and Thomadakis (1997a, 1997b) formulate a three-date model to determine the value of convenience yield for storable commodities. The model employed in the study uses two dates (0 and 1) to represent the beginning and end of the production cycle. An intermediate date \( z \) is also used. Milonas and Thomadakis posit that the convenience yield obtained by holding the inventory has a payoff structure equivalent to that of a call option. The holder of the inventory has an option either to keep the inventory over
the interval [0,1] or liquidate it if the spot price exceeds the futures price at time $z$. The payoff from this option is linked to the probability of a stockout occurring at time $z$, which is adequate to push the spot price above the contemporaneous futures price net of the carrying charges. The results show that the call option estimate of the convenience yield is negatively related to the inventory level.

Heaney (2002a) approximates the option value of the convenience yield by adopting the methodology advanced by Longstaff (1995). According to Heaney, the convenience yield on a commodity that is held from $t$ to $n$ is equal to the value of an option to sell the commodity in the intermediate period if the price rises sufficiently as to obtain profit after buying back the commodity at time $n$. To obtain the maximum value of this strategy, the trader is assumed to exhibit perfect forecasting and market timing. The payoff structure of the option thus resembles that of a lookback put option. The value of the option is expressed as a non-linear function of the volatility of cash and futures prices, and the time to maturity of the futures contracts. The study uses data for cash and 3-month futures prices of copper, lead and zinc contracts traded on the LME. The observed futures prices are then compared with the theoretical futures prices estimated using the cost of the carry model. Heaney reports that the cost of carry model that includes the estimated convenience yield measures reduces significantly the difference between observed and theoretical futures price.

Using the approach outlined in Heaney (2002a), Hochradl and Rammerstorfe (2011) estimate the implied convenience yield in European natural gas hub trading. They drop the assumption that investors can sell the commodity at a maximum price and instead assume that the commodity is sold at the average price. In this case, the value of the trading strategy is equivalent to that of an Asian put option. The study uses cash and futures data for the natural gas traded on three major trading hubs. The empirical analysis involves a comparison
between the convenience yield estimated using the lookback put option method employed by Heaney (2002a) and the Asian put option. The study finds that the convenience yield based on lookback put options diminishes at an increasing rate, whereas with the Asian option-based approach the benefits fluctuate around a constant mean. Overall the analysis indicates little difference between the two methods.

4.4 The Mechanism of Arbitrage

The most frequently cited economic function of a well-functioning futures market includes price discovery and risk transfer (Silber, 1985). The price discovery role refers to the incorporation of new information into market prices (Garbade and Silber, 1983). The risk transfer function, on the other hand, refers to the use of futures contracts by hedgers to transfer risk to other participants in the market (Working, 1948). For these roles to be performed efficiently, the futures market itself must work efficiently. Under the theory of storage, the efficiency of the futures market is explained in terms of the cost of carry framework (Brenner and Kroner, 1995). Tests of futures market efficiency rest on the assumption that there exists an equilibrium relationship between futures and spot prices, which differ by an amount that is equal to the net cost of holding the commodity until some future date. Although prices may occasionally deviate from equilibrium, these deviations are corrected in the long run by arbitrage trading.

Arbitrage trading ensures that futures and the spot markets are closely linked. In the cost of carry model, if the futures price deviates from the cost of carry price, it would result in opportunities to make riskless profit. The resulting profit taking would cause futures and/or spot prices to adjust until no further arbitrage opportunities are available. Equation (4.1) essentially describes a no-arbitrage relationship between the futures and spot market for the commodity. In the absence of arbitrage, the price of the futures contract should be equal to
the spot price and the net cost of carry. The no-arbitrage condition can also be expressed as equality between the commodity basis (i.e. the difference between futures and spot prices), and the net cost of storage. Taking natural logarithms of Equation (4.1) yields:

$$\ln F_{t,t+n} = \ln S_t + [w + rf - cy](n)$$ (4.2)

Equation (4.2) can be rearranged as:

$$\ln F_{t,t+n} - \ln S_t = [w + rf - cy](n)$$ (4.3)

Under the no-arbitrage condition, a market participant is indifferent to (i) holding inventory until the maturity date \(t + n\) and earning an expected rate of the return that is equal to the sum of the expected future spot price and the convenience yield less the cost of carry; and (ii) selling the inventory at time \(t\) and investing the proceeds at the risk-free rate. If the market is efficient, an inequality in the expected returns from the two outcomes is be corrected by arbitrage.

The cost of carry model represented by Equations (4.1), (4.2) and (4.3) can be described as a special case of temporal spread between the futures and spot prices. The model can also be applied to the spread between the prices of two futures contracts with different maturities, quoted at the same time. For instance, if we have two futures contracts at time \(t\) maturing in \(t + n\) and \(j\) months such that \(n > j\) then the spread between these contracts would be given by:

$$\ln F_{t,t+n} - \ln F_{t,t+j} = [w + rf - cy](n - j)$$ (4.4)

As indicated by Working (1949), the price of a commodity at different periods can be related by the net cost of storage (the net cost of holding the commodity over the time interval between the two contracts). Thus, the cost of carry model outlines a long-run equilibrium relationship between securities with different maturities. The strength of this relationship, however, relies on arbitrage trading to correct any deviations from equilibrium.
Moosa and Al-Loughani (1995) assert that any mispricing observed in the cost of carry relationship should cause the excess demand for futures contracts by the arbitragers to increase. The demand for futures contracts should continue to rise until the mispricing is eliminated. In their study of the futures market, Garbade and Silber (1983) describe the excess demand for futures motivated by arbitrage as the ‘supply of arbitrage services’. According to them, the supply of arbitrage services should be infinitely elastic in the face of any violation of the cost of carry relationship. This view is based on several simplifying assumptions. For instance, it is assumed that no transaction costs are faced by arbitragers who attempt to benefit from the mispricing of futures contracts. In the real world, traders incur substantial transaction costs that could outweigh the initially perceived expected profit. It is also assumed that the commodity traded in the futures and spot market is similar in terms of grade, delivery and location. In the case of crude oil, differences may exist with respect to these characteristics, which may limit cross-market substitutability (Pindyck, 2001). Further, the notion of infinitely elastic supply of arbitrage services assumes unrestricted access to arbitrage capital, which is unlikely to hold in the real world (Shleifer and Vishny, 1997). Thus, the presence of market frictions may suggest that the supply of arbitrage services is less than infinitely elastic. In this case, arbitrage would not be enough to remove mispricing, which could have important implications for the futures market.

4.5 Cointegration Theory and Futures Market Efficiency

Empirical tests of futures market efficiency generally involve determining if a long-run relationship exists between the components of the cost of carry model. In testing the efficiency hypothesis, an issue arises regarding the stationarity of variables underlying the model. Stationarity is an important property, as non-stationary variables tend to exhibit trends

---

18 Mispricing refers to the inequality between the observed and theoretical futures prices.
and have variances that change over time. Using standard statistical procedures to test the hypothesis of an equilibrium relationship under these circumstances would lead to invalid inferences. Cointegration theory overcomes this limitation by allowing for the estimation of linear combinations among non-stationary variables and examining the co-movement between these variables in the long and short run.

The concept of cointegration dates back to the study of Granger and Newbold (1974), who show that for non-stationary variables, standard estimation techniques such as linear regression may suggest a statistically significant relationship when one does not actually exist. Granger and Newbold find that even if the underlying variables are unrelated, the null hypothesis of a zero coefficient is rejected more frequently than what is inferred by theory. They conclude that significant relationships between non-stationary variables, as judged by standard econometric models, could be spurious.

To elaborate on this idea, Granger (1981) introduced the concept of order of integration of a variable. According to Granger, if variable \( z_t \) becomes stationary after it is differenced \( d \) times, then \( z_t \) is integrated of the order \( d \). Economic and financial time series are typically integrated of the order 1, which means that they are unit root processes. A popular approach adopted by econometricians to avoid the problems associated with modelling unit root processes is to transform the integrated series into stationary ones by differencing the series. However, this approach is criticised on the grounds that long-run relationships between variables as posited by economic theory are often based on the levels rather than differences of variables. Using differences would only allow examination of short-term relationships.

According to Engle and Granger (1987), two or more variables can be co-integrated if (i) the variables are integrated of the same order, and (ii) a linear combination of the variables is
stationary; that is $I(0)$. This can be illustrated by using a simple model of oil prices with a single regressor. This model can be expressed as:

$$S_t = f(X_t)$$  \hspace{1cm} (4.5)$$

where $S_t$ is the price and $X_t$ is the explanatory variable. If $S$ and $X$ are non-stationary and are integrated of the same order (that is, $S_t \sim I(1)$ and $X_t \sim I(1)$) then a cointegrating relationship between them can be expressed as:

$$S_t = \alpha + \beta X_t + \epsilon_t$$  \hspace{1cm} (4.6)$$

If the error term—which is a linear combination of these two variables—is stationary, then these variables are co-integrated. Therefore, to determine if a long-run relationship exists between the variables included in the model, they must be examined for stationarity.

Cointegration theory has been used extensively to determine the efficiency of the futures market. Efficiency is generally examined by testing the hypothesis that contemporaneous futures price provides an unbiased forecast of the subsequently observed spot price (i.e. the futures price does not constantly over- or under-predict the spot price). Failure to find cointegration between spot price and the lagged futures price is an indication of an inefficient market. This test of market efficiency involves running the following regression:

$$S_{t+n} = \alpha - \beta F_{t,t+n} + \epsilon_t$$  \hspace{1cm} (4.7)$$

and testing the coefficient restriction $(\alpha, \beta) = (0,1)$ and the hypothesis that $\epsilon_t$ is serially uncorrelated. Assuming risk neutrality, if all relevant information is incorporated into the futures prices, then, on average, the futures price at time $t$ must be equal to the realised spot price at time $t + n$ and the residuals should not contain any additional information relevant for predicting the future spot price. Several others tests of unbiasedness based on Equation (4.7) are employed in the literature. One such test involves differencing the variables in Equation (4.7) and setting $n = 1$, which provides the following regression model:
\[ \Delta S_{t+1} = \alpha - \beta \Delta F_{t,t+1} + \xi_t \]  

(4.8)

Once again, unbiasedness is examined by testing if \((\alpha, \beta) = (0,1)\) and the residuals are uncorrelated. Another popular test of unbiasedness is conducted by subtracting \(S_t\) from both sides in Equation (4.7) and setting \(n = 1\). The resulting model is:

\[ S_{t+1} - S_t = \alpha - \beta F_{t,t+1} - S_t + \varepsilon_t \]  

(4.9)

These three models have been used extensively in the literature with mixed results. For instance, Kellard et al. (1999) test unbiasedness and efficiency across several different commodity and financial futures markets using cointegration. They find cointegration between futures and spot prices, and a slope coefficient that is close to unity, implying a long-run relationship between the two price series. Haigh (2000) investigates the long-run relationship between freight cash and futures prices. The study finds significant evidence for cointegration between the freight futures and cash prices. Similarly, Yang et al. (2001) test the price discovery function of futures markets for selected storable and non-storable commodities, obtaining results that confirm cointegration for six out of eight storable commodities and for all non-storable commodities. Based on these findings, they conclude that the unbiased hypothesis cannot be rejected.

In contrast, some studies find no cointegration for some commodity markets. For instance, Chowdhury (1991) tests market efficiency for four non-ferrous metals traded on the LME. The results show that, with the exception of copper, the null hypothesis of no cointegration between futures and spot prices cannot be rejected. In the case of copper, the results indicate mixed evidence of cointegration. Chowdhury concludes that futures prices are biased predictors of subsequent spot prices, implying market inefficiency.

Crowder and Hamed (1993) test market efficiency using the no-arbitrage profit definition of efficiency. According to them, in an efficient futures market the expected return on a hedged
position in any market should be equal to the risk-free rate. The hypothesis is tested by examining cointegration between the futures price, spot price and the US Treasury bill rate. Test of stationarity performed in the study indicate that each series contains a unit root and is integrated of the same order. The study finds that the null hypothesis of zero cointegrating vectors is rejected but the null hypothesis of no more than one cointegrating vector cannot be rejected. The estimated cointegrating vector shows that once the spot price coefficient is set to 1, the futures price coefficient is not significantly different from –1 as predicted by the model. However, the parameter for the risk-free rate is not significantly different from 0. According to Crowder and Hamed, this result indicates that arbitrage fails to correct deviations from equilibrium or the possibility of an omitted variable from the arbitrage relation. They also relate the failure of the arbitrage equilibrium hypothesis to the possibility of ex-ante returns on oil futures speculation, which should be incorporated into the expectations of speculators.

Fortenbery and Zapata (1997) examine cointegration between the cash–futures prices of cheddar cheese. The study uses data from June 1993 to July 1995 to examine the cash–futures relationship. The empirical test of cointegration fails to reject the null hypothesis of no cointegration between cash and futures prices for cheddar cheese. The authors conclude that price changes in one market have little influence on prices in the other market. Sabuhoro and Larue (1997) test the efficiency of the cocoa and coffee futures markets using daily spot and futures prices obtained from the Coffee, Sugar, and Cocoa Exchange. Using the EG test, they find no evidence of cointegration between the futures and spot prices.

19 The literature provides mixed evidence on whether interest rates have stochastic trends. For instance, Bradley and Lumpkin (1992), Campbell and Shiller (1987), Engle and Granger (1987), Hall et al. (1992) and Shea (1992), among others, provide evidence favouring a stochastic trend in interest rates. Alternatively, Fama and Bliss (1987), Sanders and Unal (1988), Chan et al. (1992) and many others report contrary evidence, suggesting that interest rates are stationary processes.
Brenner and Kroner (1995) provide two explanations for the mixed results regarding cash and futures market cointegration. They argue that failure to find cointegration between the cash and futures prices of commodities may be attributed to two problems. The first is that as the time to maturity of the futures contract decreases, the residual converges on 0, which implies that the variance of residuals changes over time (the residuals are not covariance stationary). Hence, a cointegrating relationship cannot exist. The second problem pertains to misspecification of the model. Brenner and Kroner point out that the cointegration model must account for the time series properties of the net cost of carry. According to them, a test of efficiency based on unbiasedness in commodity markets would result in rejection of the null hypothesis if the net cost of carry contains stochastic trends. They suggest that a no-arbitrage assumption would be more reasonable for testing efficiency. Similarly, Dwyer and Wallace (1992) maintain that the explanation of market efficiency should rely on the condition of no-arbitrage profit.

Moosa and Al-Loughani (1995) extend the no-arbitrage test of efficiency for the crude oil futures market by including the physical cost of carrying the commodity in addition to the risk-free rate of return. The authors posit that expected return from a long position in the spot market and a short position in the futures market should be equal to the sum of the storage cost and the risk-free rate. A deviation from this relationship should be corrected by arbitrage. To test this hypothesis, a cointegration regression is estimated between observed and theoretical futures prices, with the theoretical futures price being derived from the cost of carry model. Further, the effectiveness of arbitrage is examined by restricting the intercept and slope coefficients to 0 and 1. Contrary to Crowder and Hamed (1993), the study finds a cointegrating relation at 5% significance. Restriction on the intercept is not rejected whereas restriction on the slope coefficient is rejected. According to Moosa and Al-Loughani, the results suggest that although arbitrage has a significant influence on the futures price, it is
insufficient to ensure equilibrium in the cost of carry of relation. This assertion implies that the net cost of storage does not fully explain commodity basis.

In comparison, the study of Gulen (1998) analyses a more extended sample and allows for structural breaks while testing the efficiency of the futures market. The findings confirm the role of futures prices as unbiased predictors of spot prices, indicating that the futures market works efficiently. Maslyuk and Smyth (2009) examine cointegration between the spot and futures prices of different grades of crude oil (WTI and Brent). The study reports significant evidence for rejecting the null hypothesis of no cointegration between spot and futures prices of the same, as well as different, grades of crude oil. They link the cointegration between the futures and spot prices to the effectiveness of arbitrage.

Alizadeh and Nomikos (2004) explore the arbitrage opportunities arising from the price discrepancies between different grades of crude oil. They argue that the futures price for a particular grade of crude oil (such as WTI) should be equal to the spot price of another grade of crude oil (e.g., Brent) adjusted for storage, transportation cost and quality discounts. Any deviations from this relationship should be restored through arbitrage. For example, if the cost of buying physical crude oil and carrying it to maturity is less than the futures price, then arbitragers can profit by delivering imported crude oil against WTI futures, implying that any mispricing should be eliminated by arbitrage. The study finds that fluctuations in the differential between the futures prices of WTI oil and the spot price for other grades of crude oil fail to influence transportation charges, indicating the existence of arbitrage opportunities. Further simulation analysis reveals profitable arbitrage opportunities even after incorporating transaction costs and shipment charges.
4.6 Summary and Conclusion

Efficiency of the futures market is typically assessed in terms of cointegration between futures price and the subsequently observed spot price. Under the assumption of risk neutrality, futures prices should on average provide an unbiased forecast of the spot price. Tests of futures market efficiency based on this assumption yield mixed evidence. Brenner and Kroner (1995) attribute the failure of cash and futures market cointegration to model misspecification. They argue that cointegration between futures and spot prices—and hence efficiency of the futures market—depends on the time series properties of net storage cost, which includes not only interest rate and physical storage cost but also the convenience yield. If the market is efficient, then the difference between spot and futures price should be equal to the net storage cost, which implies that market efficiency can be defined in terms of the no-arbitrage condition.

Tests of cointegration based on the cost of carry model show that arbitrage alone is insufficient to explain the difference between spot and futures prices. Several commentators associate this finding with the presence of the risk premium required by speculators for assuming the risk arising from spot price volatility (e.g., Fama and French, 1988; Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006; Hamilton and Wu, 2014). In this case, the price of the futures contract should be explained in terms of arbitrage as well as speculation.
CHAPTER 5
THE EFFECTIVENESS OF ARBITRAGE AND SPECULATION IN EXPLAINING
THE BEHAVIOUR OF CRUDE OIL FUTURES PRICE

5.1 Introduction

In the past few years, a growing body of research has revealed that participation of financial traders in the futures market has resulted in significant changes in the behaviour of crude oil prices. Motivated by this assertion, we investigate the role of arbitrage and speculation in determining the price of crude oil futures. For this purpose, we employ the framework proposed by Moosa and Al-Loughani (1995). In this chapter, we extend their study by incorporating the convenience yield into the cost of carry price of futures. The empirical model used in our analysis is given as follows:

\[ F_{t,t+n} = \beta_0 + \beta_1 F_{t,t+n} + \beta_2 E_t S_{t+n} + \varepsilon_t \]  

(5.1)

One important assumption implicit in this model is that speculators only speculate on the spot commodity. The variables used in the model are in line with established theories of futures price determination. The parameters \( \beta_1 \) and \( \beta_2 \) measure the strength of the influence exerted by arbitrage and speculation on the futures price. The model is tested using the prices of 1, 3, 6 and 9-month WTI crude oil futures contracts traded on the NYMEX. The next section outlines the measurement procedures employed for estimating the variables outlined in Equation (5.1).

---

20 See, for example, Alquist and Kilian (2010), Gilbert (2010a), Mou (2010), Singleton (2011) and Tang and Xiong (2012).
5.2 Variable Measurement

5.2.1 The Cost of Carry

The cost of carry is the costs incurred by holding the commodity for a certain time. It consists of two broad items: the physical storage cost and financial cost. The storage cost of the commodity includes expenditures such as rental on the storage space, handling, insurance and transportation. Measurement of the storage cost is difficult. Heterogeneity in storage location, delivery charges and insurance costs could make the storage cost estimate vary from one commodity holder to another. As no data set exists for storage cost, we follow Moosa and Al-Loughani (1995) in using the inflation rate as a proxy for storage charges. The inflation rate is estimated as the monthly change in the Consumer Price Index (CPI) for OECD economies over the sample period. Data for the CPI are acquired from the OECD database.

The inflation rate is calculated as follows:

\[
\pi_t = n \left[ \frac{P_t}{P_{t-1}} - 1 \right]
\] (5.2)

where \(\pi_t\) is the monthly inflation rate, \(P_t\) is \(t\)-month CPI for OECD countries and \(n\) represents the number of months until the delivery date. The consumer price index is obtained from Thomson Reuters’ Datastream database. The financial cost of carry includes the opportunity cost of the funds invested in the commodity. We follow past studies of the cost of carry model by using the risk-free rate of interest to represent the financial cost. The risk-free rate of interest is proxied by the London Interbank Offered Rate (LIBOR), which is also acquired from the Datastream database. The LIBOR rate is then de-annualised to calculate the opportunity cost over the length of the futures contract, as follows:

\[
rf_t = \frac{L_n}{100} \left( \frac{n}{365} \right)
\] (5.3)

---

21Delivery charges mainly comprise pipeline tariffs.
where \( rf_t \) is the de-annualised risk-free rate obtained by adjusting the annualised LIBOR rate \( (L_n) \) to cover maturity period of the futures contract.

5.2.2 Convenience Yield

The option pricing framework is used to estimate the convenience yield as outlined by Heaney (2002a) who proposes a change in the standard futures market arbitrage model, which is based on the cost of carry misalignments. Heaney suggests that in a standard arbitrage model, an arbitrager takes a long position in the spot commodity and a short position in the futures market whenever the futures price exceeds the cost of carry price. On the other hand, the arbitrager takes a short position in the spot market and a long position in the futures market if the futures price falls below the cost of carry price. In both cases it is assumed that spot and futures positions are held until the contract maturity. However, to formulate an approximation of the convenience yield, this assumption is relaxed so that traders are no longer required to hold their positions until maturity. According to Heaney, in this way the convenience yield would be maximum if the trader exhibits perfect foresight regarding the price path and is able to sell the commodity at the highest price. Once the commodity is sold, the trader will invest the proceeds at the risk-free rate and buy back the commodity at a lower price, when the futures contract matures. A similar approach is followed by Longtsaff (1995), who investigates the value of marketability of a security in an illiquid market. It is assumed that the commodity’s spot price follows the stochastic process:

\[
dS = \mu S dt + \sigma S dZ
\]  

(5.4)

where \( \mu \) is the drift term, \( \sigma \) is the standard deviation of the oil price and \( Z \) represents a Wiener process with mean 0 and variance \( dt \). The arbitrager who obtains the commodity at time \( t \) has the option to hold it until the delivery at time \( t + n \) or liquidate their position at time \( \tau \) such that \( \tau < n \). If the arbitrager sells the commodity at the maximum price \( S_\tau \), the revenue from
the sale of the commodity will be invested at risk-free rate \( r \) until the maturity date \( t + n \).

This maximum price over the period \( t \) to \( t + n \) compounded to futures contract maturity is:

\[
M_{t+n} = \max_{0 \leq \tau \leq t+n} \left( \exp(r(t + n - \tau))S_\tau \right)
\]  

(5.5)

The present value of this payoff is equal to:

\[
V(S_t, t + n) = \exp(-r(t + n - t))E(M_{t+n}) - \exp(-r(t + n - t))E(S_{t+n})
\]  

(5.6)

Longstaff (1995) provides an approximation for the value of this option as follows:

\[
V(S_t, t + n) = S_t \left[ \left( 2 + \frac{\sigma_S^2(t + n - t)}{2} \right) N \left( \frac{\sqrt{\sigma_S^2(t + n - t)}}{2} \right) 
+ \frac{\sigma_S^2(t + n - t)}{2\pi} \exp \left( -\frac{\sigma_S^2(t + n - t)}{2} \right) - 1 \right]
\]  

(5.7)

where \( N(.) \) is cumulative normal distribution and \( \sigma_S^2 \) is the variance of the spot price returns.

The convenience yield as a proportion of spot price would be given as:

\[
\frac{V(S_t, t + n)}{S_t} = \left[ \left( 2 + \frac{\sigma_S^2(t + n - t)}{2} \right) N \left( \frac{\sqrt{\sigma_S^2(t + n - t)}}{2} \right) 
+ \frac{\sigma_S^2(t + n - t)}{2\pi} \exp \left( -\frac{\sigma_S^2(t + n - t)}{2} \right) - 1 \right]
\]  

(5.8)

The continuously compounded convenience yield can be obtained by taking natural logs:

\[
v_{t,t+n}(S_t, t + n) = \ln \left( 1 + \frac{V(S_t, t + n)}{S_t} \right)
= \ln \left[ \left( 2 + \frac{\sigma_S^2(t + n - t)}{2} \right) N \left( \frac{\sqrt{\sigma_S^2(t + n - t)}}{2} \right) 
+ \frac{\sigma_S^2(t + n - t)}{2\pi} \exp \left( -\frac{\sigma_S^2(t + n - t)}{2} \right) \right]
\]  

(5.9)

This trading strategy is also available to the holder of the futures contract. The value of the strategy is given by:
where $\sigma^2_t$ is the variance of the futures price returns. The values of trading strategies can now be used to generate an approximation of the convenience yield on a commodity held from $t$ to $t + n$:

$$c_{yt,n} = v_{t,t+n}(S_{t,t+n}, t + n) - v_{t,t+n}(F_{t,t+n}, t + n)$$ (5.11)

Equations (5.9) and (5.10) provide estimates of the convenience yield based on the variability of the cash and futures price returns, and the time to maturity of the futures contract. According to Equation (5.11), the value of the option to sell the commodity before maturity is a positive function of the variance in the cash price returns relative to the futures price returns. When the components of the cost of carry model have been estimated, the theoretical futures price can be calculated as:

$$\bar{F}_{t,t+n} = S_t(1 + \pi_{t-1})(1 + r_f)(1 - c_{yt,n})$$ (5.12)

Once the theoretical futures price is estimated, the question then arises of how to proxy the expectations variables in Equation (5.1), which represents the expected spot price at the maturity of the futures contract. The expected spot price is estimated using the RE formation mechanism, which involves a situation where the forecast of the oil price is formed using all available information at the time of forecasting. The expected spot price of oil based on the RE mechanism can be written:

---

23Lagged inflation rate is used in Equation (5.11) as it is assumed that the cost of carry, which is determined in advance for the coming holding period, is estimated using the inflation rate available immediately before the date of the futures price observation.
\[ S_{t+n} = E_t S_{t+n} + w_{t+n} \]  \hspace{1cm} (5.13)

or

\[ E_t S_{t+n} = S_{t+n} - w_{t+n} \]  \hspace{1cm} (5.14)

where \( w_{t+n} \) is the expectations error that is realised at time \( t + n \). By substituting Equation (5.14) into Equation (5.1), we obtain:

\[ F_{t,t+n} = \beta_0 + \beta_1 \tilde{F}_{t,t+n} + \beta_2 (S_{t+n} - w_{t+n}) + \epsilon_{t+n} \]  \hspace{1cm} (5.15)

or

\[ F_{t,t+n} = \beta_0 + \beta_1 \tilde{F}_{t,t+n} + \beta_2 S_{t+n} + u_{t+n} \]  \hspace{1cm} (5.16)

where \( u_{t+n} = -\beta_2 w_{t+n} + \epsilon_{t+n} \). If we assume that \( w_{t+n} \sim I(0) \), then \( u_{t+n} \sim I(0) \) if \( \epsilon_{t+n} \sim I(0) \). This suggests that a long-run equilibrium in Equation (5.16) would imply a long-run equilibrium in Equation (5.1).

5.3 Data Description

The data required for estimating the model outlined by Equation (5.12) include spot and futures price of crude oil. We use the prices of crude oil futures contracts traded on the NYMEX. Although several different grades of crude oil are traded on NYMEX, the analysis in this study is based on WTI crude oil.\(^ {24} \) The WTI oil is also referred to as light sweet crude oil. The adoption of WTI for the analysis is primarily motivated by its use as a benchmark for oil prices. WTI is also used as the main benchmark for the pricing of imports of crude oil into the US, which is the world’s largest consumer of oil.\(^ {25} \) Although several different grades of crude oil are produced in the US, WTI has significant importance in global oil and financial markets for it constitutes one of the most actively traded commodity futures contracts. Because of its liquidity and transparency, light sweet crude oil futures also provide excellent


means of managing energy risk.

A single crude oil futures contract traded on NYMEX is specified as 1000 barrels of WTI for delivery at Cushing, Oklahoma. The prices of contracts are denominated in US dollars and cents per barrel. The contracts mature 3 days prior to the 25th calendar day of the month preceding the delivery month. If the 25th calendar day is a non-business day, trading ceases on the third business day prior to the business day preceding the 25th calendar day of the month.

The sample, which includes daily spot and futures prices of WTI crude oil covering the period January 1988 to April 2013, is obtained from the Bloomberg database. Both spot and futures prices are measured as at the close of trading on the exchange. As the majority of the fundamental variables are only available at monthly frequency, a monthly sample for futures and spot prices is constructed. Following Crowder and Hamed (1993), the settlement price on the day following the last trading day of the given contract is selected as a proxy for the monthly futures price. Following this approach, a $t$-month futures contract trading on the day following the last trading day of every month would be approximately $t$ months from the maturity of the contract. For example, to obtain the price of a 1-month futures contract with delivery in June, the closing price on the day following the last trading day for the month of April is selected. The sample comprises futures contracts with maturities of 1, 3, 6 and 9 months.

The spot price is observed on the same day as the futures price. One serious problem associated with the cash price is that it may include discounts or premiums due to long-term

---

26 A 3-day window is provided to make allowances for the physical delivery of oil in the following month (Deaves and Krinsky, 1992a).

27 In Crowder and Hamed (1993), the futures price for a 1-month contract is given by the closing price 30 days prior to the last trading day of the contract.
relationships between buyers and sellers (Pindyck, 1994). Further, the oil traded in the spot market may differ in terms of grade and delivery location, which in turn may limit the substitutability of the commodity between the spot and futures market. Hence, a second proxy is selected: the closing price on the last trading day of the nearest-to-maturity futures contract (also known as the front-month contract) is chosen as the proxy for the contract. The 1-month contract is chosen as a proxy for the contract that is nearest to maturity.

5.3.1 Sample Descriptive Statistics

This section provides a discussion of the statistical properties of the time series. The oil market has experienced a number of shocks over the study period, which in turn may bias sample statistics. Therefore, in describing the statistical behaviour of the variables, the sample is divided into three sub-periods: (i) January 1988 to February 1999; (ii) March 1999 to the end of 2003; and (iii) from the start of 2004 to the end of the sample period.

Table 5.1: Descriptive Statistics at Levels from January 1988 to February 1999

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_{t,t+1} )</td>
<td>19.13</td>
<td>35.43</td>
<td>10.81</td>
<td>3.59</td>
<td>0.97</td>
<td>6.51</td>
<td>134</td>
</tr>
<tr>
<td>( F_{t,t+3} )</td>
<td>18.89</td>
<td>32.15</td>
<td>11.41</td>
<td>3.08</td>
<td>0.75</td>
<td>5.74</td>
<td>134</td>
</tr>
<tr>
<td>( F_{t,t+6} )</td>
<td>18.69</td>
<td>29.04</td>
<td>12.20</td>
<td>2.60</td>
<td>0.48</td>
<td>5.00</td>
<td>134</td>
</tr>
<tr>
<td>( F_{t,t+9} )</td>
<td>18.56</td>
<td>27.05</td>
<td>12.77</td>
<td>2.30</td>
<td>0.31</td>
<td>4.57</td>
<td>134</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+1} )</td>
<td>19.47</td>
<td>35.25</td>
<td>10.87</td>
<td>3.78</td>
<td>0.79</td>
<td>5.34</td>
<td>134</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+3} )</td>
<td>19.81</td>
<td>36.34</td>
<td>10.96</td>
<td>3.87</td>
<td>0.81</td>
<td>5.51</td>
<td>134</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+6} )</td>
<td>20.35</td>
<td>38.08</td>
<td>11.08</td>
<td>4.04</td>
<td>0.86</td>
<td>5.78</td>
<td>134</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+9} )</td>
<td>20.91</td>
<td>39.90</td>
<td>11.24</td>
<td>4.23</td>
<td>0.91</td>
<td>6.03</td>
<td>134</td>
</tr>
<tr>
<td>( S_t )</td>
<td>19.29</td>
<td>34.71</td>
<td>10.81</td>
<td>3.72</td>
<td>0.77</td>
<td>5.26</td>
<td>134</td>
</tr>
</tbody>
</table>

Note: Normal distribution kurtosis value is 3

This behaviour of oil prices supports the Samuelson hypothesis, which asserts that the volatility of the forward prices, ceteris paribus, declines as the maturity period increases. The decline in volatility is also indicative of the manner in which oil market news or demand and
supply shocks have a greater impact on short-term than on long-term prices. This is because for long-term contracts, production is expected to adjust before the expiry of the contract. Both futures and spot prices are positively skewed, which points to the presence of extreme values at the positive end of the distribution. This implies that demand and supply shocks tend to push prices downwards, hence creating a downside risk for hedgers.

Table 5.1 provides the descriptive statistics for the observed crude oil futures prices, cost of carry price and spot price. In the first sub-period, futures prices averaged around $18 per barrel with a standard deviation of $2–$4 per barrel. In comparison to futures price, the cost of carry price and the spot price have higher means and standard deviations. Both the mean and standard deviation decline as the length of the futures contract increases. The kurtosis value in this period is 6.51 for the 1-month futures prices, which suggests a fat-tailed distribution (i.e. a higher probability of occurrence or risk of extreme prices). Similar to the mean, the kurtosis values decline with the length of the futures contract.

In the second sub-period (see Table 5.2), the volatility of futures and spot prices increases in comparison to the first sub-period. The average price level is higher as well. The price distribution is negatively skewed, which indicates a higher frequency of prices above the average price level. Kurtosis, on the other hand, is close to that of a normal distribution. This implies a slightly reduced probability of extreme shocks to prices in the underlying period. As the maturity period of the futures contract increases, the standard deviation decreases. On the other hand, both kurtosis and skewness rise as the contract length increases. These statistics suggest that although there is an increase in the variability of prices, the probability or risk of sudden price increases declines as contract maturity increases.
The oil market trends upwards in the third sub-period, with the average price reaching more than $70 per barrel (see Table 5.3). The standard deviation is significantly higher compared to the first two periods. The difference between minimum and maximum prices is also quite high. Kurtosis is closer to the normal distribution value whereas skewness is positive, which indicates a higher probability of prices falling below the average and therefore greater downside risk.

### Table 5.2: Descriptive Statistics at Levels from March 1999 to December 2003

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{t+1}$</td>
<td>26.83</td>
<td>35.58</td>
<td>15.50</td>
<td>4.63</td>
<td>-0.52</td>
<td>2.88</td>
<td>58</td>
</tr>
<tr>
<td>$F_{t+3}$</td>
<td>26.11</td>
<td>32.99</td>
<td>15.50</td>
<td>4.12</td>
<td>-0.63</td>
<td>2.99</td>
<td>58</td>
</tr>
<tr>
<td>$F_{t+6}$</td>
<td>25.05</td>
<td>31.13</td>
<td>15.30</td>
<td>3.64</td>
<td>-0.75</td>
<td>3.11</td>
<td>58</td>
</tr>
<tr>
<td>$F_{t+9}$</td>
<td>24.21</td>
<td>29.69</td>
<td>15.17</td>
<td>3.32</td>
<td>-0.81</td>
<td>3.16</td>
<td>58</td>
</tr>
<tr>
<td>$\bar{F}_{t+1}$</td>
<td>27.41</td>
<td>37.59</td>
<td>15.62</td>
<td>5.14</td>
<td>-0.47</td>
<td>2.78</td>
<td>58</td>
</tr>
<tr>
<td>$\bar{F}_{t+3}$</td>
<td>27.68</td>
<td>38.37</td>
<td>15.87</td>
<td>5.20</td>
<td>-0.45</td>
<td>2.81</td>
<td>58</td>
</tr>
<tr>
<td>$\bar{F}_{t+6}$</td>
<td>28.09</td>
<td>39.51</td>
<td>16.24</td>
<td>5.29</td>
<td>-0.41</td>
<td>2.85</td>
<td>58</td>
</tr>
<tr>
<td>$\bar{F}_{t+9}$</td>
<td>28.53</td>
<td>40.70</td>
<td>16.65</td>
<td>5.38</td>
<td>-0.37</td>
<td>2.90</td>
<td>58</td>
</tr>
<tr>
<td>$S_t$</td>
<td>27.26</td>
<td>37.20</td>
<td>15.50</td>
<td>5.10</td>
<td>-0.48</td>
<td>2.76</td>
<td>58</td>
</tr>
</tbody>
</table>

Note: Normal distribution kurtosis value is 3

### Table 5.3: Descriptive Statistics at Levels from January 2004 to April 2013

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{t+1}$</td>
<td>74.92</td>
<td>136.74</td>
<td>34.35</td>
<td>22.40</td>
<td>0.30</td>
<td>2.74</td>
<td>112</td>
</tr>
<tr>
<td>$F_{t+3}$</td>
<td>75.84</td>
<td>137.26</td>
<td>32.88</td>
<td>22.26</td>
<td>0.25</td>
<td>2.81</td>
<td>112</td>
</tr>
<tr>
<td>$F_{t+6}$</td>
<td>76.50</td>
<td>137.10</td>
<td>31.41</td>
<td>22.31</td>
<td>0.16</td>
<td>2.87</td>
<td>112</td>
</tr>
<tr>
<td>$F_{t+9}$</td>
<td>76.75</td>
<td>136.60</td>
<td>30.48</td>
<td>22.35</td>
<td>0.09</td>
<td>2.93</td>
<td>112</td>
</tr>
<tr>
<td>$\bar{F}_{t+1}$</td>
<td>74.91</td>
<td>135.73</td>
<td>33.68</td>
<td>22.84</td>
<td>0.28</td>
<td>2.63</td>
<td>112</td>
</tr>
<tr>
<td>$\bar{F}_{t+3}$</td>
<td>75.45</td>
<td>137.98</td>
<td>33.29</td>
<td>23.10</td>
<td>0.30</td>
<td>2.69</td>
<td>112</td>
</tr>
<tr>
<td>$\bar{F}_{t+6}$</td>
<td>76.31</td>
<td>141.54</td>
<td>32.69</td>
<td>23.52</td>
<td>0.34</td>
<td>2.81</td>
<td>112</td>
</tr>
<tr>
<td>$\bar{F}_{t+9}$</td>
<td>77.18</td>
<td>145.37</td>
<td>32.09</td>
<td>23.98</td>
<td>0.37</td>
<td>2.92</td>
<td>112</td>
</tr>
<tr>
<td>$S_t$</td>
<td>74.62</td>
<td>134.62</td>
<td>33.87</td>
<td>22.71</td>
<td>0.27</td>
<td>2.59</td>
<td>112</td>
</tr>
</tbody>
</table>

Note: Normal distribution kurtosis value is 3
5.3.2 Graphical Analysis

The examination of sample statistics indicates considerable variation in the mean and standard deviation for several variables over the sample period, which may suggest that the variables included in the analysis are non-stationary (i.e. they have unit roots). Prior to undertaking any econometric analysis, it is necessary to ascertain whether the underlying time series are unit root or non-stationary processes. A regression analysis based on non-stationary data may produce spurious results that cannot be used draw to valid inferences.

The approach taken in this study involves first conducting a visual inspection of the variables to identify any rising or declining levels over the study period. Changing levels over time may indicate the possibility of a trend or a drift, which is indistinguishable by visual inspection only. Therefore, a unit root test is also undertaken with different specifications involving trend component and the intercept, and with intercept but no trend.

Figure 5.1: West Texas Intermediate Spot and Futures Prices
For the purpose of visual analysis of the future’s prices, the sample is divided into the same three sub-periods used in the discussion of the sample statistics. Figure 5.1 shows the monthly spot and futures prices over the sample period. It can be seen that both price series track each other quite closely. In the first sub-period the oil prices show stable behaviour, fluctuating within a narrow range.

The standard deviation of prices in this period lies between $2 and $4 (see Table 5.1 for descriptive statistics). However, some abrupt fluctuations do result from major oil market events in this period. For example, the Gulf War in 1990 caused a sharp spike in oil prices amid declining production. A few years later, the Asian financial crisis, combined with the Ruble crisis, resulted in a significant decline in oil prices in 1998 as the Organisation of the Petroleum Exporting Countries (OPEC) increased its production (Williams, 2013). As oil consumption declined, the prices fell to around $12 per barrel by the end of 1998. In this period the oil price averaged around $18 per barrel. However, the skewness of oil prices in this period is high, indicating a fat-tailed distribution and suggesting a higher probability of extreme prices.

The second sub-period can be characterised as a transition period in which prices became less stable with no long-lasting trends. This is also evident in the higher mean and standard deviation of prices in comparison to the first sub-period. At the start of this period, oil prices began to rise as OPEC changed its pricing strategy, but the terrorist attacks on 11 September 2001 caused both futures and spot prices to decline. The downwards trend in prices was later reversed due to political unrest in Venezuela in the first half of 2002. The invasion of Iraq in the following year saw a sharp increase in price as OPEC’s excess production capacity

---

28 The 3-month futures prices are used to represent futures prices in the graph. The behaviour of futures prices with other maturities is almost identical.
dropped to two million barrels per day. The increase in prices was further intensified by speculation on the destruction of Iraq’s oil reserves in the days leading up to the invasion.

The third sub-period witnessed an increase in oil price volatility around a positive trend. During the first few years of this period, excess production capacity remained below one million barrels per day, which resulted in a sharp increase in oil prices in 2004 and 2005 (Williams, 2013). The hurricanes of 2005 also contributed to higher prices during this period. Oil prices continued to soar in 2006 and 2007 as world oil production failed to adjust to the increase in demand from Asian economies, and the popularity of commodities as an alternative investment also increased (Hamilton, 2009). According to an estimate by the Commodity Trading Futures Commission (CFTC), the institutional holdings of commodities increased from $15 billion in 2003 to more than $200 billion in 2008 (CFTC, 2008). As this period approximately coincides with the boom in oil prices, this observation has led to a debate on whether the influx of institutional investments into the commodity markets impacted the oil prices (for example, Masters, 2008; Singleton, 2011).

Growth in the oil market continued in 2008 as prices rose to a record high of $145.29 at the close of trading on NYMEX on 3 July 2008. However, with the onset of the global financial crisis, oil prices dropped below $40 per barrel in December 2008. In the following year, the oil market recovered amid increase in demand from Asian economies. The positive trend in oil prices continued in 2010 and 2011 as global oil supplies were constrained by unrest in the Middle East (Libyan war, Arab Spring), coupled with political tensions in the region (sanctions on Iran). Oil prices remained high in 2012 and 2013 amid concerns regarding possible supply disruptions caused by unrest in the Middle East.
Overall crude oil prices exhibited greater variability in the third sub-period than in the first and second sub-periods. To ascertain this, the trend and cyclical component of the price series are obtained using a Hodrick–Prescott filter. The cyclical component displayed in Figure 5.2, shows an increase in variation in prices in the third sub-period.

A clear positive trend is also evident in this period, which is present even when crude oil price is expressed in natural logs, as shown in Figure 5.3. However, it must be emphasised here that the majority of the oil shocks over the sample period occurred in the third sub-period as well. These shocks (resulting from armed conflicts, political tensions and supply disruptions) have contributed to the instability of the oil market, prompting an increase in the need for hedging.
The graphical analysis of WTI crude oil price behaviour over the sample period shows that oil price is characterised by several upwards and downwards trends and volatility that varies over the sample period. In the first half of the sample period, prices remain generally stable, fluctuating within a narrow range with no noticeable trend. The latter half, however, is characterised by increased volatility and upwards drifts that vary in magnitude. Visual inspection supports the findings obtained from the analysis of sample statistics. An econometric analysis based on these data may produce spurious results and impede the effectiveness of oil price forecasting.
The graphical analysis of commodity basis is also conducted over the sample period. Figure 5.4 plots the cost of carry measured as $\bar{F}_t - S_t$ against basis $F_t - S_t$ for the 3-month futures contract. The figure shows that the cost of carry varies considerably from the spot–futures basis, suggesting that the former does not account accurately for the difference between the spot and futures price over the sample period. The difference between the basis of and the cost of carry appears to be increasing with variability of the basis. Overall, Figure 5.3 demonstrates the interaction between arbitrage and speculation over the sample period.

5.4 Methodology

5.4.1 Unit Root Tests

Following the graphical analysis, the study proceeds with unit root tests to identify the non-stationary variables and determine their order of integration. The Augmented Dickey–Fuller (ADF) test is used to test for stationarity of the variables. The ADF is based on the following equation:
\[
\Delta y_t = u + \beta t + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \varepsilon_t
\]  
(5.17)

where \( y_t \) is the time series under investigation, \( t \) is a time trend, \( k \) is the number of lags and \( \varepsilon_t \) is the error term with mean of 0 and constant variance. Stationarity is determined by testing the null hypothesis of \( \alpha = 0 \) against an alternative, \( \alpha < 0 \). Rejection of the null hypothesis implies a stationary series.

One critical limitation of the ADF test is that failure to account for structural breaks in the data may lead to false rejection of the null hypothesis (Perron, 1989).\(^{29}\) For this reason, the Phillips–Perron (PP) test is employed in addition to the ADF test. The PP test includes a dummy variable to account for a single known structural break. In this test, the assumption of independent errors and constant variance are relaxed. Like the ADF test, the PP test involves testing a null hypothesis of a unit root against the alternative of a stationary series.

To determine the optimal lag length, the Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) are used. Two specifications for both the ADF and PP tests are used. The first specification includes a constant term, whereas the second specification includes a constant as well as a time trend, which is included to determine if any of the series is trend stationary.

5.4.2 Testing for Cointegration

The aim of this study is to investigate the role of arbitrage and speculation in the determination of futures prices. As shown previously, the intensity of arbitrage activity is shown to be positively related to the magnitude of the deviation of the observed futures price from the theoretical futures price (Section 3.2). On the other hand, the influence of

\(^{29}\) A structural break involves an unexpected change in the data series.
speculation is shown to be positively related to the difference between the observed futures price and the expected spot price at the maturity of the futures contract (Section 3.3). The impact of both speculation and arbitrage on the futures price is expressed in the form of a testable model as represented by Equation (5.1). However, to account for trend in the variables (as observed in the visual inspection), a time trend variable is included in Equation (5.1). Hence:

\[ F_t = \beta_0 + \beta_1 \bar{F}_t + \beta_2 S_{t+n} + \beta_3 t + u_{t+n} \]  

Equation (5.18) describes a positive relationship between the observed futures price, the cost of carry price and expected spot price. The variables in Equation (5.18) are measured in natural logarithms; hence the coefficients \( \beta_1 \) and \( \beta_3 \) should be interpreted as the price elasticity of demand for futures contract by arbitragers and speculators. If a small deviation of the future price from the cost carry equilibrium is immediately corrected, futures demand by arbitragers is said to be highly elastic. On the other hand, if mispricing persists without any substantial change in the quantity demanded, the demand is said to be inelastic. Markets where arbitrage is highly inelastic move independently with weakly correlated prices. In the case of the futures market, this indicates a failure to perform the price discovery and risk transfer functions.

To determine if the futures price, the cost of carry price and the expected spot price in Equation (5.18) are co-integrated, the study employs the widely used Engle and Granger (EG) test for cointegration. Developed in 1987 by Engle and Granger, the EG method allows for examining the long-run relationship between variables as well as short-run deviations from equilibrium. One of the benefits of this method is that the long-run equilibrium can be estimated between the variables at levels. The first step involves the estimation of a cointegration regression between the underlying variables. To determine if the variables are co-integrated, the residuals obtained from the estimated equation are tested for stationarity. A
cointegrating relationship is said to exist between the variables if the residuals from the estimated regression are stationary.

To test for a unit root in the estimated residuals, the EG and Phillips and Ouliaris (PO) cointegration tests are used. The two methods differ in the manner in which serial correlation in the residuals is corrected. The EG test uses a parametric correction for higher-order correlation in the residuals. It assumes that the residual series follows an AR\((p)\) process; thus \(p\)-lagged differences of the dependent variable are included in the regression of the form:

\[
\Delta u_t = (\rho - 1)u_{t-1} + \sum_{j=1}^{p} \delta_j \Delta u_{t-j} + v_t
\]  

(5.19)

The EG cointegration test is based on the following two test statistics for unit root:

\[
t = \frac{\hat{\rho} - 1}{se(\hat{\rho})}
\]  

(5.20)

\[
\hat{z} = \frac{T(\hat{\rho} - 1)}{1 - \sum \delta_j}
\]  

(5.21)

where \(T\) is the sample size and \(se(\hat{\rho})\) is the estimator of the standard error of \(\hat{\rho}\). The PO test, on the other hand, uses a non-parametric PP methodology employing the following unaugmented Dickey–Fuller regression to obtain an estimate of \(\rho\):

\[
\Delta u_t = (\rho - 1)u_{t-1} + \omega_t
\]  

(5.22)

Test statistics corresponding to Equations (5.21) and (5.22) are:

\[
Z_t = \frac{(\hat{\rho}^* - 1)}{se(\hat{\rho}^*)}
\]  

(5.23)

\[
Z_\alpha = T(\hat{\rho}^* - 1)
\]  

(5.24)

where \(\hat{\rho}^*\) is the bias-corrected autocorrelation coefficient. The test statistics obtained from the EG and PO tests do not follow the limiting distribution tabulated by Dickey and Fuller. As a result, the critical values outlined by Mackinnon (1991) are used for significance testing. It must be noted that the ordinary least squares (OLS) method is used for estimating the
cointegration regression: the \( t \)-statistics of the estimated coefficients will not have asymptotic normal distribution and hence they cannot be used to conduct valid inference. This problem arises because simple OLS regression does not take into account the presence of endogeneities and serial correlation in the regressors, which makes the estimates asymptotically biased as their limiting distributions are shifted away from the true parameters Phillips and Hansen (1990). For this reason, Equation (5.18) is estimated using the Phillips and Hansen (1990) fully modified ordinary least squares method (FMOLS), which overcomes this problem.

To determine whether the individual influence of arbitrage and speculative trading is sufficient for the determination of WTI crude oil futures prices, several key restrictions are applied to the parameters of Equation (5.18). These restrictions include \( \beta_0 = 0 \), \( \beta_1 = 1 \), \( \beta_2 = 0 \) and \( \beta_1 + \beta_2 = 1 \). The Wald coefficient restriction test is used to test these restrictions, such that the rejection of the null hypothesis would imply that the respective coefficient restriction fails to hold.

### 5.5 Empirical Results

In this section, the results of the empirical analysis are now presented, commencing with a discussion of unit root tests. Table 5.4 provides the unit root tests results for observed futures price, the cost of carry price and the spot price at log levels.

The null hypothesis for the unit root is not rejected at the 5% level with or without the trend component. This indicates that the mean, variance or both are not constant over the sample period. To determine the order of integration of the variables included in the model, unit root tests are undertaken once more for all the variables at first difference. Table 5.5 shows that all of the variables are stationary at first difference. The null hypothesis is rejected at the 5%
level with and without the trend component. As these variables are integrated of the same order, \((I(1))\), a long-run relationship may exist if some linear combination of the variables is stationary.

### Table 5.4: Unit Root Tests at Log Levels

<table>
<thead>
<tr>
<th>Variable</th>
<th>Augmented Dickey–Fuller</th>
<th>Phillips–Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td>(F_{t, t+1})</td>
<td>-1.0955</td>
<td>-2.7949</td>
</tr>
<tr>
<td>(F_{t, t+3})</td>
<td>-0.9155</td>
<td>-2.5426</td>
</tr>
<tr>
<td>(F_{t, t+6})</td>
<td>-0.7277</td>
<td>-2.2794</td>
</tr>
<tr>
<td>(F_{t, t+9})</td>
<td>-0.6115</td>
<td>-2.1067</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+1})</td>
<td>-1.2376</td>
<td>-2.9112</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+3})</td>
<td>-1.2694</td>
<td>-2.9521</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+6})</td>
<td>-1.3247</td>
<td>-3.0199</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+9})</td>
<td>-1.3899</td>
<td>-3.0974</td>
</tr>
<tr>
<td>(S_t)</td>
<td>-1.2218</td>
<td>-2.8916</td>
</tr>
</tbody>
</table>

Note: All variables are expressed natural logs.

### Table 5.5: Unit Root Tests at First Difference

<table>
<thead>
<tr>
<th>Variable</th>
<th>Augmented Dickey–Fuller</th>
<th>Phillips–Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td>(F_{t, t+1})</td>
<td>-16.4932**</td>
<td>-16.4688**</td>
</tr>
<tr>
<td>(F_{t, t+3})</td>
<td>-16.0976**</td>
<td>-16.0734**</td>
</tr>
<tr>
<td>(F_{t, t+6})</td>
<td>-15.8350**</td>
<td>-15.8121**</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+1})</td>
<td>-16.7512**</td>
<td>-16.7241**</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+3})</td>
<td>-16.6708**</td>
<td>-16.6441**</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+6})</td>
<td>-16.5483**</td>
<td>-16.5221**</td>
</tr>
<tr>
<td>(\ddot{F}_{t, t+9})</td>
<td>-16.4370**</td>
<td>-16.4112**</td>
</tr>
<tr>
<td>(S_t)</td>
<td>-16.8146**</td>
<td>-16.7873**</td>
</tr>
</tbody>
</table>

Note: All variables are expressed natural logs. * and ** indicate significance at the 10 and 5% levels.
Having determined the order of integration of variables included in the model, the analysis continues by testing for cointegration between the observed futures price, the cost of carry price and the expected spot price. Prices of 1, 3, 6 and 9-month WTI crude oil futures contracts are used in the analysis. A cointegration regression is estimated using FMOLS, and a unit root test is performed on the residual series. The FMOLS method provides asymptotically unbiased estimates of the standard errors, thus allowing for a more robust testing of the significance of the regression coefficients. The presence of unit root in the residual series obtained from regression of Equation (5.18) leads to the acceptance of the null hypothesis of no cointegration between observed and theoretical futures prices. The results of the cointegration regression and coefficient restrictions are presented in Table 5.6. The results show that the null hypothesis of no cointegration is rejected at the 5% level for all maturities by the EG and PO test statistics, which is consistent with the results obtained by Moosa and Al-Loughani (1995). This indicates that arbitrage exerts a significant influence on the observed futures price. The magnitudes of the coefficients for cost of carry price for the four maturities are 0.9185, 0.9237, 0.8594 and 0.8678. This suggests that over 85% of the variability in the observed futures price can be explained by arbitrage. The restriction $\beta_1 = 1$ is rejected in all cases, which implies that although arbitrage plays a statistically significant role in the determination of the WTI futures price, it is not sufficient for achieving equilibrium represented by the cost of carry model.
The expected spot price has a positive and significant impact on the futures price at the 5% level for all maturities. The restriction $\beta_2 = 0$ is rejected whereas the restriction $\beta_1 + \beta_2 = 1$ cannot be rejected at the 5% significance level. These results imply that the determination of crude oil futures prices can be explained on the basis of arbitrage and speculation. The results also confirm our propositions regarding the role of hedging. The magnitude of the coefficient for the expected spot price increases with the maturity period, which indicates greater explanatory power of speculation at longer maturities. Also noteworthy is that the expected profit from speculation, reported in Table 5.7, also increases with the maturity.
Table 5.7: Expected Profit from Speculation

|                  | \( EP_{t,t+n} \)  
|------------------|---------------------
|                  | \( (n=1) \) | \( (n=3) \) | \( (n=6) \) | \( (n=9) \) |
| Mean             | 0.018* | 0.024** | 0.0586** | 0.093* |
| Std. dev.        | 0.078  | 0.164   | 0.238    | 0.268  |

Note: \( EP_{t,t+n} \) represents the average expected profit from speculation at time \( t \) over period \( t + n \). The expected profit is calculated as the natural logarithm of the difference between the futures price at time \( t \) and expected spot price prevailing at the maturity of the futures contract, \( \ln(F_{t,t+n} - E_{t}S_{t+n}) \); * and ** indicate significance at the 10 and 5% levels.

This result is in line with the findings of Pagano and Pisani (2009) who report a positive association between the futures risk premium and speculative positions in the crude oil futures market in 1999 and 2000 and late 2003.\(^{30}\) Similarly, Melolinna (2011) find the risk premium in the crude oil futures market to be negative and increasing with contract maturity.\(^{31}\) The change in the expected spot price coefficient is also consistent with variation in the futures market risk premium observed by Hamilton and Wu (2014). Hong (2000) attributes this finding to heterogeneity in the informational endowments of traders. He argues that information asymmetry rises as a futures contract nears maturity. As a result, less private information is incorporated into the futures prices, which decline as the contract nears maturity. He calls this phenomenon the ‘speculative effect’.

---

\(^{30}\) Pagano and Pisani (2009) follow Pindyck (2001) by estimating the risk premium as the difference between the futures price at time \( t \) and the expected spot price at the maturity of the futures contract.

\(^{31}\) The statistical significance of the risk premium, which is estimated by using a rolling regression, changes over the sample period.
To determine if the inclusion of convenience yield plays a significant role in determining the impact of arbitrage activity, Equation (5.18) is estimated by excluding the convenience yield measure from the net cost of storage. Table 5.8 provides the results of this estimation. The null hypothesis of no cointegration is rejected based on the four test statistics. The null hypotheses of $\beta_1 = 1$ is also rejected, whereas the hypothesis $\beta_1 + \beta_2 = 1$ cannot be rejected at the 5% level. It can be seen that the values of the coefficient for the theoretical futures price in the absence of convenience yield are similar to those obtained by including the

<table>
<thead>
<tr>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n=1$</td>
<td>$n=3$</td>
<td>$n=6$</td>
<td>$n=9$</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>-0.0237</td>
<td>-0.054</td>
<td>-0.0783</td>
</tr>
<tr>
<td></td>
<td>(-1.01)</td>
<td>(-1.33)</td>
<td>(-0.94)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9229</td>
<td>0.9236</td>
<td>0.8652</td>
</tr>
<tr>
<td></td>
<td>(33.12)**</td>
<td>(30.92)**</td>
<td>(20.89)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0766</td>
<td>0.0688</td>
<td>0.1115</td>
</tr>
<tr>
<td></td>
<td>(2.69)**</td>
<td>(2.27)**</td>
<td>(2.56)**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>(2.69)**</td>
<td>(2.56)**</td>
<td>(2.49)**</td>
</tr>
</tbody>
</table>

$R^2$ 0.9971 0.9941 0.9831 0.9732
Adj.$R^2$ 0.997 0.994 0.9829 0.9729

$t$ -11.30** -5.39** -5.22** -4.81**
$z$ -180.42** -64.46** -50.12** -43.06**
$Z_\alpha$ -11.62** -10.21** -5.05** -4.66**
$Z_\tau$ -196.87** -163.51** -46.33** -39.94**

$\beta_0 = 0$ -1.01 -1.33 -0.94 -1.41
$\beta_1 = 1$ -2.77** -2.56** -3.25** -2.51**
$\beta_2 = 0$ 2.69** 2.27** 2.56** 2.11**
$\beta_1 + \beta_2 = 1$ -0.05 -0.46 -0.7 -0.3

$N$ 303 301 298 295

Note: * and ** indicate significance at the 10 and 5% levels.
convenience yield. These findings suggest that excluding convenience yield does not significantly alter the impact of arbitrage activity on the futures price.

A test of the equivalence of the coefficients from the two regressions must be conducted to determine if the difference between the coefficients is statistically significant. For this purpose, we use the z-test for the equality of two coefficients outlined by Clogg et al. (1995) and Paternoster et al. (1998). The formula for the Z-statistic is given by:

\[ Z = \frac{b_1 - b_2}{\sqrt{(SEb_1)^2 + (SEb_2)^2}} \]  (5.25)

where \( b_1 \) and \( b_2 \) are the estimated values of the coefficients and \( SEb_1 \) and \( SEb_2 \) represent standard errors of the coefficients. We test the null hypothesis \( b_1 = b_2 \) against the alternative, \( b_1 \neq b_2 \). The results of the z-test are displayed in Table 5.9.

<table>
<thead>
<tr>
<th>Table 5.9: Test of Coefficient Equality</th>
</tr>
</thead>
<tbody>
<tr>
<td>t + 1</td>
</tr>
<tr>
<td>Z-statistic</td>
</tr>
</tbody>
</table>

According to the test results, the null hypothesis cannot be rejected for any maturity, which means that exclusion of the convenience yield has no significant impact on the role of arbitrage activity. This finding also confirms the assertion of Moosa and Al-Loughani (1995) that financial traders do not realise any benefits from holding the commodity.

To further gauge the importance of convenience yield in the model, a comparison is drawn between observed futures price and theoretical futures price. The theoretical futures price takes two forms, the first of which is based on the simple cost of carry model; the second includes the convenience yield approximation. Table 5.10 reports the results of the actual and
percentage difference between the observed futures price and the two sets of theoretical futures prices over 1, 3, 6 and 9-month maturity periods.

**Table 5.10: Impact of Estimated Convenience Yield on West Texas Intermediate Futures Price**

<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Futures Price Difference</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple cost of carry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>–0.255</td>
<td>–0.582</td>
<td>–1.333</td>
<td>–2.167</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.638</td>
<td>2.312</td>
<td>3.567</td>
<td>4.461</td>
</tr>
<tr>
<td>Cost of carry adjusted for convenience yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>–0.232</td>
<td>–0.463</td>
<td>–1.027</td>
<td>–1.656</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.653</td>
<td>2.327</td>
<td>3.618</td>
<td>4.612</td>
</tr>
<tr>
<td><em>Percentage of Futures Price Difference</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple cost of carry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>–0.011</td>
<td>–0.029</td>
<td>–0.061</td>
<td>–0.094</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.037</td>
<td>0.069</td>
<td>0.107</td>
<td>0.138</td>
</tr>
<tr>
<td>Cost of carry adjusted for convenience yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>–0.009</td>
<td>–0.025</td>
<td>–0.051</td>
<td>–0.078</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.208</td>
<td>0.060</td>
<td>0.104</td>
<td>0.135</td>
</tr>
<tr>
<td>$N$</td>
<td>304</td>
<td>304</td>
<td>304</td>
<td>304</td>
</tr>
</tbody>
</table>

Note: Mean and standard deviation of the dollar and percentage difference between the actual futures price and theoretical futures price. The percentage difference is calculated with respect to the futures price.

Although the standard deviation of the percentage difference shows slight reduction, the average pricing error after the convenience yield adjustment is not statistically significant. This shows that convenience yield estimates based on the framework of Heaney (2002a) fail to explain the cost of carry pricing error. One potential explanation for these findings is that the assumption may be too extreme that the spot commodity holder demonstrates perfect timing ability when buying and selling the commodity. To draw further comparisons with the
study of Moosa and Al-Loughani (1995), Equation (5.18) is estimated over the sample period used by them.\textsuperscript{32}

Table 5.11: Cointegration Regression, January 1986 to December 1991

<table>
<thead>
<tr>
<th></th>
<th>$F_{t,t+n}$ (n=1)</th>
<th>$F_{t,t+n}$ (n=3)</th>
<th>$F_{t,t+n}$ (n=6)</th>
<th>$F_{t,t+n}$ (n=9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>$-0.0196$</td>
<td>$0.0537$</td>
<td>$0.2926$</td>
<td>$0.4686$</td>
</tr>
<tr>
<td></td>
<td>$(-0.20)$</td>
<td>$-0.3$</td>
<td>$-1.17$</td>
<td>$-1.64$</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>$0.8568$</td>
<td>$0.8404$</td>
<td>$0.7261$</td>
<td>$0.6544$</td>
</tr>
<tr>
<td></td>
<td>$(19.46)^{**}$</td>
<td>$(15.68)^{**}$</td>
<td>$(11.49)^{**}$</td>
<td>$(8.97)^{**}$</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>$0.143$</td>
<td>$0.1232$</td>
<td>$0.1424$</td>
<td>$0.1427$</td>
</tr>
<tr>
<td></td>
<td>$(3.21)^{**}$</td>
<td>$(2.16)^{**}$</td>
<td>$(1.98)^{*}$</td>
<td>$(1.67)^{*}$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>$0.9632$</td>
<td>$0.9198$</td>
<td>$0.8374$</td>
<td>$0.7331$</td>
</tr>
<tr>
<td>Adj.$R^2$</td>
<td>$0.9621$</td>
<td>$0.9174$</td>
<td>$0.8326$</td>
<td>$0.7252$</td>
</tr>
<tr>
<td>$\hat{t}$</td>
<td>$-5.76^{**}$</td>
<td>$-3.64^{*}$</td>
<td>$-3.55^{*}$</td>
<td>$-3.24$</td>
</tr>
<tr>
<td>$\hat{z}$</td>
<td>$-45.02^{**}$</td>
<td>$-22.93^{*}$</td>
<td>$-21.03^{*}$</td>
<td>$-18.95$</td>
</tr>
<tr>
<td>$\hat{Z}_\alpha$</td>
<td>$-5.82^{**}$</td>
<td>$-3.72^{*}$</td>
<td>$-3.73^{*}$</td>
<td>$-3.48$</td>
</tr>
<tr>
<td>$\hat{Z}_t$</td>
<td>$-45.73^{**}$</td>
<td>$-23.79^{*}$</td>
<td>$-23.38^{*}$</td>
<td>$-21.91$</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>$-0.2$</td>
<td>$0.3$</td>
<td>$1.17$</td>
<td>$1.64$</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>$-3.25^{**}$</td>
<td>$-2.98^{**}$</td>
<td>$-4.33^{**}$</td>
<td>$-4.73^{**}$</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>$3.21^{**}$</td>
<td>$2.16^{**}$</td>
<td>$1.98^{*}$</td>
<td>$1.67^{*}$</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>$-0.01$</td>
<td>$-0.6$</td>
<td>$-1.57$</td>
<td>$-2.13^{**}$</td>
</tr>
<tr>
<td>$N$</td>
<td>$71$</td>
<td>$69$</td>
<td>$66$</td>
<td>$63$</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

As the crude oil price does not exhibit any trends over this period (see Figure 5.3), the trend variable is excluded from the model over this period. The results for the cointegration regression with the modified sample period are outlined in Table 5.11. With the exception of a 9-month maturity period, the null hypothesis of no cointegration is rejected at the 10% level. The coefficient of the cost of carry price with a 3-month maturity period (0.8404) is

numerically smaller than the coefficient value of 0.925 reported by Moosa and Al-Loughani (1995), but the difference is not significant.\(^{33}\)

In general, the results presented in Tables 5.7 and 5.11 are consistent with those of Moosa and Al-Loughani (1995) in that the arbitrage is responsible for the majority of the short-term changes in the crude oil prices across different maturity periods. The impact of arbitrage activity is higher for near-term futures contracts in comparison to distant contracts. This can be attributed to the higher cost of arbitraging and uncertainty associated with distant contracts (Garbade and Silber, 1983).

5.6 Summary and Conclusion

In this study, we investigate the effectiveness of arbitrage and speculation in determining WTI crude oil futures prices. In line with the proposed extension to the study of Moosa and Al-Loughani (1995), we also specify a measure of convenience yield using the option pricing framework proposed by Heaney (2002a). The main conclusion emerging from this analysis is that although arbitrage plays a significant role in determining WTI crude oil futures price, this role is not exclusive. The results show that the inclusion of convenience yield does not lead to any significant change in the impact of arbitrage activity on futures price. This finding does not provide convincing evidence in support of the convenience yield effect in understanding the role of arbitrage. Also, it reinforces the argument of Moosa and Al-Loughani (1995) that market participants, such as financial institutions, merely hold the title to crude oil inventories and therefore do not realise any convenience yield. It may be that assumptions drawn while estimating the convenience yield are stringent, as the inventory holder is required to have perfect foresight with respect to future price movements, and to demonstrate perfect timing in selling the inventory at profit.

\(^{33}\) The test of coefficient equality results in a \(z\)-statistic value of 1.3887.
We also find that mispricing in futures contracts cannot be fully corrected by arbitrage alone. This calls into question the adequacy of the cost of carry model in explaining the price of crude oil futures. Finally, the findings support the view that speculative activity has predictive power in the determination of crude oil futures prices. However, it is important to note that this finding may vary depending on the expectation formation mechanism used to proxy the expected spot price of crude oil.
6.1 Introduction

Developing an understanding of the expectation-generating processes used by agents has received considerable attention in the literature investigating the role of trading behaviour in driving market prices. The rational expectation hypothesis (REH) (attributed to the early works of Muth (1961), Lucas (1972) and Sargent (1973) has long been the leading paradigm in modelling expectations. However, due to the limited success of representative agent models in predicting or even explaining market behaviour, recent research has witnessed a shift away from the assumptions underlying REH.

The literature has instead focused on alternative behavioural approaches where the expectations of traders are boundedly rational (see, e.g., Conlisk (1996) for a survey). Broadly speaking, a boundedly rational agent can be described as being able to choose from a set of alternatives for making economic decisions, without knowing the exact structure of the economy (Simon, 1957). Consequently, the use of simple rules of thumb for making decisions under uncertainty is a more precise representation of behaviour than fully optimal decision rules based on perfect rationality. In forming expectations of variables, bounded rationality implies that agents do not know the true equilibrium process generating those variables, which means that ex-post realisations on average need not coincide with ex-ante expectations. Instead, boundedly rational agents employ simple heuristics and update their expectation formation process as they learn through the outcomes from past decisions.
In this chapter, we relax the assumption that expectations are formed rationally, in that the agents have complete knowledge of the true economic model determining the price of crude oil. Put simply, we assume that agents only possess knowledge of their objectives given their constraints and that they base their trading decisions on subjective expectations of the future evolution of the oil price. As a result, when faced with the same economic problem, different speculators make different decisions based on their expectation formation mechanisms. These mechanisms can differ in terms of their sophistication and information processing requirements. Different models of expectation formation mechanism are outlined in this chapter. The aim is to investigate how the influence of speculation on the price of crude oil varies according to the expectation formation mechanisms employed by speculators.

6.2 Theoretical Aspects of Expectations Formation

Literature on the dynamics of crude oil has mostly focused on the use of fundamental factors in explaining the behaviour of prices. Empirical models based on market fundamentals (such as those employed by Energy Information Administration) typically employ data on inventory levels, production capacity and other economic fundamentals to formulate expectations of the future oil prices. These models perform well in explaining the real demand and supply for the proportion of oil used for production and consumption purposes. However, commodity markets have been increasingly used for investment purposes (Domanski and Heath, 2007) for they can yield more attractive returns than can alternative markets (Edwards and Caglayan, 2001; Gorton and Rouwenhorst, 2006). Fundamental models, as a result, are limited in their ability to explain the price dynamics caused by speculative activity, as their impact by definition is not explained by fundamentals.

Fundamental models are based on several traditional and well-known theories of economic representations of reality. These theories have attracted a substantial amount of criticism over
the years, and include the REH, the efficient market hypothesis (EMH) and the representative agent theory. According to Hommes (2006), these assumptions are logically inconsistent and the stylised facts produced by these models do not coincide with the day-to-day empirical behaviour of the market.

The REH has been confidently applied to different economic models and theory since the advent of Lucas (1976) critique of the adaptive or error learning approach. The REH states that, on average, the expectations of investors on an economic variable coincide with the mathematical conditional expectation of the variable, which means that the estimates of investors are not systematically wrong (Muth, 1961). This suggests that no additional profit can be accrued by trading on the basis of trends, which would be known to all investors in a semi-strong efficiency market. Earlier arguments in support of the rational expectations framework were put forward by Alchian (1950) and Friedman (1953). They argued that irrational traders cannot sustain profit in a competitive market, as they would lose profit to their rational counterparts.

However, Hommes (2006) points out that if traders behave rationally, no trades will take place as all traders will have the same expectations. If one agent wants to sell on the basis of superior information, the potential buyer infers the decline in value and will not take the opposite position necessary for this trade to take place. This signalling mechanism prevents the trade from taking place, which is inconsistent with the observed excessive trading volume across different markets.

Closely associated with the REH is the concept of market efficiency posited by Fama (1965): the idea of efficient markets is based on the concept of informational efficiency. The EMH characterises a market as efficient if asset prices fully reflect the available information.
relevant to the pricing of the asset. In other words, efficiency can be defined as the equality of
the probability distribution of expected prices conditional on the information set relevant to
the underlying asset and the true distribution of prices. This relationship can be expressed as:

\[
\int_a^b S_t f_m(S_{t+1}|I_t) dS = \int_a^b S_{t+1} f(S_{t+1}|I_t) dS
\]  
(6.1)

where \(f_m(S_{t+1}|I_t)\) is the market-estimated conditional probability distribution and \(f(S_{t+1}|I_t)\)
is the true distribution of \(S_{t+1}\). The definition closely relates to the concept of Muthian
rationality, which can be stated mathematically as:

\[
S_{t+1} = E(S_{t+1}|I_t) + \varepsilon_t
\]  
(6.2)

where \(S_{t+1}\) is a continuous random variable, \(E(S_{t+1}|I_t)\) represents the mathematical
expectation for \(S_{t+1}\) conditional on the information set available at time \(t\); and \(\varepsilon_t\) is the
forecasting error. Sheffrin (1996) outlines two important properties of forecasting errors.
First, forecasting errors have a mean of 0, which follows directly from the mathematical
equality between the conditional expectations of agents and the actual values of the variable
on average (i.e. \(S_{t+1} = E(S_{t+1}|I_t)\)). Second, the errors must be uncorrelated with the
information set available at time \(t\). If forecasting errors are related to the available
information, it would be possible to reduce them by incorporating this correlation (Sheffrin,
1996).

If markets do operate efficiently, then investment strategies based on discerning systematic
patterns in historical prices (technical analysis) or analysis of factors influencing supply and
demand for the asset (fundamental analysis) should not yield returns greater than the returns
obtained from a randomly selected portfolio. As a result, no profitable opportunities will be
available due to the presence of rational arbitragers in the market. Price changes should
precede changes in fundamental economic factors pertaining to supply and demand. In
contrast, observed behaviour of financial prices in foreign exchange, stock and commodity
markets cannot be explained only by movements in the underlying fundamentals. For instance, Shiller (2000) observes a disconnection between stock and bond prices and their fundamental values. Similar evidence documented in the foreign exchange market has given rise to a well-known anomaly, the *volatility disconnect puzzle* (Sarno, 2005).

In the futures market, the notion of market efficiency can be summarised in the form of one important property: the expected price should be an unbiased predictor of the actual price. However, Gjølberg (1985) and Moosa and Al-Loughani (1994) reject the efficiency respectively of crude oil futures and oil-based products, and conclude that the WTI futures price is a biased predictor of the future spot price.

However, tests of unbiasedness are subject to some criticism due to the use of the actual price as a proxy for the expected price. The actual price proxy assumes that the change in the price arises due to the arrival of new information. Consequently, the equilibrium return should be equal to a random error term with 0 mean. To avoid this assumption, other studies employ survey-based data. However, tests of market efficiency based on survey data have also led to the rejection of unbiasedness (MacDonald (2000) for exchange rates; Pearce (1984) for stock prices; MacDonald and Marsh (1993) for crude oil prices) and the homogeneity of expectations (Frankel and Froot, 1987; MacDonald and Marsh, 1996; Sarno and Taylor, 2002).

Other explanations for price dynamics based on non-fundamental sources include the convenience yield (Working 1949; Schwartz and Smith, 2000). However, even this does not provide a conclusive explanation. Pindyck (2004) examines the relationship between volatility in commodity markets and fundamental factors such as inventory and convenience.

---

34 The proxy is necessitated by the unavailability of data on agents’ expectation of financial and speculative prices.
yields. Pindyck reports that the fundamentals can only be weakly related to spot price volatility and suggests that other factors such as investor irrationality and herding behaviour might account for volatility.

In modelling economic behaviour, the use of the representative agent approach is mainly attributed to the need for simplifying reality. However, a major challenge to the validity of this approach is that if agents share identical prior beliefs and draw common inferences, as suggested by the rational expectations approach, they will arrive at similar valuations for the same underlying asset. Kirman (1991) suggests that one possible alternative to the representative agent approach to economics could be agent-based simulation models, which are capable of dealing with heterogeneous agents. The behavioural finance literature addresses this gap by providing an alternative perspective based on cognitive biases and bounded rationality. The REH implies that mispricing cannot persist and is corrected by offsetting positions taken by rational arbitragers. However, the behavioural finance approach allows mispricing to continue due to the presence of traders with irrational beliefs (also known as ‘noise’ traders) (Barberis and Thaler, 2002). This is in contrast with the RE approach in which agents share identical prior beliefs, draw common inferences and correct any over- or undervaluation. De Long et al. (1990a, 1990b) contradict this notion and argue that irrational traders may overlook market signals and analyst forecasts and select assets based on personal research derived from limited information. These traders can drive prices away from fundamental values for considerable lengths of time. If noise traders persists, they may force rational investors to take losses or liquidate their positions early, even if they are correct about fundamental values.
6.2.1 Biases

Before evaluating the influence of noise trading it is necessary to understand what drives this activity. Kahneman and Tversky (1974) use the idea of cognitive biases to explain the persistence of irrational behaviour. Their findings show that investors do indeed demonstrate irrational behaviour. They allow their test subjects to make investment decisions under uncertainty, and find that their decisions are significantly influenced by cognitive biases. As these biases are used to construct expectations in their proposed model, Kahneman and Tversky propose a thorough analysis of the role of cognitive biases in decision making.

Representativeness or cognitive bias works in conjunction with confirmation bias (Einhorn and Hogarth, 1981), which is relevant for understanding trading behaviour that is consistent with extrapolation strategies. Representativeness bias pertains to the tendency to form expectations that conform to existing beliefs. Confirmation bias, on the other hand, involves assigning higher weight to events that confirm the underlying hypothesis and give lower weight to events that contradict the hypothesis. Together, they explain why people tend to disregard past events and their associated probabilities when such events are in conflict with their prior notions. In an investment scenario, these heuristics may encourage the selection of information that is in line with existing beliefs.

Consider the case of an investor who expects a stock to appreciate in the future. This investor may rely on a succession of positive returns to form their expectation of a positive return in the future, without taking into consideration the negative returns realised after positive returns (confirmation bias). As a result, the investor disregards the actual distribution of returns and assigns a higher probability to a positive return in the next period (representativeness bias).
Availability bias occurs when events that can be easily recalled from memory form the basis upon which the likelihood of recalled events is extrapolated to the future. If events are simple to recall, they would be expected to occur more frequently, and vice versa. This may cause a high likelihood to be associated with frequent events and a low likelihood to events that are difficult to recall in an investment setting, which may ultimately lead to incorrect decisions. Anchoring bias means that the most recent value of an asset will influence an investor’s estimate of its value in the next period. Kahneman and Tversky (1974) show that the initial observed value is adjusted to form an estimate of the asset price that is unrelated to the underlying fundamentals. In an investment setting, recent prices or returns are seen as starting points for future estimates.

Another kind of cognitive bias, attributed to Samuelson and Zeckhauser (1988), is the status quo bias, which implies that people prefer to stay in their current situation and do not change their behaviour unless they are provided with very strong incentives. This proposition contradicts the microeconomic literature on the rational agent theory, which asserts that investors react even to minor changes in variables, to optimise their utility. Status quo bias forms the basis of heterogeneous agent models with switching regimes, as it helps explain why agents do not behave optimally in switching between strategies.

Another influential study, by Kahneman and Tversky (1979), is based on prospect theory, which explains how subjects make investment decisions under uncertainty. Under the traditional approach, this problem is resolved by applying the expected utility theory, which postulates that the utility of an outcome is determined by its weighted probability as derived from the values of the final asset. Instead of evaluating an investment based on final asset values, Kahneman and Tversky evaluate the potential outcomes of the investment on the basis of gains and losses, and find that investors tend to assign a smaller weight to outcomes
that have lower probability in comparison to the outcomes that are certain. This tendency, known as the certainty effect, can contribute to risk aversion to certain gains and risk preference certain losses, such that the value function for gains is concave, whereas the value function for losses is convex. According to Kahneman and Tversky, the average investor prefers the gamble of a small, yet probable gain, over a large gain with smaller probability and equal expected value. This implies that in the case of a loss, investors prefer a small probability of a large loss over a small loss with high probability of equal expected value. This behaviour is inconsistent with the rational agent theory, in which investors assign the same decision rules to both gains and losses.

6.2.2 Bounded Rationality

Closely related to the concept of biases is the theory of bounded rationality, which posits that boundedly rational agents are constrained by their cognitive limitations and limited available information. As a result, they form different expectations regarding the future behaviour of prices. The theory of bounded rationality dates back to Simon (1957), who notes that simple heuristics used for decisions under uncertainty provide a better approximation of behaviour than do fully optimal decision rules. Boundedly rational agents do not have information on the true process governing the evolution of variables; rather they use time series observations to form expectations and update model parameters as new information becomes available. Through this trial-and-error process, agents learn about their environment through feedback on their past decisions.

Some researches argue that this form of adaptive learning by boundedly rational agents, whose beliefs coincide with the economic environment, would ultimately lead to the same outcome as when they behave rationally. Such learning may converge on RE equilibria. However this learning process may also lead to non-RE equilibria, such as those discussed by
Bullard (1994). According to Bullard (2006), ‘some rational expectations equilibria are learnable while others are not. Furthermore, convergence will in general depend on all the economic parameters of a given system, including policy parameters’. Hence, the convergence towards rationality will depend on the feedback mechanism relating individual expectations to the economic environment, as the information may be interpreted differently by agents attempting to move towards RE equilibria.

6.2.3 Heterogeneous Expectations

The bounded rationality approach has become a useful paradigm for understanding investment behaviour outside rationality constraints. Further support for heterogeneous expectations comes from the ‘no trade’ argument. According to this argument, in a world where all agents acted rationally, there would be no trade. Although several no trade theorems are proposed in the literature, they are in sharp contrast to the daily trading volume observed across different markets in the real world.35 Contradictory evidence may be attributed to incomplete learning process (Frankel and Froot, 1987) or the voluntary use of limited information due to information costs (Feige and Pearce, 1976). Nevertheless, behavioural finance models with bounded rationality and heterogeneous expectations have emerged as plausible alternatives to the rational agent approach. Moreover, abundant evidence for heterogeneity in individual expectations is documented in the literature. For example, Frankel and Froot (1987, 1990) and Allen and Taylor (1990) note that forecasters employ different forecasting strategies to predict exchange rate movement. Using survey data on inflation expectations, Carroll (2003), Mankiw et al. (2003), Branch (2004) and Pfajfar and Santoro (2010) provide strong evidence for heterogeneous behaviour. Reitz and Slopek

35 For a discussion of the no-trade theorems, see Milgrom and Stokey (1982) and Fudenberg and Tirole (1991).

6.3 Expectation Formation Mechanisms

Although the importance of heterogeneous expectations is clearly recognised in the literature, the question that arises is how to model the expectations of speculators when they behave heterogeneously? The literature documents substantial evidence of speculators employing simple strategies to derive trading decisions. For instance, according to Smidt (1965), a large proportion of speculators apply price charts to predict the behaviour of commodity prices. Similar findings are reported by Canoles et al. (1998). Moreover, Sanders et al. (2000) document evidence of positive feedback trading in several commodity markets and Weiner (2002) reports herding behaviour in the petroleum market. These findings indicate that in forming expectations of futures price changes, speculators rely on technical and fundamental analysis to form price expectations.

Technical analysis involves extrapolation of historical price trends to form price predictions, whereas fundamental analysis is based on the assumption that prices eventually revert towards their long-run equilibrium or fundamental values. A number of studies have successfully used technical and fundamental analysis to explain the behaviour of traders in foreign exchange and stock markets (Day and Huang, 1990; De Grauwe et al., 1993; Brock and Hommes, 1998; LeBaron et al., 1999; Lux and Marchesi, 2000; Farmer and Joshi, 2002; Chiarella et al. 2002; Rosser et al., 2003). The core premise of technical and fundamental analysis is that asset prices embody a substantial endogenous component in addition to random exogenous shocks.
This assumption is supported by a large body of empirical research. For instance, Tversky and Kahneman (1974) use laboratory experiments to show that agents are boundedly rational and employ simple rules of thumb. Further, experiments conducted by Smith (1991) and Sonnemans et al. (2004) suggest that market participants employ simple forecasting rules based on extrapolative or regressive predictors. Survey studies performed by Taylor and Allen (1992) and Menkhoff (1997) reveal that foreign exchange traders rely on technical and fundamental trading rules to make investment decisions. In another study, Pilbeam (1995) shows that exchange rate models under extrapolative and adaptive expectations provide significant gains in explaining exchange rate movements.

It may be concluded that market prices are determined by complex interactions between traders who employ a mix of adaptive, extrapolative and regressive expectation formation mechanisms to predict price movements irrespective of the market in which they are trading. Guided by this evidence, this study employs several expectation formation mechanisms to generate oil price expectations of speculators. These include (i) extrapolative, (ii) adaptive and (iii) regressive expectations. Further, the random walk model is used as a benchmark to gauge the forecasting performance of these expectation formation mechanisms. A detailed description of these expectation formation mechanisms and the estimation procedures are now provided.

6.3.1 Extrapolative Expectations

If speculators use extrapolative expectations, the spot oil price at the maturity of the futures contract would be determined by rate of change in prices observed in the previous periods. A simple linear approximation for extrapolative expectations can be expressed as:

\[ E_tS_{t+n} - S_t = + \sum_{i=1}^{k} \beta_i (S_{t-i+1} - S_{t-i}) \]  

(6.3)
where \( n = 1, 3, 6 \) or 9. The expected change in the spot price in the period between \( t \) and \( t + n \) is equal to the weighted average of the monthly change in the previous \( n \) months. It is assumed that the expectations are formed at time and \( t \). The AIC and SBC are used to determine the optimal value for \( k \). In Equation (6.4), if \( \beta_i \) is positive, this would indicate that a price increase in the previous period would lead to increase in the price the next period. A negative \( \beta_i \), on the other hand, indicates that a previous period increase in price results in a decrease in the subsequent period. A \( \beta_i \) that is equal to 1 represents static expectations; whereas if \( \beta_i > 1 \), the expectations are explosive.

### 6.3.2 Adaptive Expectations

If the expectations of speculators are formed adaptively, oil price forecasts for the next period can be represented by the current expectations of oil prices adjusted for the prediction error. The standard adaptive expectations hypothesis takes the form:

\[
E_t S_{t+n} = E_{t-n} S_t + \gamma (S_t - E_{t-n} S_t),
\]

which can be rearranged as:

\[
E_t S_{t+n} - E_{t-n} S_t = \gamma (S_t - E_{t-n} S_t)
\]

where \( E_{t-n} S_t \) is the \( n \)-months-ahead expectation for spot price at time \( t - n \) for \( n = 1, 3, 6 \) and 9. According to Equation (6.6), forecasters base their expectations on the expected spot price at time \( t - n \) and a fraction \( \gamma \) of the forecasting error realised after \( n \) months. Equation (6.6) may be rewritten as:

\[
E_t S_{t+n} = \gamma S_t + (1 - \gamma) E_{t-n} S_t
\]
The expectations of the future spot price can now be shown as the weighted average of the current observed spot price and the expected spot price. If \( \gamma = 1 \) then expectations are static and the expected spot price is determined by the observed spot price at time \( t \). In the standard adaptive expectations process, the forecasting horizons correspond with the frequency of the observations. This study uses monthly observations, whereas forecasting horizons vary from 1 to 9 months. However, Prat and Uctum (2011) argue that forecasters may not wait until the \( n \)-month horizon to complete before they revise their estimates of the spot price. As a result, forecasters may compare the observed price at time \( t \) with the price expected in the previous month and not with the price expected \( k \) months before. This assumption leads to an adaptive process in which the expected spot price is expressed as a weighted average of the observed spot price and expected price in the past month, which gives:

\[
E_t S_{t+n} = \gamma S_t + (1 - \gamma) E_{t-1} S_t
\]  
(6.8)

Lagging Equation (6.8) by 1 month and multiplying both sides by \((1 - \gamma)\), we obtain:

\[
(1 - \gamma)E_t S_{t+n-1} = \gamma (1 - \gamma) S_{t-1} + (1 - \gamma)^2 E_{t-2} S_{t-1}
\]  
(6.9)

By continuously substituting the value of the expected spot price into Equation (6.9), we obtain:

\[
E_t S_{t+n} = \gamma [S_t + (1 - \gamma) S_{t-1} + (1 - \gamma)^2 S_{t-2} + \ldots \ldots]
\]  
(6.10)

or

\[
E_t S_{t+n} = \gamma \sum_{i=1}^{\infty} (1 - \gamma)^i S_{t-i+1}
\]  
(6.11)

The adaptive expectations process as represented by Equation (6.11) resembles a geometrically distributed lag model. According to this model, the expected spot price is the weighted average of the current and the previous period prices, and the impact of past prices decays geometrically.
6.3.3 The Random Walk Model

This study also includes one of the most widely used models for forming expectations in the early literature—the random walk model. According to this model, the change in oil price is purely random and is unpredictable (Copeland, 2000). Its simplest form, also known as the naïve random walk model, suggests that the expected price of crude oil is equal to the current price:

\[ E_t S_{t+n} = S_t + \varepsilon_{t+n} \]  (6.12)

where \( \varepsilon_{t+n} \sim IID(0, \sigma^2) \). According to Equation (6.12), the difference in crude oil price between two periods, \( \varepsilon_{t+n} \), is completely random with no discernible patterns. Thus, the random walk implies weak-form efficiency where the future price path cannot be determined on the basis of past history. In Equation (6.12), if the average difference in price between the two periods is non-zero then the process is said to follow a random walk with drift, which is specified as:

\[ E_t S_{t+n} = S_t + d + \varepsilon_{t+n} \]  (6.13)

where \( d \) is the drift term, which represents the average change in crude oil price from one period to another. If \( d > 0 \) then the process follows an upward trend, whereas \( d < 0 \) implies a negative trend.

The question arises as to whether oil price expectations should be generated using random walk with or without the drift term. Following Engel and Hamilton (1990) and Engel (1994), the choice between random walk with and without drift is based on the within-sample statistical significance of the drift term. If the drift term is significant then the out-of-sample forecast should be generated using random walk with drift. In line with Alquist and Kilian (2010), the drift term is estimated by regressing the current spot price (spot price at time \( t \)) on the spot price at time \( t-n \) where \( n = 1, 3, 6 \) and 9.
Table 6.1: Statistical Significance of the Drift Term

<table>
<thead>
<tr>
<th>$t-n$</th>
<th>$n = 1$</th>
<th>$t-n$</th>
<th>$n = 3$</th>
<th>$t-n$</th>
<th>$n = 6$</th>
<th>$t-n$</th>
<th>$n = 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.024</td>
<td>0.107</td>
<td>0.218</td>
<td>0.263</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(0.11)</td>
<td>(1.04)</td>
<td>(1.12)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1 shows that the drift term is not significantly different from 0 at the 5% confidence level, suggesting that the naïve random walk, rather than the random walk with drift, should be used for forecasting crude oil prices.

6.3.4 Regressive Expectations

Under a regressive specification mechanism, expectations for future spot price are influenced by deviations of the current spot price from its equilibrium value. This mechanism can be specified as:

$$E_t S_{t+n} - S_t = \beta_0 + \beta_1 (S_t - \bar{S}_t)$$  \hspace{1cm} (6.14)

where $\bar{S}_t$ represents the equilibrium oil price at time $t$ and $\beta_1$ is a measure of the speed of adjustment of the actual oil price to its equilibrium value. A negative $\beta_1$ indicates expectations of movement oil price back to its fundamental value whereas a positive sign would signify movement away from the fundamental value.

The next step in the estimation of regressive expectations mechanism involves specifying the equilibrium price of crude oil. A common practice in this regard involves the estimation of equilibrium price of crude oil as a moving average of the prices in previous periods (e.g., Ellen and Zwinkels, 2010). However, by using this approach stabilisation is more likely to occur towards the fundamental value. As a result, the sign of $\beta_1$ is likely to be negative. This study adopts an alternative approach by estimating the equilibrium price using oil market and
macroeconomic variables. The next section provides a brief overview of the literature to identify the key variables necessary for the estimation of equilibrium oil price.

6.4 Equilibrium Price of Crude Oil

The equilibrium or long-run price of crude oil is typically explained in terms of the macroeconomic fundamentals underlying the supply and demand of oil (Bacon, 1991; Dées et al., 2007; Alquist et al., 2011). With this in mind, we identify several key macroeconomic and oil market variables that have been used extensively in the literature. These variables are categorised into supply side and demand side determinants. The following is a brief description of these variables as well as the key findings of related studies.

6.4.1 Oil Demand

Although there is little consensus in the literature on the key drivers of crude oil demand, those pertaining to global economic activity have been shown to have a significant influence on the oil market. For example, Hamilton (2009) attributes the rally in oil prices from 2003 to 2008 to the corresponding higher energy demand resulting from the growth of global Gross Domestic Product (GDP). Similarly, Kilian and Hicks (2009) find that unexpected growth in emerging and advanced economies is associated with fluctuations in oil prices during the period 2000 to 2008. A strong influence of economic growth on oil prices is also observed by Kilian (2009), who finds that oil price shocks since 2001 have been primarily driven by changes in real economic activity, whereas previous period shocks are consistent with rising precautionary demand for oil. Reitz et al. (2012) analyse the demand for oil in emerging economies and find that changes in Chinese oil imports explain a significant portion of the fluctuations in oil prices.
The evidence provided by the literature indicates that movement in oil prices can be reasonably attributed to global economic output. A suitable measure of economic activity is therefore needed to approximate the fundamental value of oil. This study follows Kilian (2009) by using ocean freight rates for industrial commodities as a measure of worldwide economic growth. The freight rates are proxied by the Baltic Dry Index (BDI) obtained from the Bloomberg database. Provided by the Baltic Exchange, the BDI is based on the daily averages of the shipping costs for major raw materials by sea. This measure is chosen as because cycles in economic activity have been shown to have a significant impact on the demand for seaborne transport services (Klovland, 2004). Moreover, the BDI has been extensively used by market practitioners to assess global economic activity and the demand for shipping services (Coleman, 2012). The availability of the index at daily as well as monthly frequencies make it a suitable candidate for analysing short-term demand fluctuations that may be unobservable from quarterly indicators such as GDP or world oil consumption.

A second measure of oil demand included in the model is the exchange rate of the US dollar. The US dollar is used frequently as the invoicing currency for international crude oil trading. If the US dollar depreciated, the demand for oil would appreciate as foreign buyers would be willing to pay more dollars for oil. In contrast, appreciation in the US dollar would reduce the demand for oil.36 The strong relationship between exchange movements and commodity prices is reported in a number of studies. Pindyck and Rotemberg (1990) find a significant negative relationship between oil prices and an equally weighted index of the dollar in terms of pound sterling, the Yen and the Deutsche mark. Sadorsky (2000) finds the crude oil futures prices to be co-integrated with the trade-weighted US dollar index. Similarly, Chen and Chen

---

36 The demand for oil is highly inelastic (e.g., Hamilton, 2008), which means that the impact of exchange rates on the demand for crude oil price is likely to be small.
(2007) identify a significant long-run relationship between the exchange rate of the US dollar and different grades of crude oil prices. To determine the role of the exchange rate in explaining oil prices, the model employs the trade-weighted US dollar index, which is a weighted average of the bilateral exchange rates against the currencies of major trading partners. This data set is sourced from the Federal Reserve Bank of St Louis (the St Louis Fed).

6.4.2 Oil Supply

Factors affecting oil supply typically pertain to the available quantities of oil, and the cost of exploration and production. Coleman (2012) argues that measures of oil supply should include global oil reserves, but the relevant data cannot be reliably obtained. This study follows Merino and Ortiz (2005) by using the OECD oil inventory as a proportion of its production to represent oil supply. Using the inventory level to explain oil price changes can be justified both theoretically (Pyndick, 1994; Pyndick, 2001; Considine and Larson, 2001) and empirically (Ye et al., 2002; Dées et al., 2007). This measure implies that low oil stock levels may lead to expectations on future supply shortages and consequently add a premium to current prices. Although total oil stocks would be a more comprehensive measure of the oil supply, data are only available at a quarterly frequency.

A second proxy for oil supply is OPEC’s production. This particular proxy accounts for OPEC’s control over global oil prices via its decisions on quotas, production and capacity utilisation (Bacon, 1991). This measure follows the study of Coleman (2012) who argues that OPEC’s ability to influence prices is dependent on its share of global oil production. Following Kaufmann (2011), Kilian and Park (2009), and Prat and Uctum (2011), OPEC’s spare production is also included to proxy crude oil supply. Spare capacity is often used to measure the impact of supply shocks on oil prices.
6.5 Summary and Conclusion

Expectations regarding future outcomes of variables play a central role in determining the investment and trading decisions of market participants. For many years the rational expectations (RE) approach was the dominant paradigm for modelling economic expectations. Tests of the RE model have revealed that empirical observations contradict assumptions underlying the model. The recent literature therefore shifts emphasis towards boundedly rational models with heterogeneous expectations. The survey of the literature shows that the behaviour of market participants is at odds with the assumption of homogeneity (postulated by the REH) and indicates that expectations are formed heterogeneously.

The survey also shows that models of expectation formation based on the historical behaviour of prices (technical analysis) and market fundamentals (fundamental analysis) are successfully applied to explain price dynamics across different markets. In line with the evidence obtained from the preceding review, several different expectation formation mechanisms are considered here. We now move on to focus on the estimation of oil price expectations and their implications for how speculative activity affects crude oil prices.
7.1 Introduction

The analysis conducted in Chapter 5 indicates that speculative activity (as measured by the expected spot price) plays a significant role, along with arbitrage, in the determination of crude oil futures price. The analysis is now extended and a more complete assessment of the relationship between expectation formation mechanisms and the impact of speculative activity is provided by allowing expectations to be formed heterogeneously. Several expectation formation mechanisms to proxy the expected spot price prevailing on the maturity of the futures contracts are described in the preceding chapter.

We assume that speculators select an expectation formation mechanism on the basis of its profitability. This is based on a simple idea: if a particular expectation formation mechanism is used by the majority of futures traders—who base their decisions to buy or sell on the signals obtained from the expectation mechanism—they influence the supply and demand for futures contracts and, in the process, drive the price of crude oil. For instance, if an expectation formation mechanism that is adopted by the majority of traders generates a buy signal, then the traders will react by taking a long position on the commodity, causing the price of crude oil to rise. The dominant expectation formation mechanism is the one that generates the highest level of profit for speculators. This reasoning is based on the assertion that expectations regarding the future behaviour of prices leads to actions, which in turn create events (Davidson, 1982; Harvey, 1999).
The analysis commences with a description of the data, which is followed by the methodology used for the estimation of expected spot price proxies. The predictive accuracy of expectation formation mechanisms is also evaluated. Finally, the results of cointegration analyses involving the observed futures price, the theoretical futures price and the expected spot price of crude oil are discussed.

7.2 Data

This section provides a description of the data sample and the time series characteristics of the variables used in this analysis. Data description is conducted in conjunction with theory and the previously predicted relationship between the variables and the price of crude oil. This study uses monthly spot and futures prices of the WTI crude oil traded on the NYMEX from January 1988 to April 2013. The sample used for the estimation of expectation formation mechanisms is constructed such that the date of observation of the spot price corresponds with the observation date of the futures price. In addition to the futures and spot prices, the sample contains macroeconomic and oil market-specific variables that are used in the estimation of the equilibrium oil price. These variables include the BDI (representing economic activity), the US dollar exchange rate, OECD crude oil inventories, OPEC crude oil production and OPEC’s surplus production capacity. Specific details of these variables are presented in Table 7.1 along with a brief description of the data and sources. The majority of the data are sourced from the Bloomberg and the EIA, supplemented by data sourced from the St Louis Fed.

The sample period includes monthly observations of the variables from January 1988 to April 2013. However, the monthly data for OPEC surplus production capacity are only available

---

37 Refer to Section 5.3 for a detailed description of the futures price sample.
for January 1994. Therefore, the sample period used for the estimation of equilibrium oil price and regressive expectation mechanism is January 1994–April 2013.

**Table 7.1: Data Description**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Data description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baltic Dry Index</td>
<td>BDI</td>
<td>Monthly values of average shipping costs for major raw materials by sea available immediately before the date of observation of the oil price</td>
<td>Bloomberg</td>
</tr>
<tr>
<td>Trade-weighted US dollar index</td>
<td>ERATE</td>
<td>End of the month values of the weighted average of the foreign exchange value of the US dollar against a subset of the broad index currencies that circulate widely outside the country of issue</td>
<td>St Louis Fed</td>
</tr>
<tr>
<td>OECD crude oil inventory</td>
<td>INV</td>
<td>End of the month values of the crude oil inventories held by OECD countries measured in thousands of barrels</td>
<td>EIA</td>
</tr>
<tr>
<td>OPEC crude oil production</td>
<td>PROD</td>
<td>Monthly averages of the daily crude oil production by OPEC member countries measured in thousands of barrels</td>
<td>EIA</td>
</tr>
<tr>
<td>OPEC spare capacity</td>
<td>CAPACITY</td>
<td>Monthly surplus capacity of OPEC member countries measured in thousands of barrels</td>
<td>EIA</td>
</tr>
<tr>
<td>WTI crude oil spot price</td>
<td>S</td>
<td>Monthly spot price of crude oil available at the expiry of the futures contract</td>
<td>Bloomberg</td>
</tr>
</tbody>
</table>

The variables are transformed into natural logarithms to ensure that the data sample approximates more closely the assumptions of linear regression analysis. This also helps prevent some observations from exerting an extreme influence on the estimated coefficients. Moreover, specifying the variables in logarithmic form should facilitate drawing comparisons with previous studies. The data sample is then visually inspected to gain some initial insights into the time series characteristics of the variables. In this regard, the time series graph of natural logarithms of each variable is examined for long-term trends and changes in volatility. This approach provides useful information for determining the appropriate method for detecting non-stationarity.
Figure 7.1: Baltic Dry Index

Figure 7.1 shows the time series behaviour of BDI relative to the spot price of crude oil. The BDI appears to track oil prices, indicating a positive relationship between the two time series. The correlation matrix in Table 7.2 reveals a positive correlation of 0.48 between the oil price and the BDI, which confirms a positive relation between shipping cost for commodities and oil price.

Table 7.2: Correlation Matrix for Oil Price Fundamentals in Log Levels

<table>
<thead>
<tr>
<th></th>
<th>SPOT</th>
<th>BDI</th>
<th>ERATE</th>
<th>PROD</th>
<th>INV</th>
<th>CAPACITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>1</td>
<td>0.48</td>
<td>-0.78</td>
<td>0.87</td>
<td>0.74</td>
<td>-0.43</td>
</tr>
<tr>
<td>BDI</td>
<td>0.48</td>
<td>1</td>
<td>-0.44</td>
<td>0.38</td>
<td>0.26</td>
<td>-0.56</td>
</tr>
<tr>
<td>ERATE</td>
<td>-0.78</td>
<td>-0.44</td>
<td>1</td>
<td>-0.62</td>
<td>-0.67</td>
<td>0.36</td>
</tr>
<tr>
<td>PROD</td>
<td>0.87</td>
<td>0.38</td>
<td>-0.62</td>
<td>1</td>
<td>0.79</td>
<td>-0.52</td>
</tr>
<tr>
<td>INV</td>
<td>0.74</td>
<td>0.26</td>
<td>-0.67</td>
<td>0.79</td>
<td>1</td>
<td>-0.18</td>
</tr>
<tr>
<td>CAPACITY</td>
<td>-0.43</td>
<td>-0.56</td>
<td>0.36</td>
<td>-0.52</td>
<td>-0.18</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 7.1 also suggests that the Baltic Dry Index index captures major events in the oil market over the sample period, as indicated by the shaded portions of the graphs. In
particular, the index experienced an expansionary period in the late 1990s, and during the increase in oil prices after 2002 and the global financial crisis in 2007. Moreover, the graph displays a rising trend after 2002 amid strong demand from emerging economies, whereas a downward trend is observed in 2008 due to the reduction in global output. The variance and mean of the BDI appear to vary over time, which suggests the possibility of non-stationarity. Figure 7.2 shows crude oil production by OPEC member countries, which exhibits a positive long-term trend over the sample period, indicating a non-constant mean. In the first half of the sample, oil production experienced a few abrupt changes (indicated by the shaded areas) due to the Gulf War in 1900 as well as labour and political unrest in Venezuela in 2002; as a result, oil prices rose sharply. Oil production fell again in 2003 due to military action in Iraq. Subsequently, OPEC’s production increased to meet growing international demand, but was cut back in 2006 and 2007 due to concerns over the increase in OECD inventories. A sharp decline in oil production is apparent in 2008–09 as OPEC responded to reduction in oil demand due to the global financial crisis. Oil production was reduced again in 2011 due to the military intervention in Libya, and then increased in 2012 and 2013 to normal trend levels to meet growing demand from Asian economies. Overall, oil production of OPEC members shows a positive long-term trend with sharp fluctuations resulting from disruptions to production facilities and global demand shocks.

Figure 7.2: OPEC Crude Oil Production

Figure 7.3 shows the trade-weighted index of the US dollar against other currencies. The index is calculated as the geometric average of the foreign exchange value of the US dollar against a subset of the currencies of major trading partners.\(^40\) The index exhibits relatively high variability over the sample period with several upwards and downwards trends as indicated by the shaded areas. The index reaches its peak level in 2002 when the oil prices are relatively low. After 2002, the index continues to decline until the end of the sample period. Oil prices, on the other hand, exhibit a positive trend over the same period. The graph therefore indicates an inverse relationship between oil prices and the exchange rate index. This is confirmed by a negative correlation coefficient of \(-0.78\) (see Table 7.2) between the two time series.

The relationship between OECD crude oil inventories and oil price is plotted in Figure 7.4, in which the inventory level exhibits a seasonal pattern while conforming to a positive long-term trend. The seasonal behaviour of crude oil inventories typically results from variations in inventory level in response to changes in oil demand during the year. The graph also shows some sharp fluctuations in inventory level that are consistent with OPEC production changes. For example, the declines in inventory level in 2003 and 2008 correspond to OPEC’s production cuts that occurred in those periods due to the war in Iraq and the global financial crisis. This indicates strong dependence of OECD oil inventories on OPEC’s production. A correlation coefficient of 0.79 (see Table 7.2) provides further evidence of a strong relationship between the two variables. Fluctuations in crude oil inventory also appear to vary in magnitude over the sample period, which may suggest the presence of a unit root.
Figure 7.5 shows a plot of OPEC’s spare production capacity against the price of crude oil. The EIA defines the spare or surplus capacity as ‘the volume of production that can be brought on within 30 days and sustained for at least 90 days’. Spare capacity is essentially used as a measure of the market’s ability to respond to potential supply disruptions. During periods of tight oil supply, spare production can also act as an indicator of OPEC’s ability to exert an upwards influence on oil prices. The oil price appears to increase during periods when spare capacity is low and vice versa. For instance, from 2000 to 2003 the oil price declines as spare capacity builds up. Subsequently, from 2004 to 2008, the spare production levels remains low, limiting OPEC’s ability to respond to rising demand. During this period, the oil price displays a sharp increase implying an inverse relationship between the two variables. The correlation matrix in Table 7.2 also provides a negative correlation of –0.43 between OPEC’s spare production capacity and crude oil price. Although the spare capacity does not exhibit any clear trend, the volatility appears to vary over the sample period, which suggests a non-stationary process.
7.3 Modelling the Equilibrium Price of Oil

The equilibrium oil price represents a fair value to which market participants expect the oil price to converge over time. The procedure adopted for estimating the equilibrium price involves testing a long-run equilibrium relationship between the WTI oil cash price and oil market fundamentals, using the cointegration approach. This relationship between the oil price and explanatory variables can be expressed by the following model:

\[ S_t = \beta_0 + \beta_1 BDI_t + \beta_2 ERATE_t + \beta_3 PROD_t + \beta_4 INV_t + \beta_5 CAPACITY_t + \beta_6 t + \nu_t \]  \hspace{1cm} (7.1)

where \( BDI \) is the Baltic dry index, \( ERATE \) represents the trade-weighted index of the US dollar, \( PROD \) is OPEC’s crude oil production, \( INV \) is OECD crude oil inventories, \( CAPACITY \) is OPEC’s spare capacity and \( t \) is the linear time trend, which is included to account for the behaviour observed in the graphical analysis. It should be noted that the residual series obtained from the estimated model, \( \nu_t \), is equal to the regressive or mean reversion component \( (S_t - \bar{S}) \) in Equation 6.14.
The Phillips and Hansen (1990) FMOLS method is used to estimate Equation (7.1). Although it is possible to estimate the regression using the OLS approach, the resulting \( t \)-statistics for the estimated coefficients will not have asymptotic normal distribution because the regressors are integrated variables. Therefore, the \( t \)-statistics cannot be used to evaluate the statistical significance of the coefficients. The FMOLS technique overcomes this problem by modifying the OLS standard errors to take into account endogeneity as well as serial correlations. The EG and PO tests are then conducted on the residuals obtained from the estimated regression to determine if Equation (7.1) presents an equilibrium relationship.

7.4 Modelling Expectations

The next step in the analysis involves constructing a suitable proxy for oil price expectations of speculators. For this purpose the study uses the expectation formation mechanisms outlined in Section 6.3. It is assumed that forecasters update the weights assigned to the respective expectation formation mechanism in every period based on the most recent available information. Therefore, to obtain a forecast of the spot price of oil, the expectation formation mechanism included in the analysis (RE excluded) is estimated recursively over a portion of the sample period. Recursive estimation is preferred over rolling estimation for two important reasons. First, recursive estimation has been shown to be a more robust estimation technique for macroeconomic variables in the presence of structural changes (see Stock and Watson, 1996). Second, a number of studies suggest that inclusion of all available prior data provides gains in predictive performance.\(^{41}\)

Recursive estimation is carried out by estimating the model over part of the sample period \( t = 1, 2, \ldots, m \). Then, an out-of-sample forecast for the spot price is generated for the next \( n \)

\(^{41}\) For example, see MacDonald and Marsh (1993), Stock and Watson (2003), Pesaran et al. (2006) and Clark and McCracken (2009).
periods. The process is then repeated by estimating the model for sample period $t = 1, 2, ... , m + 1$ until the end of the sample period. Forecasts of the spot price for the 3, 6 and 9-month horizons are then constructed by averaging the monthly forecast values over these periods. For example, the average of the forecasts for April, May and June generated in March would constitute the forecast of the spot price for a 3-month horizon. One advantage of this procedure is that the monthly forecasts of the spot price are consistent with long-term forecasts (Baumeister and Kilian, 2013). Following the recursive estimation, a forecasting error is calculated as the difference between the actual spot price and the expected spot prices for $t + n$.

The initial estimation window comprises the first 100 observations of the sample, which covers the period January 1988–April 1996. However, as the data for OPEC spare capacity are only available from January 1994, the initial estimation window for the regressive expectations mechanism is based on the first 50 observations (January 1994–February 1998). A forecast for the oil price is then generated for the remaining period using each of the expectation formation mechanisms. It must be noted that because the forecast for oil price is obtained by averaging the monthly forecasts, the expected spot price sample is reduced by $n - 1$ where $n$ is the forecasting horizon. Once the expected spot price sample is constructed, the expected profit for a speculative strategy can be estimated as the difference between the expected spot price and the observed futures price at time $t$.

Once the expected spot price $E_t S_{t+n}$ and the corresponding actual spot price $S_{t+n}$ series are obtained, the forecasting performance of the expectations formation mechanisms can be

---

42 This choice of out-of-sample forecast is in line with Tashman (2000), Fildes and Makridakis (1995) and Moosa (2013).
43 The evaluation period is selected as it encompasses some of the most relevant events in the crude oil market history (e.g., the 2001 terrorist attacks in the US, the invasion of Iraq in 2003, the impact of the global financial crisis in 2008 and the European sovereign debt crisis of 2011).
evaluated for the $n$-month forecasting horizon, where $n = 1, 3, 6$ and $9$. Forecast evaluation is performed using the root mean squared error (RMSE), which is widely employed in the literature for evaluating predictive accuracy.\(^{44}\) The RMSE is the square root of the average forecasting error $(E_tS_{t+n} - S_{t+n})$ over the forecasting period:

$$RMSE = \sqrt{\frac{\sum_{t=m+k}^{n} (E_tS_{t+n} - S_{t+n})^2}{T - m - n - 1}}$$  \hspace{1cm} (7.2)

where $T$ represents the total number of observations. The RMSE of the random walk model is calculated as:

$$RMSE = \sqrt{\frac{\sum_{t=m+k}^{n} (S_{t+n} - S_t)^2}{T - m - n - 1}}$$  \hspace{1cm} (7.3)

The RMSE measures the variability of the forecasting errors expressed in terms of the original unit of measurement. The lower the value of RMSE, the more accurate is the respective expectation formation mechanism in predicting the magnitude of change in the underlying time series. The RMSE provides a quadratic loss function as the errors are squared and subsequently averaged. As a result, significantly more weight is given to large errors than small ones.

A comparison of the forecasting methods cannot be based purely on the numerical values of forecast errors. As a result, a formal test of forecast evaluation is employed to measure the statistical differences between the errors obtained from the random walk model and those from expectation formation mechanisms. The Ashley–Granger–Schmalancee (AGS) test, designed by Ashley et al. (1980), allows for measurement of the statistical significance of the

difference between the mean square errors of two competing models. The test involves estimating the following linear regression:

\[ D_t = a_0 + a_1(M_t - \bar{M}) + u_t \]  

(7.4)

where \( D_t = w_{1t} - w_{2t} \), \( M_t = w_{1t} - w_{2t} \). \( \bar{M} \) represents the mean of \( M_t \), \( w_{1t} \) is the forecasting error at time \( t \) of the model with the higher RMSE value and \( w_{2t} \) is the time \( t \) forecasting error of the model with the lower RMSE. The forecasting errors must be multiplied by \(-1\) if the sample mean is negative.

The estimated values of the regression parameters \( a_0 \) and \( a_1 \) in Equation (7.4) can be used to test whether the difference between the mean forecasting errors of the two models is statistically significant. If both \( a_1 \) and \( a_4 \) are positive, then a Wald test of the joint hypothesis \( H_0: a_0 = a_1 = 0 \) can be conducted. The test statistic has a chi-squared distribution with two degrees of freedom. If one of the estimates is significantly negative, the test is inconclusive. However, if one of estimates is only weakly negative (and not significant), then the test is conclusive. In this case, significance is determined on the basis of the upper tail of the distribution on the positive coefficient estimate.

The RMSE relies entirely on the magnitude of the prediction error and does not account for the ability of the forecasting method to predict the direction of change. Therefore, in addition to RMSE, a measure of directional accuracy is also employed to evaluate the forecasts. The directional accuracy metric helps determine the accuracy of the underlying expectation formation mechanism in predicting the direction of change (Cheung et al., 2005). Directional accuracy (DA) is determined by calculating the number of times a positive predicted change in the spot price corresponds to a positive realised change, and vice versa. It is calculated as:
\[ DA = \frac{\sum_{t=m+n}^{T} a_t}{T - m - n - 1} \]  

(7.5)

where

\[ a = \begin{cases} 1 & \text{if } \{(E_t S_{t+n} - S_t) (S_{t+n} - S_t) > 0 \} \\ 0 & \text{if } \{(E_t S_{t+n} - S_{t+n}) (S_{t+n} - S_t) < 0 \} \end{cases} \]  

(7.6)

DA of the forecast is expressed in terms of a percentage. A high value of DA indicates a more accurate forecast of the spot price.

The statistical significance of DA is tested using two null hypotheses—H1: DA=0 and H2: DA=0.5—against the alternative DA > 0 and DA > 0.5. The test statistic is given by:

\[ z = \frac{DA - \vartheta}{\sqrt{DA(1-DA)/T - m - n - 1}} \]  

(7.7)

One of the key advantages of this measure is that it is not sensitive to outliers. To test the hypothesis of zero DA, the parameter \( \vartheta \) is set to equal 0. In this way, the DA of the model can be tested against that of the random walk. Following Cheung et al. (2005), the parameter \( \vartheta \) is then set equal to 0.5 to determine if the model correctly predicts the direction of change on at least 50% of all occasions.

However, it must be noted that DA is not used as an alternative to measures based on the magnitude of errors. In the proposed speculative demand function, the oil price forecast is conducted to predict the profit of the speculative strategy. To predict the profit, speculators must predict not only the magnitude change, but also the direction of change. Hence, the DA measure must be used in conjunction with other forecast evaluation techniques to ascertain the accuracy of the forecasts.

A further test of DA is also conducted. This is the Pesaran and Timmermann (1992) test (PT test), which is a non-parametric test performed to determine the ability of the forecast to
predict the change in actual price. It is essentially a test of association between two series. The null hypothesis of the test is that the actual and the forecasted series are distributed independently. The test statistic, which has a chi-squared distribution with one degree of freedom, can be estimated as follows:

$$PT = \left( \frac{P_c(1 - P_e)}{n} \right)^{-\frac{1}{2}} (\hat{P} - P_e)$$  \hspace{1cm} (7.8)

where $\hat{P}$ represents the proportion of occasions on which the sign of the change is predicted correctly and $P_e$ is the probability that $(x_t, y_t) > 0$, where $x_t$ and $y_t$ represent the predicted and the actual change in spot price.

7.5 Equilibrium Price Estimation Results

The presentation of empirical results commences with unit root tests of the variables used for estimating the equilibrium oil price. Although the time series graphs suggest that each variable represents an $I(1)$ process, it is still necessary to undertake a formal test of unit root to identify non-stationary variables. For this purpose, unit root tests are performed on the natural logarithms of variables in their levels and first differences. To account for the long-term trends observed in the visual inspection of variables, the unit root tests are carried out by including an intercept component, and then an intercept as well as a trend component. The ADF test and the PP tests are employed to detect the presence of a unit root. The AIC and SBC are used to determine optimal lag length.
Table 7.3: Unit Root Test at Levels

<table>
<thead>
<tr>
<th></th>
<th>Augmented Dickey–Fuller</th>
<th>Phillips–Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td>$S$</td>
<td>-1.3257</td>
<td>-3.0268</td>
</tr>
<tr>
<td>$BDI$</td>
<td>-1.6556</td>
<td>-2.5837</td>
</tr>
<tr>
<td>$ERATE$</td>
<td>-1.2444</td>
<td>-1.944</td>
</tr>
<tr>
<td>$PROD$</td>
<td>-1.7583</td>
<td>-3.0625</td>
</tr>
<tr>
<td>$INV$</td>
<td>-1.779</td>
<td>-2.2856</td>
</tr>
<tr>
<td>$CAPACITY$</td>
<td>-2.447</td>
<td>-2.4943</td>
</tr>
</tbody>
</table>

The critical values at 95% and 90% are 2.88 and 2.58 with intercept and 3.43, 3.14 with intercept and trend. * indicates significance at the 10% level.

The results of unit root tests at the log level are reported in Table 7.3. The results are broadly consistent across the ADF and PP tests. With the exception of spare capacity, the null of a unit root cannot be rejected at the 5% level for the variables with or without the trend component. The unit root test for spare capacity leads to the rejection of the null hypothesis at the 10% level using the PP test when only an intercept is included. However, once the trend component is included, the null hypothesis cannot be rejected at the 5% level.

Table 7.4: Unit Root Test at First Difference

<table>
<thead>
<tr>
<th></th>
<th>Augmented Dickey–Fuller</th>
<th>Phillips–Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td>$S$</td>
<td>-12.0594**</td>
<td>-12.0328**</td>
</tr>
<tr>
<td>$BDI$</td>
<td>-12.8333**</td>
<td>-12.8311**</td>
</tr>
<tr>
<td>$ERATE$</td>
<td>-11.0383**</td>
<td>-11.0206**</td>
</tr>
</tbody>
</table>

The critical values at 95% and 90% are 2.88 and 2.58 with intercept and 3.43, 3.14 with intercept and trend. * and ** indicate significance at the 10 and 5% levels.

Table 7.4 presents the results of units root tests of the oil market variables in their first differences. The null hypothesis of non-stationarity is rejected at the 5% level for all variables with or without the trend component. This suggests that the variables reported in Table 7.4
are integrated at order 1, or $I(1)$. As a result, a long-run relationship can be estimated between these variables using the cointegration approach. For this purpose, the spot price of crude oil is expressed as a linear function of oil market fundamentals and macroeconomic indicators reported in the literature as having a significant influence on oil price.

The results of the equilibrium price estimation are presented in Table 7.5. The goodness of fit measures are also reported in the table. These include the coefficient of determination $R^2$ and the adjusted coefficient of determination Adj-$R^2$, which accounts for degrees of freedom. Model 1 displays the results are obtained from the estimation of Equation (7.1). With the exception of OPEC’s crude oil production, the effect of all variables is significant at the 10% level. The values for $R^2$ and Adj-$R^2$ are 0.94 and 0.93, suggesting that Model 1 explains a substantial portion of the total variation in oil prices. However, according to the EG and PO cointegration tests, the null hypothesis of a unit root in the error terms cannot be rejected at the 10% level, which implies there is no cointegration relationship between oil price and the explanatory variables in Equation (7.1).
Model 2 omits the variables that are not significant in Model 1. The EG and PO tests reject the null hypothesis of non-stationarity at 5% significance; hence confirming the presence of cointegration. Moreover, all the explanatory variables are significant at the 5% level. The coefficient for BDI is positive, indicating that dry cargo shipping rates, which proxy the level of economic activity, exert a significant and positive influence on oil price. A 1% increase in BDI is accompanied by an approximate 0.16% increase in oil price. This finding is in line with the results obtained in previous studies investigating the relationship between commodity freight rates and oil prices. For example, using the Kilian Economic Index, He at
al. (2010) find the dry cargo rates to be closely related to the real crude oil prices. Similarly, Tamakoshi and Hamori (2012) report a strong positive relationship between oil prices and economic activity, as measured by the Kilian Economic Index.

The coefficient for the trade-weighted US dollar index is negatively related to oil prices, consistent with the results obtained by He et al. (2010), who find that a 1% increase in this index is in line with a decrease of 0.89% in the long-run price of oil. Consistent with the theory of storage, a strong negative relationship is observed between the crude oil inventory and WTI oil price. According to the results, an increase of 1% in crude oil inventory causes the oil price to decline by 3.6%. Finally, OPEC’s spare capacity, which serves as an excess supply indicator, is also negatively related to the oil price. A 1% increase in spare capacity leads to a decrease in oil price of approximately 0.05%. Similarly, Prat and Uctum (2011) find that on average the oil price increases by 0.08% for a 1% increase in spare capacity. They also find that the impact of spare capacity is greater (0.15%) during the period 2004–08.

The residuals obtained from the regression are used to proxy the mean reversion component in the regressive expectations hypothesis.

### 7.6 Oil Price Forecast Evaluation

For the purposes of evaluating oil price forecasts, the study takes the view of a speculator who aims to earn a positive return by successfully predicting the spot price of oil prevailing at maturity of the futures contract. In this case, the forecast loss function of the speculator must take into account the ability of the expectation formation mechanism to accurately predict not only the direction but also the magnitude of the change in the spot price. The approach taken therefore involves evaluating predictive accuracy on the basis of the magnitude of the

---

45 The Kilian Economic Index is constructed by deflating growth rates of freight charges for bulk dry cargoes by the US consumer price index; see Kilian (2009) for more details.
forecast errors as well as the DA of the forecast. The forecast error magnitude is evaluated using the RMSE measure.

The oil price forecasts are evaluated over the period May 1996–April 2013 with the exception of the regressive expectation formation mechanism, which is evaluated over the period March 1998–April 2013. Using the forecast evaluation techniques described earlier, an analysis is presented of the forecasting performance of extrapolative, adaptive and regressive expectation mechanisms, and the benchmark random walk is presented. The analysis commences with a visual comparison of forecasts with each of the expectation formation mechanisms with that of the random walk model.
Figure 7.6: Oil Price Forecasts—1 and 3-month Horizons
The graphs for the 1 and 3-month-ahead forecasts are shown in Figure 7.6. Over the 1-month horizon, the forecast of oil price series is indistinguishable from the actual price in the case of all the expectation formation mechanisms. However, over the 3-month horizon, the forecasted series clearly deviates from the actual price. On the basis of these graphs, it can be inferred that neither the random walk nor the other expectation formation mechanisms can predict the sign of change in oil price and follow the actual price series in a lagged manner. It is also evident that the performance of all the expectations formation mechanisms is identical to the random walk, which indicates that no additional gains can be achieved by predicting the spot price of oil over 1 and 3-month periods using these mechanisms.

Figure 7.7 shows the graphs of the actual price as well as the 6 and 9-month-ahead forecasts of the spot price. The magnitude of deviation of the forecast series from the actual price increases as the forecasting horizon lengthens. Over longer horizons the forecast accuracy of the adaptive, extrapolative and regressive expectations mechanisms is marginally higher during periods of high volatility. However, this may be due to deterioration in the accuracy of the random walk—as noted in past studies—rather than improvement in the accuracy of the other expectation formation mechanisms over long horizons.\(^{46}\) Similar to Figure 7.6, the forecasting performance of all the expectation formation mechanisms over the same forecasting horizon is identical.

Figure 7.7: Oil Price Forecasts—6 and 9-month Horizons
Table 7.6 reports the RMSEs for the extrapolative expectation mechanism and the random walk, along with the results of the AGS test. The RMSEs for the extrapolative expectation mechanism are numerically smaller than those for the random walk method for all forecast horizons. In line with the visual analysis of the graphs, the RMSEs increase with the length of the forecasting horizon. The RMSEs for the extrapolative expectation mechanism increase by a smaller amount over the forecasting horizons than do those of the random walk.

According to the AGS test results, the intercept term $\alpha_0$ is significantly positive for all forecasting horizons, whereas the slope $\alpha_1$ is not significantly different from 0, with the exception of the 9-month forecast horizon where it is significantly positive at the 5% level. The Wald chi-square test statistics indicate that the null hypothesis $\alpha_0 = \alpha_1 = 0$ can be rejected at the 5% level, which means that the extrapolative expectation generates a statistically smaller forecasting error than the random walk.

<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE (extrapolative expectation)</td>
<td>2.77</td>
<td>5.69</td>
<td>8.71</td>
<td>10.51</td>
</tr>
<tr>
<td>RMSE (random walk)</td>
<td>2.73</td>
<td>4.92</td>
<td>7.73</td>
<td>10.86</td>
</tr>
<tr>
<td>$t(\alpha_0)$</td>
<td>(0.339)**</td>
<td>(0.604)**</td>
<td>(0.983)**</td>
<td>(1.381)**</td>
</tr>
<tr>
<td>$t(\alpha_1)$</td>
<td>0.0248</td>
<td>0.022</td>
<td>0.028</td>
<td>(0.0362)**</td>
</tr>
<tr>
<td>$\alpha_0 = \alpha_1 = 0$</td>
<td>(26.45)**</td>
<td>(23.23)**</td>
<td>(26.23)**</td>
<td>(35.49)**</td>
</tr>
</tbody>
</table>

Note: $t(\alpha_0)$ and $t(\alpha_1)$ represent the $t$-values for the intercept ($\alpha_0$) and slope ($\alpha_1$) estimates. The test statistic follows a $\chi^2(2)$ distribution; * and ** indicate significance at the 10 and 5% levels.

The RMSE and AGS test results for the adaptive expectation formation method are displayed in Table 7.7. Once more, the RMSE confirms the findings of the graphical analysis in that the forecasting accuracy deteriorates as the forecasting horizon lengthens. In contrast to extrapolative expectations, the adaption expectations mechanism produces higher RMSEs than does the random walk model without drift.
The intercept parameter $a_0$ of the AGS test regression is significantly positive at 5% across all forecasting horizons. The slope parameter $a_1$, on the other hand, is not significant across all horizons. With the exception of the 1-month horizon, the Wald chi-square test statistic shows that the null hypothesis is rejected at the 5% level in each case. This suggests that the random walk method with drift is not superior to the adaptive expectations model at 1-month horizon, whereas over the 3, 6 and 9-month horizons the random walk model outperforms the adaptive expectations model.

Table 7.7: Root Mean Squared Error of Adaptive Expectations

<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE (adaptive expectation)</td>
<td>2.80</td>
<td>5.84</td>
<td>9.16</td>
<td>11.07</td>
</tr>
<tr>
<td>RMSE (random walk)</td>
<td>2.73</td>
<td>4.92</td>
<td>7.73</td>
<td>10.86</td>
</tr>
<tr>
<td>$t(a_0)$</td>
<td>(0.218)</td>
<td>(0.658)</td>
<td>(1.3)</td>
<td>(1.858)</td>
</tr>
<tr>
<td>$t(a_1)$</td>
<td>−0.027</td>
<td>−0.013</td>
<td>−0.006</td>
<td>−0.007</td>
</tr>
<tr>
<td>$a_0 = a_1 = 0$</td>
<td>4.28</td>
<td>(21.73)**</td>
<td>(39.65)**</td>
<td>(46.91)**</td>
</tr>
</tbody>
</table>

Note: $t(a_0)$ and $t(a_1)$ represent the $t$-values for the intercept ($a_0$) and slope ($a_1$) estimates. The test statistic follows a $\chi^2(2)$ distribution; * and ** indicate significance at the 10 and 5% levels.

The RMSE and AGS test for forecast equivalence between regressive expectations and random walk are summarised in Table 7.8. The RMSE for regressive expectations is numerically higher than the RMSE for the random walk over all forecasting horizons. The RMSE for the regressive expectations mechanism is also higher than that of the extrapolative and adaptive formation mechanisms. It must be noted that in this framework, the speculative demand for futures contract is expressed as a positive function of the expected profit ($F_{t,t+n} - E_S_{t+n}$) obtained on the speculative position. A higher forecasting error increases the uncertainty associated with the expected profit. As a result, an expectation formation mechanism that generates a higher RMSE may lead to a smaller impact of speculative demand on crude oil futures prices for that mechanism.
<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE (regressive expectation)</td>
<td>3.10</td>
<td>5.56</td>
<td>10.67</td>
<td>13.25</td>
</tr>
<tr>
<td>RMSE (random walk)</td>
<td>2.73</td>
<td>4.92</td>
<td>7.73</td>
<td>10.86</td>
</tr>
<tr>
<td>$t(a_0)$</td>
<td>(0.225)*</td>
<td>(0.693)**</td>
<td>(1.398)**</td>
<td>(2.023)**</td>
</tr>
<tr>
<td>$t(a_1)$</td>
<td>-0.028</td>
<td>-0.013</td>
<td>0.007</td>
<td>(0.008)*</td>
</tr>
<tr>
<td>$a_0 = a_1 = 0$</td>
<td>(22.68)**</td>
<td>3.87</td>
<td>(39.33)**</td>
<td>(49.94)**</td>
</tr>
</tbody>
</table>

Note: $t(a_0)$ and $t(a_1)$ represent the $t$-values for the intercept ($a_0$) and slope ($a_1$) estimates. The test statistic follows a $\chi^2(2)$ distribution; * and ** indicate significance at the 10 and 5% levels.

The AGS test results confirm that the random walk model outperforms the regressive expectation formation mechanism over the 1, 6 and 9-month horizons. Over the 3-month horizon, the null hypothesis of equivalence of the two RMSEs cannot be rejected, which indicates that the random walk does not outperform the regressive expectations method over the 3-month horizon. This result highlights the importance of conducting a statistical test of forecast equivalence, rather than a mere comparison of the numerical values of RMSEs for the two models for evaluating predictive accuracy.

The forecast of crude oil is also evaluated on the basis of DA, which is measured by calculating the number of times, as a proportion of total observations, in which the forecasting model correctly predicts the direction of change in the price of oil.
Table 7.9: Directional Accuracy of the Forecasts

<table>
<thead>
<tr>
<th></th>
<th>t + 1</th>
<th>t + 3</th>
<th>t + 6</th>
<th>t + 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extrapolative expectation</td>
<td>0.57</td>
<td>0.74</td>
<td>0.48</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>(11.6)**</td>
<td>(15.5)**</td>
<td>(9.63)**</td>
<td>(9.11)**</td>
</tr>
<tr>
<td>Adaptive expectation</td>
<td>0.49</td>
<td>0.48</td>
<td>0.42</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(9.93)**</td>
<td>(9.56)**</td>
<td>(8.48)**</td>
<td>(8.25)**</td>
</tr>
<tr>
<td>Regressive expectation</td>
<td>0.79</td>
<td>0.59</td>
<td>0.55</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(16.6)**</td>
<td>(11.3)**</td>
<td>(10.4)**</td>
<td>(10.6)**</td>
</tr>
</tbody>
</table>

H0: DA=0.5

<table>
<thead>
<tr>
<th></th>
<th>t + 1</th>
<th>t + 3</th>
<th>t + 6</th>
<th>t + 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extrapolative expectation</td>
<td>0.57</td>
<td>0.74</td>
<td>0.48</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>(1.60)</td>
<td>(5.18)**</td>
<td>(-0.2)</td>
<td>(-0.6)</td>
</tr>
<tr>
<td>Adaptive expectation</td>
<td>0.49</td>
<td>0.48</td>
<td>0.42</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(-0.0)</td>
<td>(-0.2)</td>
<td>(-1.4)</td>
<td>(-1.5)</td>
</tr>
<tr>
<td>Regressive expectation</td>
<td>0.79</td>
<td>0.59</td>
<td>0.55</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(6.14)**</td>
<td>(1.82)*</td>
<td>(1.02)</td>
<td>(1.30)</td>
</tr>
</tbody>
</table>

Note: t-statistics are reported in parentheses; * and ** indicate significance at the 10 and 5% levels.

The results of the DA are presented in Table 7.9, which shows that DA deteriorates as the forecasting horizon becomes larger. There are a few exceptions; for example, in the case of extrapolative expectations where DA increases at the 3 and 9-month horizons. The null hypothesis of zero DA is rejected at the 5% level in all cases.

An additional test is also conducted to determine if the model correctly predicts at least 50% of all changes in the direction of oil price. The results of the test are presented in Table 7.9. On most occasions, the forecasting models fail to correctly predict the sign of change in 50% or more of the cases. For extrapolative expectations, DA exceeds 50% at the 3-month horizon, whereas in the case of the regressive expectation formation mechanism the DA exceeds 50% at the 1 and 3-month horizons.

In general, the forecasting performance in terms of error magnitude is consistent with the forecasting performance in terms of DA. A notable exception occurs in the case of the
regressive expectation mechanism, which performs relatively poorly in terms of RMSEs but outperforms the other models in predicting the direction of change at the 1-month horizon.

Table 7.10: Directional Accuracy of the Forecasts using the Pesaran–Timmermann (PT) Test

<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extrapolative expectation</td>
<td>0.57</td>
<td>0.74</td>
<td>0.48</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>(5.76)**</td>
<td>(6.17)**</td>
<td>(2.75)*</td>
<td>(1.15)</td>
</tr>
<tr>
<td>Adaptive expectation</td>
<td>0.49</td>
<td>0.48</td>
<td>0.42</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(7.73)**</td>
<td>(10.1)**</td>
<td>(11.8)**</td>
<td>(12.8)**</td>
</tr>
<tr>
<td>Regressive expectation</td>
<td>0.79</td>
<td>0.59</td>
<td>0.55</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(8.48)**</td>
<td>(3.30)*</td>
<td>(2.27)</td>
<td>(1.82)</td>
</tr>
</tbody>
</table>

Note: The PT test statistic follows a $X^2(2)$ distribution and has critical values of 3.84 and 2.71 for significance at the 5(**) and 10%(*) levels.

According to the PT test (Table 7.10), DA increases with forecast horizon in most cases. Further, no predictable relationship is observed between the predicted and realised change (the null hypothesis cannot be rejected) at the 9-month horizon for extrapolative expectations, and at the 6 and 9-month horizons for regressive expectations. In all other cases, the null hypothesis is rejected, which implies that predicted and realised changes are not independent. The results support the findings of the DA test discussed earlier in suggesting that DA deteriorates at long horizons.

In an overall sense, the findings of the forecast evaluation suggest that the extrapolative expectation mechanism provides the most accurate forecast of crude oil price—even outperforming the random walk. This result is consistent with the study of Hendry (2006) who observes that the forecasting performance of the random walk can be improved upon by extrapolating the current price at the most recent growth rates.

However, the results do contrast with the findings of MacDonald and Marsh (1993), who evaluate expectation generation mechanisms in the G7 countries. They find adaptive and
static expectations to be most successful in explaining oil price behaviour for the majority of countries. However, note that MacDonald and Marsh use survey data as a proxy for oil price expectations. Moreover, their analysis is conducted over the period 1989–91, which entails different oil price dynamics than occurred in the sample period used in this study.

Forecast evaluation also shows that the random walk model fails to outperform regressive expectations method at the 3-month horizon. This finding is slightly surprising as the random walk tends to be more accurate at short horizons relative to approaches that are market fundamentals based (see, e.g., Baumeister and Kilian 2012, 2013; Baumeister et al., 2013). However Reitz et al. (2009) assert that agents may find the predictive accuracy of the random walk relative to the mean reverting or regressive expectations to be higher only when the actual price lies within a ‘reasonable range’ around the fundamental value. They find that oil price forecasts exhibit greater reliance on the mean reverting expectations model in 2005 when oil prices rise sharply. To determine if the higher accuracy of regressive expectation at the 3-month horizon is due to sharp movements in oil prices, a visual inspection of the forecasting errors is conducted.
Figure 7.8 is plot of recursive RMSEs for all forecasting methods over the forecast evaluation period. The predictive accuracy of the expectation formation processes varies considerably over the sample period. For the 1-month horizon, the RMSEs for all of the expectation mechanisms are relatively similar over the late 1990s and the early part of the 2000s when oil prices were relatively stable. However, these similarities are to be expected as all the models include the lagged value of the spot price as an independent variable. After 2003, the forecast accuracy of the regressive expectations method decreases in comparison with that of the
random walk and adaptive expectations. The gap between the RMSEs for regressive expectations and the random walk becomes more apparent after the oil price shock of 2008.

The graphical analysis of recursive RMSEs does lend some support to the findings of Reitz et al. (2012). The predictive accuracy of regressive expectations tends to rise when sharp movements in oil prices are observed. However, the evidence is not conclusive as this phenomenon is only observed at the 3-month horizon. Moreover, the graphs for the RMSEs of the 1, 6 and 9-month forecasts indicate significant deterioration in the accuracy of regressive expectations relative to other mechanisms.

7.7 Cointegration Results

Before testing any long-run relationship between the forecast spot price and the futures price, their order of integration must be ascertained. For this purpose, unit root tests are performed on the expected spot prices generated using the adaptive, extrapolative and regressive expectation mechanisms in their log levels and in their first differences. Unit root tests are also performed on the revised samples of observed futures price and the theoretical futures price. To account for the long-term trends observed in the visual inspection of oil market variables, the unit root tests are carried out by including an intercept component, and then an intercept plus trend component. The ADF and PP test are employed to determine if the variables are stationary at log levels.
Table 7.11: Unit Root Test for Expected Spot Price and Futures Price at Log Levels

<table>
<thead>
<tr>
<th>Augmented Dickey– Fuller</th>
<th>Phillips– Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td>with intercept</td>
<td>with trend and</td>
</tr>
<tr>
<td></td>
<td>intercept</td>
</tr>
<tr>
<td>Extrapolative expectations</td>
<td></td>
</tr>
<tr>
<td>$E_t S_{t+1}$</td>
<td>-1.1157</td>
</tr>
<tr>
<td></td>
<td>-3.1635</td>
</tr>
<tr>
<td>$E_t S_{t+3}$</td>
<td>-1.1204</td>
</tr>
<tr>
<td></td>
<td>-3.0767</td>
</tr>
<tr>
<td>$E_t S_{t+6}$</td>
<td>-1.1108</td>
</tr>
<tr>
<td></td>
<td>-3.0330</td>
</tr>
<tr>
<td>$E_t S_{t+9}$</td>
<td>-0.9899</td>
</tr>
<tr>
<td></td>
<td>-3.0193</td>
</tr>
<tr>
<td>Adaptive expectations</td>
<td></td>
</tr>
<tr>
<td>$E_t S_{t+1}$</td>
<td>-1.5409</td>
</tr>
<tr>
<td></td>
<td>-3.0744</td>
</tr>
<tr>
<td>$E_t S_{t+3}$</td>
<td>-1.4890</td>
</tr>
<tr>
<td></td>
<td>-3.1141</td>
</tr>
<tr>
<td>$E_t S_{t+6}$</td>
<td>-1.4918</td>
</tr>
<tr>
<td></td>
<td>-3.0689</td>
</tr>
<tr>
<td>$E_t S_{t+9}$</td>
<td>-1.4402</td>
</tr>
<tr>
<td></td>
<td>-3.2463*</td>
</tr>
<tr>
<td>Regressive expectations</td>
<td></td>
</tr>
<tr>
<td>$E_t S_{t+1}$</td>
<td>-1.5156</td>
</tr>
<tr>
<td></td>
<td>-3.0727</td>
</tr>
<tr>
<td>$E_t S_{t+3}$</td>
<td>-1.4891</td>
</tr>
<tr>
<td></td>
<td>-3.2202</td>
</tr>
<tr>
<td>$E_t S_{t+6}$</td>
<td>-1.4666</td>
</tr>
<tr>
<td></td>
<td>-3.0998</td>
</tr>
<tr>
<td>$E_t S_{t+9}$</td>
<td>-1.3737</td>
</tr>
<tr>
<td></td>
<td>-3.0151</td>
</tr>
<tr>
<td>Futures price (May 1996–April 2013)</td>
<td></td>
</tr>
<tr>
<td>$F_{t,t+1}$</td>
<td>-1.2094</td>
</tr>
<tr>
<td></td>
<td>-2.883</td>
</tr>
<tr>
<td>$F_{t,t+3}$</td>
<td>-1.2852</td>
</tr>
<tr>
<td></td>
<td>-2.9239</td>
</tr>
<tr>
<td>$F_{t,t+6}$</td>
<td>-1.2725</td>
</tr>
<tr>
<td></td>
<td>-2.9917</td>
</tr>
<tr>
<td>$F_{t,t+9}$</td>
<td>-1.3507</td>
</tr>
<tr>
<td></td>
<td>-3.0692</td>
</tr>
<tr>
<td>Futures price (April 1998–April 2013)</td>
<td></td>
</tr>
<tr>
<td>$F_{t,t+1}$</td>
<td>-1.2216</td>
</tr>
<tr>
<td></td>
<td>-2.8952</td>
</tr>
<tr>
<td>$F_{t,t+3}$</td>
<td>-1.2954</td>
</tr>
<tr>
<td></td>
<td>-2.9361</td>
</tr>
<tr>
<td>$F_{t,t+6}$</td>
<td>-1.2887</td>
</tr>
<tr>
<td></td>
<td>-3.0039</td>
</tr>
<tr>
<td>$F_{t,t+9}$</td>
<td>-1.3669</td>
</tr>
<tr>
<td></td>
<td>-3.0814</td>
</tr>
<tr>
<td>Theoretical futures price (May 1996–April 2013)</td>
<td></td>
</tr>
<tr>
<td>$\bar{F}_{t,t+1}$</td>
<td>-1.6473</td>
</tr>
<tr>
<td></td>
<td>-3.0100</td>
</tr>
<tr>
<td>$\bar{F}_{t,t+3}$</td>
<td>-1.1260</td>
</tr>
<tr>
<td></td>
<td>-3.0537</td>
</tr>
<tr>
<td>$\bar{F}_{t,t+6}$</td>
<td>-1.2407</td>
</tr>
<tr>
<td></td>
<td>-2.9116</td>
</tr>
<tr>
<td>$\bar{F}_{t,t+9}$</td>
<td>-1.3040</td>
</tr>
<tr>
<td></td>
<td>-2.9710</td>
</tr>
</tbody>
</table>
The null hypothesis of a unit root cannot be rejected even at the 10% level for any variable with or without a trend. On a few occasions, the null hypothesis is rejected at the 10% level using the ADF test when the trend component is included. At the same time, the PP test cannot reject the null hypothesis at the 10% level. It must be noted that the ADF test can lead to false rejection of the null hypothesis in the presence of a structural break (Perron, 1989). It is therefore possible that the structural change in the crude oil price, particularly the sharp decline in 2008, may have influenced the results of the test. The unit root tests are conducted on the first differences of the variables and the results are presented in Table 7.12.

The null hypothesis of non-stationarity is rejected at the 5% level for all variables with and without a trend. The results reported in Table 7.12 show that variables are integrated of order 1 or $I(1)$. As a result, a long-run relationship can be estimated between these variables using the cointegration tests outlined in Section 5.4.2.
Table 7.12: Unit Root Test for Expected Spot Price and Futures Price at First Difference

<table>
<thead>
<tr>
<th></th>
<th>Augmented Dickey–Fuller</th>
<th>Phillips–Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and intercept</td>
</tr>
<tr>
<td>Extrapolative expectations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_t S_{t+1}$</td>
<td>$-10.6099^{**}$</td>
<td>$-10.5916^{**}$</td>
</tr>
<tr>
<td>$E_t S_{t+3}$</td>
<td>$-10.4590^{**}$</td>
<td>$-10.4391^{**}$</td>
</tr>
<tr>
<td>$E_t S_{t+6}$</td>
<td>$-10.2481^{**}$</td>
<td>$-10.2273^{**}$</td>
</tr>
<tr>
<td>$E_t S_{t+9}$</td>
<td>$-10.0967^{**}$</td>
<td>$-10.0746^{**}$</td>
</tr>
<tr>
<td>Adaptive expectations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_t S_{t+1}$</td>
<td>$-8.56767^{**}$</td>
<td>$-8.55199^{**}$</td>
</tr>
<tr>
<td>$E_t S_{t+3}$</td>
<td>$-15.6104^{**}$</td>
<td>$-15.5753^{**}$</td>
</tr>
<tr>
<td>$E_t S_{t+6}$</td>
<td>$-15.4077^{**}$</td>
<td>$-15.3757^{**}$</td>
</tr>
<tr>
<td>$E_t S_{t+9}$</td>
<td>$-15.0390^{**}$</td>
<td>$-15.0042^{**}$</td>
</tr>
<tr>
<td>Regressive expectations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Futures price (May 1996–April 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_{t+6}$</td>
<td>$-12.9228^{**}$</td>
<td>$-12.8922^{**}$</td>
</tr>
<tr>
<td>$F_{t+9}$</td>
<td>$-12.6568^{**}$</td>
<td>$-12.6278^{**}$</td>
</tr>
<tr>
<td>Futures price (April 1998–April 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_{t+1}$</td>
<td>$-14.2588^{**}$</td>
<td>$-14.2241^{**}$</td>
</tr>
<tr>
<td>$F_{t+3}$</td>
<td>$-13.5180^{**}$</td>
<td>$-13.4847^{**}$</td>
</tr>
<tr>
<td>$F_{t+6}$</td>
<td>$-12.9863^{**}$</td>
<td>$-12.9551^{**}$</td>
</tr>
<tr>
<td>$F_{t+9}$</td>
<td>$-12.7115^{**}$</td>
<td>$-12.6817^{**}$</td>
</tr>
<tr>
<td>Theoretical futures price (May 1996–April 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{F}_{t+1}$</td>
<td>$-13.9393^{**}$</td>
<td>$-13.9069^{**}$</td>
</tr>
<tr>
<td>$\bar{F}_{t+3}$</td>
<td>$-11.3276^{**}$</td>
<td>$-11.3023^{**}$</td>
</tr>
<tr>
<td>$\bar{F}_{t+9}$</td>
<td>$-13.5609^{**}$</td>
<td>$-13.5288^{**}$</td>
</tr>
</tbody>
</table>
### Table 7.12 (Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept without trend and intercept</td>
<td>with intercept without trend and intercept</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+1} )</td>
<td>(-13.0118^{<strong>} ) (-12.9901^{</strong>} )</td>
<td>(-13.0107^{<strong>} ) (-12.9889^{</strong>} )</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+3} )</td>
<td>(-10.6252^{<strong>} ) (-10.6106^{</strong>} )</td>
<td>(-10.6089^{<strong>} ) (-10.6338^{</strong>} )</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+6} )</td>
<td>(-12.8299^{<strong>} ) (-12.8071^{</strong>} )</td>
<td>(-12.8299^{<strong>} ) (-12.8071^{</strong>} )</td>
</tr>
<tr>
<td>( \bar{F}_{t,t+9} )</td>
<td>(-12.7490^{<strong>} ) (-12.7246^{</strong>} )</td>
<td>(-12.7478^{<strong>} ) (-12.7232^{</strong>} )</td>
</tr>
</tbody>
</table>

Note: All variables are expressed as natural logs. * and ** indicate significance at the 10 and 5% levels.

Cointegration testing involves investigating the existence of a long-run relationship between the observed crude oil futures price, the cost of carry price and the expected spot price of oil. Several proxies for the expected spot price are constructed using the expectation formation mechanisms described earlier. To determine how variation in the expectations of speculators impacts the prices of crude oil futures, a cointegration regression is estimated for each of the proxies for expected spot price.

#### 7.7.1 Extrapolative Expectations

The estimated cointegration regression is presented in Table 7.13. In this case, the expected spot price is estimated using the extrapolative expectation formation mechanism. Based on the EG test statistic, the null hypothesis of no cointegration is rejected at the 5% level for 1 and 3-month maturities, whereas for 6 and 9-month maturities, the null hypothesis cannot be rejected. On the other hand, the PO test statistics provide significant evidence of cointegration at all maturities. The coefficients \( \beta_1 \) and \( \beta_2 \) are significant for all maturities, indicating that both arbitrage and speculation have significant impacts on futures price. The coefficient restriction \( \beta_0 = 0 \) cannot be rejected for all maturities. Similar to the RE method, the restrictions \( \beta_1 = 0 \) and \( \beta_2 = 0 \) are rejected at the 5% level whereas the restriction \( \beta_1 + \beta_2 = 1 \) is rejected for all with the exception of the 1-month maturity. This indicates that, in
addition to arbitrage and speculation, other variables may be important for the determination of futures price. However, the main finding emerging from these results is that the extrapolative expectation formation mechanism does not rule out the possibility of speculative pressures on crude oil futures price. The impact of speculative demand on futures price increases as the maturity of futures contract becomes longer. In other words, the sensitivity of futures prices to changes in the expected spot price increases at longer maturities. As previously observed, the forecasting errors (as measured by the RMSE) increase with the forecast horizon. Therefore, this finding supports our earlier assertion that the impact of futures demand by speculators (which is positively related to the expected profit) is positively associated with the predictive accuracy of the underlying mechanism. However, Leitch and Tanner (1991) argue that measures of predictive accuracy that rely on error magnitude criteria bear little relationship to the profitability of forecasts. This means that the magnitude of the forecast error is an unreliable indicator of expected profits.

To investigate these results further, the expected profit obtained by speculators over 1, 3, 6 and 9-month periods, using the extrapolative expectation formation mechanism, is examined. The expected profit is calculated as the absolute value of difference between the observed futures price at time $t$ and the expected spot price at the expiry of the futures contract. Table 7.14 outlines the expected profit expressed as a percentage of the futures price over different maturity periods.
Table 7.13: Cointegration Regression—Extrapolative Expectations

<table>
<thead>
<tr>
<th></th>
<th>$F_{t,t+n}$ $(n=1)$</th>
<th>$F_{t,t+n}$ $(n=3)$</th>
<th>$F_{t,t+n}$ $(n=6)$</th>
<th>$F_{t,t+n}$ $(n=9)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>-0.0057</td>
<td>-0.0059</td>
<td>-0.0626</td>
<td>-0.0892</td>
</tr>
<tr>
<td></td>
<td>(-0.26)</td>
<td>(-0.14)</td>
<td>(-1.03)</td>
<td>(-1.15)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9259</td>
<td>0.7432</td>
<td>0.7477</td>
<td>0.5994</td>
</tr>
<tr>
<td></td>
<td>(34.36)**</td>
<td>(15.00)**</td>
<td>(9.55)**</td>
<td>(6.40)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.055</td>
<td>0.2053</td>
<td>0.1894</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>(1.92)*</td>
<td>(3.78)**</td>
<td>(2.23)**</td>
<td>(3.00)**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0003</td>
<td>0.0008</td>
<td>0.0012</td>
<td>0.0017</td>
</tr>
<tr>
<td></td>
<td>(2.39)**</td>
<td>(3.18)**</td>
<td>(3.20)**</td>
<td>(3.47)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9973</td>
<td>0.994</td>
<td>0.9896</td>
<td>0.9844</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.9973</td>
<td>0.9939</td>
<td>0.9895</td>
<td>0.9842</td>
</tr>
<tr>
<td>$t$</td>
<td>-10.66**</td>
<td>-7.16**</td>
<td>-3.36</td>
<td>-3.02</td>
</tr>
<tr>
<td>$z$</td>
<td>-145.64**</td>
<td>-81.36**</td>
<td>-25.2</td>
<td>-19.61</td>
</tr>
<tr>
<td>$\hat{z}_x$</td>
<td>-10.81**</td>
<td>-7.17**</td>
<td>-5.76**</td>
<td>-4.93**</td>
</tr>
<tr>
<td>$\hat{z}_t$</td>
<td>-153.33**</td>
<td>-81.05**</td>
<td>-55.40**</td>
<td>-41.57**</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>-0.26</td>
<td>-0.14</td>
<td>-1.03</td>
<td>-1.15</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>-2.75**</td>
<td>-5.18**</td>
<td>-3.22**</td>
<td>-4.28**</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>1.92*</td>
<td>3.78**</td>
<td>2.23**</td>
<td>3.00**</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>-1.55</td>
<td>-2.12**</td>
<td>-1.83*</td>
<td>-2.02**</td>
</tr>
<tr>
<td>$N$</td>
<td>203</td>
<td>201</td>
<td>198</td>
<td>195</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

Table 7.14: Summary Statistics for Expected Profit—Extrapolative Expectations

<table>
<thead>
<tr>
<th></th>
<th>$EP_{t,t+n}$ $(n=1)$</th>
<th>$EP_{t,t+n}$ $(n=3)$</th>
<th>$EP_{t,t+n}$ $(n=6)$</th>
<th>$EP_{t,t+n}$ $(n=9)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.021**</td>
<td>0.072**</td>
<td>0.093**</td>
<td>0.119**</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.068</td>
<td>0.073</td>
<td>0.080</td>
<td>0.088</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

The average expected profit is statistically significant over all maturities. However, it must be noted that the expected profit and its statistical significance vary strongly over the sample period. Also, the variability of the expected profit increases with the maturity period. Figure
7.9 shows the expected profit rise in 2007 and 2008—a period of financial market turbulence that had a dramatic impact on crude oil price forecasts.

**Figure 7.9: Expected Profit—Extrapolative Expectations**

Based on these findings, it may be suggested that the increase in the impact of speculative demand on futures prices at longer maturities is simply due to the higher profitability of speculation at those maturity periods. Moosa and Shamsuddin (2004) argue that the actions of traders and the consequent impact of their actions on the price should, in principle, be determined by the underlying profitability of the trading strategy. A similar view is expressed by Hodgson (1985, p.13) who states that “actions flow from judgments about the future, which often lack a firm objective empirical foundation as well as from observation of convention that is formed by the action of others”.

On the other hand, the higher sensitivity of futures price speculation at longer maturities could be attributed to the risk premium embedded in futures price. For instance, Bessembinder (1993) notes that the changes in the futures price represent a linear function of the futures’ systematic risk and the risk premiums arising from future uncertainty. Based on
the term structure of futures prices, Bhar and Lee (2011) show that speculators receive a risk premium that increases with contract maturity. Gorton et al. (2013), confirm that changes in the long-term futures prices incorporate an ex-ante compensation for the risk arising from concerns regarding the futures state of crude inventories.

Figure 7.10: Futures Price Differentials

Figure 7.10 displays price differentials of the 9 and 1-month futures, and the 3 and 1-month futures. These differentials suggest that between 1999 and 2005, the crude oil market exhibited strong backwardation, because the prices of the 9 and 3-month futures contracts are well below the 1-month futures price. This suggests that speculators received a risk premium as an insurance against future price decline. Moreover, the price differential between the 9 and 1-month futures contracts is higher in comparison to the price differential between the 3 and 1-month futures prices. This provides support for the argument that the higher sensitivity of long-term futures to speculative activity may be due to the high risk premium.
7.7.2 Adaptive Expectations

Table 7.15 presents the results from the cointegration regression involving the expected spot price estimated using adaptive expectations. The EG and PO test statistics lead to rejection of the null hypothesis of no cointegration at the 5% level for all maturities. The coefficient for the expected spot price is not significantly different from 0 for the 1-month maturity. The coefficient for the theoretical futures price and the expected spot price are significant at the 5% level. The Wald test results of the coefficient restrictions are also presented. The restriction $\beta_0 = 0$ cannot be rejected, whereas restrictions $\beta_1 = 1$ and $\beta_2 = 0$ are rejected in all cases. Further, with the exception of the 3-month futures contract, the restriction $\beta_1 + \beta_2 = 1$ cannot be rejected. This result differs from the one obtained using extrapolative formations, where the restriction $\beta_1 + \beta_2 = 1$ is rejected in all cases with the exception of the 1-month contract.

The expected profits generated by adaptive expectations are presented in Table 7.16. The table shows that the profits generated by the adaptive expectations are statistically significant and smaller in comparison to the profit obtained by the extrapolative expectations mechanism.
Table 7.15: Cointegration Regression—Adaptive Expectations

<table>
<thead>
<tr>
<th></th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=1)</td>
<td>(n=3)</td>
<td>(n=6)</td>
<td>(n=9)</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>-0.0067</td>
<td>-0.0103</td>
<td>-0.0619</td>
<td>-0.0977</td>
</tr>
<tr>
<td></td>
<td>(-0.31)</td>
<td>(-0.24)</td>
<td>(-1.07)</td>
<td>(-1.42)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9201</td>
<td>0.7136</td>
<td>0.6132</td>
<td>0.4353</td>
</tr>
<tr>
<td></td>
<td>(33.96)**</td>
<td>(14.76)**</td>
<td>(8.43)**</td>
<td>(5.75)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0618</td>
<td>0.2406</td>
<td>0.3342</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>(2.15)**</td>
<td>(4.54)**</td>
<td>(4.25)**</td>
<td>(5.99)**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0003</td>
<td>0.0008</td>
<td>0.0011</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>(2.34)**</td>
<td>(3.10)**</td>
<td>(3.13)**</td>
<td>(3.53)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9973</td>
<td>0.994</td>
<td>0.9898</td>
<td>0.9861</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.9973</td>
<td>0.9939</td>
<td>0.9896</td>
<td>0.9859</td>
</tr>
<tr>
<td>$\hat{t}$</td>
<td>-10.66**</td>
<td>-7.41**</td>
<td>-3.94*</td>
<td>-4.93**</td>
</tr>
<tr>
<td>$\hat{z}$</td>
<td>-145.78**</td>
<td>-85.79**</td>
<td>-35.30**</td>
<td>-49.06**</td>
</tr>
<tr>
<td>$\hat{Z}_{xt}$</td>
<td>-10.81**</td>
<td>-7.46**</td>
<td>-6.73**</td>
<td>-6.81**</td>
</tr>
<tr>
<td>$\hat{Z}_{t}$</td>
<td>-153.18**</td>
<td>-86.88**</td>
<td>-73.17**</td>
<td>-74.81**</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>-0.31</td>
<td>-0.24</td>
<td>-1.07</td>
<td>-1.42</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>-2.95**</td>
<td>-5.92**</td>
<td>-5.32**</td>
<td>-7.45**</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>2.15**</td>
<td>4.54**</td>
<td>4.25**</td>
<td>5.99**</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>-1.49</td>
<td>-1.94*</td>
<td>-1.6</td>
<td>-1.65</td>
</tr>
<tr>
<td>$N$</td>
<td>203</td>
<td>201</td>
<td>198</td>
<td>195</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

Table 7.16: Summary Statistics for Expected Profit—Adaptive Expectations

<table>
<thead>
<tr>
<th></th>
<th>$EP_{t,t+n}$</th>
<th>$EP_{t,t+n}$</th>
<th>$EP_{t,t+n}$</th>
<th>$EP_{t,t+n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=1)</td>
<td>(n=3)</td>
<td>(n=6)</td>
<td>(n=9)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.02**</td>
<td>0.05**</td>
<td>0.09**</td>
<td>0.08**</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.033</td>
<td>0.061</td>
<td>0.073</td>
<td>0.076</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

Figure 7.11 shows the expected profit from speculation using the adaption expectation mechanism. It appears that the variability of the expected profit is lower in comparison to the variability of the expected profits obtained from using extrapolative expectations.
Figure 7.11: Expected Profit—Adaptive Expectations

Note: $EP_{t+1}$, $EP_{t+3}$, $EP_{t+6}$ and $EP_{t+9}$ represent expected profit at 1, 3, 6 and 9-month maturities.

7.7.3 The Random Walk

A cointegrating regression is also estimated based on the random walk model without drift. The results of the regression are presented in Table 7.17. The EG and PO test statistics show that all variables are co-integrated with the exception of the 9-month futures contract. The coefficient for the arbitrage price is significant in all cases, with the exception of the 1-month maturity, whereas the coefficient for the expected spot price is not significant in any case. The restriction $\beta_1 = 1$ cannot be rejected for any maturity, which implies that arbitrage alone is sufficient for determination of the futures price. Also, the restriction $\beta_2 = 0$ cannot be rejected, indicating that speculation plays no role in the determination of futures price.
Table 7.17: Cointegration Regression—Random Walk

<table>
<thead>
<tr>
<th></th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
<th>$F_{t,t+n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>($n=1$)</td>
<td>($n=3$)</td>
<td>($n=6$)</td>
<td>($n=9$)</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>-0.0011</td>
<td>0.0226</td>
<td>-0.0268</td>
<td>-0.0287</td>
</tr>
<tr>
<td></td>
<td>(-0.05)</td>
<td>-0.49</td>
<td>(-0.42)</td>
<td>(-0.34)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>1.7941</td>
<td>1.9065</td>
<td>1.2086</td>
<td>1.2318</td>
</tr>
<tr>
<td></td>
<td>-1.4</td>
<td>(2.11)**</td>
<td>(6.01)**</td>
<td>(5.16)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.824</td>
<td>-1.0053</td>
<td>-0.3075</td>
<td>-0.3878</td>
</tr>
<tr>
<td></td>
<td>(-0.64)</td>
<td>(-1.09)</td>
<td>(-1.55)</td>
<td>(-1.61)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0005</td>
<td>0.0014</td>
<td>0.0016</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>(2.78)**</td>
<td>(4.14)**</td>
<td>(4.44)**</td>
<td>(5.14)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9973</td>
<td>0.9935</td>
<td>0.9892</td>
<td>0.9828</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.9973</td>
<td>0.9934</td>
<td>0.989</td>
<td>0.9825</td>
</tr>
<tr>
<td>$\hat{t}$</td>
<td>-10.60**</td>
<td>-6.45**</td>
<td>-2.93</td>
<td>-2.43</td>
</tr>
<tr>
<td>$\hat{\tau}$</td>
<td>-145.53**</td>
<td>-69.86**</td>
<td>-20.38</td>
<td>-14.47</td>
</tr>
<tr>
<td>$\hat{Z}_x$</td>
<td>-10.77**</td>
<td>-6.34**</td>
<td>-4.87**</td>
<td>-3.85</td>
</tr>
<tr>
<td>$\hat{Z}_t$</td>
<td>-153.67**</td>
<td>-66.35**</td>
<td>-43.10**</td>
<td>-28.66</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>-0.05</td>
<td>0.49</td>
<td>-0.42</td>
<td>-0.34</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>0.62</td>
<td>1</td>
<td>1.04</td>
<td>0.97</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>-0.64</td>
<td>-1.09</td>
<td>-1.55</td>
<td>-1.61</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>-2.26**</td>
<td>-3.53**</td>
<td>-2.93**</td>
<td>-3.53**</td>
</tr>
<tr>
<td>$N$</td>
<td>203</td>
<td>203</td>
<td>203</td>
<td>203</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

The expected profit over the different maturity periods is presented in Table 7.18. The table shows that the expected profit is not significantly different from 0, with the exception of the 3-month maturity.

Table 7.18: Summary Statistics for Expected Profit—Random Walk

<table>
<thead>
<tr>
<th></th>
<th>$EP_{t,t+n}$</th>
<th>$EP_{t,t+n}$</th>
<th>$EP_{t,t+n}$</th>
<th>$EP_{t,t+n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>($n=1$)</td>
<td>($n=3$)</td>
<td>($n=6$)</td>
<td>($n=9$)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.005</td>
<td>0.015**</td>
<td>0.007</td>
<td>0.045</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.025</td>
<td>0.041</td>
<td>0.059</td>
<td>0.073</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.
The expected profit in Figure 7.12 constantly fluctuates within a narrow range over the sample period. However, this is not surprising as futures prices at shorter maturities are almost always indistinguishable from the corresponding spot prices (the proxy for expected spot price in this case). Based on these findings, it could be argued that the impact of speculation on crude oil futures price may be based on the profitability of expectation formation mechanisms employed by speculators.

**Figure 7.12: Expected Profit—Random Walk**

![Graph showing expected profit over time]

Note: $EP_{t+1}, EP_{t+3}, EP_{t+6}$ and $EP_{t+9}$ represent expected profit at 1, 3, 6 and 9-month maturities.

**7.7.4 Regressive Expectations**

The results of the cointegration regression based on the regressive expectation formation technique are provided in Table 7.19. The $\hat{Z}_c$ and $\hat{Z}_t$ test statistics confirm the presence of cointegration. Similar to the results for the previous regression, the coefficient of the theoretical futures price is significant at the 5% level whereas the expected spot price coefficient is not significant even at the 10% level for 6 and 9-month futures contracts.
Table 7.19: Cointegration Regression—Regressive Expectations

<table>
<thead>
<tr>
<th></th>
<th>$F_{t,t+n}$ $(n=1)$</th>
<th>$F_{t,t+n}$ $(n=3)$</th>
<th>$F_{t,t+n}$ $(n=6)$</th>
<th>$F_{t,t+n}$ $(n=9)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>-0.0045</td>
<td>0.0255</td>
<td>-0.0389</td>
<td>-0.0396</td>
</tr>
<tr>
<td></td>
<td>(-0.20)</td>
<td>(-0.57)</td>
<td>(-0.58)</td>
<td>(-0.44)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9244</td>
<td>0.7622</td>
<td>0.8767</td>
<td>0.7874</td>
</tr>
<tr>
<td></td>
<td>(30.70)**</td>
<td>(12.37)**</td>
<td>(8.34)**</td>
<td>(6.35)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0521</td>
<td>0.1528</td>
<td>0.0154</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(1.73)*</td>
<td>(2.60)**</td>
<td>-0.17</td>
<td>-0.33</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0004</td>
<td>0.0012</td>
<td>0.0018</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(2.52)**</td>
<td>(4.16)**</td>
<td>(4.19)**</td>
<td>(5.15)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.997</td>
<td>0.9934</td>
<td>0.9889</td>
<td>0.9832</td>
</tr>
<tr>
<td>Adj.$R^2$</td>
<td>0.997</td>
<td>0.9933</td>
<td>0.9887</td>
<td>0.9829</td>
</tr>
<tr>
<td>$\hat{t}$</td>
<td>-10.02**</td>
<td>-6.49**</td>
<td>-3.25</td>
<td>-2.97</td>
</tr>
<tr>
<td>$\hat{z}$</td>
<td>-129.72**</td>
<td>-64.80**</td>
<td>-21.08</td>
<td>-15.75</td>
</tr>
<tr>
<td>$\hat{Z}_a$</td>
<td>-10.17**</td>
<td>-6.40**</td>
<td>-5.27**</td>
<td>-4.27**</td>
</tr>
<tr>
<td>$\hat{Z}_t$</td>
<td>-136.63**</td>
<td>-61.75**</td>
<td>-43.49**</td>
<td>-28.37</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>-0.2</td>
<td>0.57</td>
<td>-0.58</td>
<td>-0.44</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>-2.51**</td>
<td>-3.86**</td>
<td>-1.17</td>
<td>-1.71*</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>1.73*</td>
<td>2.60**</td>
<td>0.17</td>
<td>0.33</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>-1.85*</td>
<td>-3.41**</td>
<td>-2.90**</td>
<td>-3.69**</td>
</tr>
<tr>
<td>$N$</td>
<td>181</td>
<td>179</td>
<td>176</td>
<td>173</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

The restriction $\beta_2 = 0$ is rejected at 3 and 6-month maturities whereas the restriction $\beta_1 + \beta_2 = 1$ is rejected in all cases, which implies that arbitrage and speculation are not sufficient for future market equilibrium. The expected returns obtained by speculators using the regressive expectation formation mechanism are presented in Table 7.20. The expected profit is significant in all cases and is higher in comparison to profits obtained by the expectations formation methods.
Table 7.20: Summary Statistics for Expected Profit—Regressive Expectations

<table>
<thead>
<tr>
<th></th>
<th>$EP_{t,t+n}$ ($n=1$)</th>
<th>$EP_{t,t+n}$ ($n=3$)</th>
<th>$EP_{t,t+n}$ ($n=6$)</th>
<th>$EP_{t,t+n}$ ($n=9$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.085**</td>
<td>0.087**</td>
<td>0.109**</td>
<td>0.132**</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.071</td>
<td>0.071</td>
<td>0.072</td>
<td>0.114</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

Figure 7.13 shows that expected profit varies considerably over time, and that speculators expect the spot price prevailing at the contract maturity date to be higher than the contemporaneous futures price from 2000 to 2008. By the end of 2008, the expected price is lower than the futures price, consistent with the results of Baumeister and Killian (2014) who report that between 2003 and 2008, oil price expectations were consistently lower than the actual futures prices over long horizons.

The findings derived from this regression are contradictory to our earlier assertion that the impact of speculation can be explained on the basis of the expected profit obtained using the underlying expectation formation mechanism. A possible explanation for this contradictory result in this specific case could be the poor forecasting performance of regressive expectations (as observed earlier) relative to the other expectation formation mechanisms. Therefore, a thorough analysis must be conducted to determine whether the influence of speculation on futures prices varies with the predictive accuracy of the expectation formation mechanism.
Figure 7.13: Expected Profit—Regressive Expectations

Overall, the results of the cointegration analysis indicate that speculators use both technical and fundamental analysis in forming expectations. The impact of speculation varies according to the expectation formation mechanism selected by speculators. Moreover, the impact of speculation appears to be related to the profitability of the speculative strategy, which in turn is based on the expected formation mechanism employed by the speculators. However, this assertion is complicated by the lack of significant impact of expected spot price on the price of 6 and 9-month futures contracts when the expectation formation process takes into account crude oil market fundamentals. One potential explanation for this result could be that speculators expect changes in market fundamentals to influence prices only in the short run. For instance, Fama and French (1988) argue that oil supply shocks, such as those resulting from changes in inventory, have a smaller impact on forward prices relative to spot price due to future inventory response. As a result, less of the change in the current spot prices (due to changes in market fundamentals) would be observed in futures prices with longer maturities.
Support for this argument is found in the results obtained by Baumeister and Kilian (2014), who note that expectation processes based on oil market fundamentals fail to quantify speculative pressures on oil prices.\footnote{The speculative impact on oil prices is measured by the presence of a time-varying risk premium in futures prices.} Prat and Uctum (2011) find that regressive expectations based on macroeconomic and crude oil market variables have limited value in explaining expectations of traders in the presence of structural breaks. In a study on the grains market, Robles et al. (2009) state that ‘changes in supply and demand fundamentals cannot fully explain the recent drastic increase in food prices’, and ‘rising expectations, speculation, hoarding and hysteria also played a role in the increasing level and volatility of food prices’.

It is also observed that the expectation formation mechanisms that rely on technical analysis are important in quantifying the impact of speculation on futures price. This result highlights the role of the theory of storage which postulates that changes in the spread between spot and futures price that exceed the net storage cost should be eliminated by arbitrage. Consequently, changes in the observed futures price should only be explained by the theoretical futures price (in other words, the actions of arbitragers).

In our empirical results, the inability of the theoretical futures price to explain fully the observed futures price can be attributed to the price risk resulting from the expectations of technical traders, which may deter rational arbitragers. A similar argument is made by Kaufmann (2011), who attributes the breakdown in the cointegrating relationship between spot and futures prices of different grades of crude oil to noise traders. He asserts that growth in the participation of noise traders (as measured by the percentage of open interest originating from non-commercial traders) has resulted in exacerbation of oil price
movements, which is responsible for the failure of price models based on market fundamentals.

Also consistent with our results are the findings of Cifarelli and Paladino (2010), who note the presence of a significant relationship between large swings in crude oil price and positive feedback trading. Similarly, by using a regime switching model, Vansteenkiste (2011) shows that the increase in speculative activity after 2004 is consistent with the increase in the probability of traders using a chartist regime. Further support for our findings is observed in the study of Gorton et al. (2012) who find that portfolios of non-commercial traders formed on the basis of prior spot and futures price returns are correlated with futures risk premiums. They also find momentum portfolios to be the main source of speculative risk premiums, which are compensated for in average returns.

At this point it may be argued that the impact of speculative activity is only indicative of the risk premium flowing to speculators. Proponents of the efficient market hypothesis (EMH) may argue that because speculators cannot have a superior forecasting ability, returns must accrue from the compensation for bearing risk. This implies that a naïve trader who earns profit with no forecasting ability considers the return to a speculator as a risk premium. However, using the difference between the commodity basis and the net cost of storage as the measure of risk premium, Chatrath et al. (1997) show that speculators in the commodity futures market generate a profit without receiving a premium for assuming risk. Moreover, their profits entirely arise from their superior forecasting ability and market timing.

---

48 For example, see Chang (1985).
7.8 Summary and Conclusion

This chapter extends the analysis conducted in Chapter 5 by incorporating the role of heterogeneous expectations in the determination of futures price. The main conclusion emerging from the current findings is that speculators in the crude oil futures market behave heterogeneously. This conclusion is formed on the basis of the empirical evidence indicating that the expected spot price of crude oil, simulated on the assumption of heterogeneous expectations, has a significant influence on crude oil price. However, we also find that arbitrage still plays a dominant role in the determination of futures price.

In line with the literature, we find that expectation formation based on historical price patterns (technical analysis) performs better (compared to fundamentals-based expectations) in explaining the impact of speculative activity. These results also shed some light on the usefulness of fundamentals in explaining crude oil price movements. Moreover, the findings suggest that the influence of speculation on crude oil prices varies according to the expectation formation mechanism employed. Based on the reported results, two possible explanations can be stated for this variation. The first is the profitability of the trading strategy. It appears that the expectation formation mechanism that generates significant profits also results in a significant impact of speculative activity on the price of crude oil futures. However, it may be worthwhile to examine whether profitability from these trading strategies varies over the sample period and whether changes can be related to major oil market events. The second possible explanation pertains to the predictive accuracy of the expectation formation method. We find that the fundamentalist expectations formation mechanism performs relatively poorly in predicting the spot price of crude oil. Subsequently, it is observed that the impact of speculative activity based on fundamentalist expectation formation is insignificant for longer maturities. Hence, it may be argued that the elasticity of speculative demand for futures depends on the predictive accuracy of the expectation
formation mechanism. However, a formal analysis is required to establish if there is any link between the predictive accuracy of the expectation process and the elasticity of speculative demand.
8.1 Introduction

With the decline in the popularity of EMH and the representative agent approach, the role of heterogeneous agent behaviour has gained attention in recent models of asset pricing. These models typically involve two types of traders. The first type is the rational trader who adheres to the belief that the price of an asset is determined by its fundamental or equilibrium value. The second is a chartist or a technical trader (sometimes also called a noise trader) who believes that asset prices are not solely determined by the fundamentals and can be predicted on the basis of simple technical trading rules. These traders, in every period, may switch between predictors of future price behaviour, based on the past performance of the predictors. The switching mechanism employed by traders is an interesting feature as it introduces non-linear dynamics into the expectation formation process. Awareness of such heterogeneous behaviour among traders has led to the rise of heterogeneous agents models.49

In this chapter, we employ a heterogeneous agent model (HAM) to generate crude oil price expectations. This analysis is motivated by the finding in Chapter 7 that the impact of speculation on the futures price is sensitive to the expectation formation adopted by speculators. However, it was assumed that speculators conform to the same expectation formation mechanism (regardless of its predictive accuracy) in every period for forecasting crude oil prices. This implies that speculators using fundamentalist expectations continue to

49 See Hommes (2006) for an overview.
expect the oil price to return to its fundamental value. Similarly, speculators using chartist forecasting methods continue to extrapolate past price trends to form futures price expectations. In this chapter, this limitation is addressed by allowing speculators to switch between different expectation formation mechanisms based on their past forecasting performance. According to Brock and Hommes (1997, 1998), switching between prediction models can lead to market instability and in the process, generate returns for speculators that may differ significantly from the returns obtained by non-switching models.

8.2 Heterogeneous Agents Models

The limitations associated with the representative agent approach have led to the popularity of HAMs, which allow traders with different expectations or trading strategies to be modelled jointly without introducing any of the biases observed under the representative agent approach. The interaction of varying expectations determines price formation, such that the resulting market price is a weighted average of the expectations of different groups of traders.

In the heterogeneous agent literature, two main groups of traders are identified: the chartists and the fundamentalists. The chartists, also known as technical traders, base their expectations on historical behaviour of prices. Such a trading style is consistent with successful identification of trends and patterns in past prices to determine their future movement. Fundamentalists, on the other hand, rely on market fundamentals and economic indicators to formulate expectations about future prices. Under this approach, the expectation of future prices depends on the transaction price of an asset relative to its fundamental value. Traders can switch dynamically between chartist and fundamentalist strategies to form expectations based on the past performance of the underlying strategy. The weights allocated to individual strategies vary over time, conditional on the strategy’s goodness of fit. Interaction among different trader groups, combined with the switching between strategies,
yields a non-linear dynamic asset pricing model capable of mimicking the stylised facts of financial markets (De Grauwe and Grimaldi, 2006).

Models based on heterogeneous behaviour of agents have been successfully applied to various markets. For instance, by using behavioural assumptions about chartists and fundamentalists, Zeeman (1974) estimates a heterogeneous model for the equity pricing to explain bull and bear markets. Frankel and Froot (1987, 1990) develop a HAM for exchange rates with time-varying weights for the expectation formation mechanism. Pilbeam (1995) implements a similar model by dividing the fundamentalist and chartist approach into several expectation-forming mechanisms (static, extrapolative, adaptive, regressive, rational and heterogeneous). The study finds that fundamentals, combined with different extrapolative strategies, lead to a significant increase in the probability of forecasting exchange rates. Moosa and Shamsuddin (2004) test for heterogeneity in the foreign exchange rate market by simulating exchange rate series based on different trading strategies. The study finds that the simulated exchange rate series exhibits similar volatility to that of the actual series, thereby confirming heterogeneity in trading behaviour.

An example of HAMs in the commodity market is that of Frechette and Weaver (2001), who find statistical significance for heterogeneity in the US soybean futures market. Westerhoff and Reitz (2005) model the behaviour of heterogeneous interacting agents in the US corn market, revealing that an increase in the participation of chartists in the market is associated with deviations of prices from their fundamental values and suggesting that technical trading exerts a destabilising influence on prices. On the other hand, Reitz and Westerhoff (2007) find that fundamentalists become more active as prices deviate from their long-run equilibrium values. However, their impact becomes relatively low as prices move towards
fundamentals. The study concludes that the destabilising influence of heterogeneous traders may be responsible for strong and persistent misalignments in commodity markets.

Despite substantial literature documenting heterogeneity in various markets, research into heterogeneous trading in the crude oil market is relatively sparse. An important exception in this regard is the study of Ellen and Zwinkels (2010), who analyse heterogeneity among speculators in the WTI spot market, finding that speculators switch between fundamentalist and chartist strategies. Further, deviation of spot prices from their intrinsic values are found to have a negative influence on price changes in subsequent periods, which is indicative of mean reversion in oil prices. In comparison, price changes in previous periods are positively related to the changes in current spot price. The study also reports that HAMs outperform other forecasting models in out-of-sample tests. Although the study differentiates between chartists and fundamentalists, it assumes that agents within each category act homogenously. Further, the empirical methodology only incorporates a single forecasting scheme to model technical trading behaviour.

It can be concluded that speculators use a mix of strategies in forming expectations, irrespective of the market in which they are trading. A contrasting view suggests that expectations can be represented more accurately using the REs approach. However, according to both views, the sensitivity of speculative behaviour to the expectations formation mechanism is crucial for understanding the role of speculators in determination of futures prices. Microstructural studies have applied heterogeneous agent-based models to various markets, but their usefulness in explaining crude oil future prices remains relatively under-researched. Although trader heterogeneity has been associated with episodes of mispricing, the effectiveness of arbitrage in realigning prices remains unaddressed in models explaining trader heterogeneity.
8.3 The Model

We use the empirical model represented in Equation (5.1) for the analysis of futures price, but in order to proxy the expectations of the speculators, a HAM is formulated. The model is adopted from the study of Ellen and Zwinkels (2010), who evaluate the effect of heterogeneous expectations formed by speculators on the prices of the WTI and Brent crude oil. The study of Ellen and Zwinkels is motivated by prior research pertaining to agent behaviour by Brock and Hommes (1997, 1998), Reitz and Westerhoff (2007) and Reitz and Slopek (2009).

The model rests on the assumption that the price of crude oil is determined by the demand of real and speculative market participants. Real participants include consumers and suppliers of oil and oil-based products, whose demand for oil stems from oil market fundamentals. Speculators, on the other hand, have no physical use for crude oil and do not participate in the production process. Speculators in the crude oil market are then placed into two main groups: fundamentalists and the chartists. The next two sections provide a detailed description of the demand functions and the expectation generation process of these two groups of traders.

8.3.1 Fundamentalists

Fundamentalists base their demand for oil on the expected spot price at time $t + 1$ and the actual price at time $t$. This is expressed as:

$$D_t^F = a^F [E_t^F S_{t+n} - S_t] \quad (8.1)$$

where $D_t^F$ is demand for oil by fundamentalists, $S_t$ is the spot price of crude oil at time $t$, $E$ is the expectations operator and $a^F$ measures the sensitivity of the demand. According to

---

50 This is consistent with the model specified in Chapter 3. According to this model, both speculators and hedgers act on the same parameters, hence the demand for hedgers does not need to be specified separately.
Equation (8.1), demand from fundamentalist traders increases with the magnitude of the difference between the expected and the current spot price.

Traders with a fundamentalist approach base their expectations of the spot price on fundamental values as determined by the long run or the equilibrium condition. Fundamentalists expect the oil price to revert back to its long-run value, which means that they would expect the oil price to decline whenever it is above the fundamental value, and to rise whenever it is below the fundamental value. The expected price can therefore be expressed as:

\[ E_t^S t+n = S_t + \beta_1^F (S_t - \bar{S})^+ + \beta_2^F (S_t - \bar{S})^- \]  

(8.2)

where \( \bar{S} \) is the long-run equilibrium value of oil. Equation (8.2) can be used to distinguish between undervaluation and overvaluation. If \( (S_t - \bar{S}) \geq 0 \) then \( (S_t - \bar{S}) = (S_t - \bar{S})^+ \) and if \( (S_t - \bar{S}) < 0 \) then \( (S_t - \bar{S}) = (S_t - \bar{S})^- \). The term \( (S_t - \bar{S})^+ \) indicates overvaluation, whereas \( (S_t - \bar{S})^- \) indicates undervaluation. Equation (8.2) shows that movement in prices is caused by the deviation of oil price from the fundamental value. An undervaluation leads to an expectation of a rise in oil price in the next period whereas overvaluation leads to an expectation of a decline in the next period.

The reaction of fundamentalists to overvaluation (undervaluation) is measured by the terms \( \beta_1^F \in [-1, 0] \) and \( \beta_2^F \in [-1, 0] \). These terms represent the extent to which oil price expectations respond to the degree of undervaluation and overvaluation. For the oil price to revert back to the equilibrium value these parameters must be negative, as fundamentalists will expect the oil price to decline (rise) if the current price is above (below) the long-run value. When \( \beta_1^F \) and \( \beta_2^F \) are equal, there is no net change in the oil price to undervaluation and overvaluation. This distinction provides more flexibility in analysing the behaviour of traders. For example, Kahneman and Tversky (1979) show that the response of investors to potential profit differs
from their response to potential losses. Traders are found to respond more to the notion of potential losses from undervaluation (i.e. they are more reluctant to sell an overvalued asset than to buy an undervalued asset). Canoles et al. (1998) draw similar conclusions in their analysis of the behaviour of speculators in commodity markets. Ellen and Zwinkels (2010) state that due to short-selling constraints, buying pressure may be higher than selling pressure even with similar degrees of undervaluation and overvaluation. As a result, we expect the magnitude of the coefficient for the undervaluation term to be greater than the magnitude of the overvaluation coefficient \( |\beta^1| < |\beta^2| \).

### 8.3.2 Chartists

The second group of speculators are chartists. In the model proposed by Ellen and Zwinkels (2010), it is assumed that chartist use a very simple form of technical analysis to form expectations on the future behaviour of prices. As for fundamentalists, demand is expressed as a positive linear function of the of the expected price change:

\[
D^C_t = \alpha^C[E_{t}^C(S_{t+n} - S_t)] \tag{8.3}
\]

where \( D^C_t \) is demand for oil by chartists and \( \alpha^C \) is a measure of the response of chartists to the difference between the expected and actual price. The question that arises is what functional form of technical analysis should be used to represent chartist expectations, as they may use several different forms (e.g., see Brock et al., 1992). To identify the relevant form of technical analysis to represent chartist expectations, Ellen and Zwinkels rely on two important characteristics: (i) chartists base their expectations on past trends and fluctuations in prices without taking into account other exogenous information, and (ii) chartists expect past trends and movements to continue in the same direction. Traders using this style of technical analysis are also known as ‘momentum traders’, who tend to extrapolate past and current trends irrespective of the equilibrium price, exerting a destabilising effect on prices.
Consistent with the heterogeneous agents literature (e.g., Brock and Hommes, 1998) and following Ellen and Zwinkels (2010), we choose the simplest and most general extrapolative expectations rule that incorporates both of these characteristics. This choice simplifies the identification of positive and negative changes in prices by only including the most recent change in the price, implying the use of an AR(1) specification. This is consistent with the findings of Gjølberg (1985), who reports that autocorrelation in oil price changes does not extend beyond the first lag. Hence chartist expectations can be expressed as:

\[ E_t^c S_{t+n} = S_t + \beta_1^c (S_t - S_{t-1})^+ + \beta_2^c (S_t - S_{t-1})^- \]  

(8.4)

where \((S_t - S_{t-1})^+\) indicates a positive trend or an increase in price over the last period and \((S_t - S_{t-1})^-\) indicates a negative trend. In Equation (8.4) if \((S_t - S_{t-1}) \geq 0\) then \((S_t - S_{t-1}) = (S_t - S_{t-1})^+\) and 0 otherwise. If \((S_t - S_{t-1}) < 0\) then \((S_t - S_{t-1}) = (S_t - S_{t-1})^-\) and 0 otherwise. If \(\beta_1^c\) and \(\beta_2^c\) are positive then traders would expect the most recent trend to continue over the forecasting horizon. However, a negative value for \(\beta_1^c\) and \(\beta_2^c\) would be indicative of contrarian behaviour. Also, if \(\beta_1^c > \beta_2^c\) then traders are more likely to take a long position in an upwards trend than to take a short position in a downwards trend.

8.3.3 Learning Mechanism

For this analysis it is assumed that speculators do not conform to one forecasting strategy, but switch between strategies based on the past forecasting performance of the strategy. In this regard, Ellen and Zwinkels (2010) formulate a weighting rule based on the study of Brock and Hommes (1998), according to which the performance of a strategy is based on its past predictive accuracy. The forecasting performance is gauged in terms of the magnitude of the squared forecasting errors in the previous \(k > 0\) months.

The specification of the learning mechanism is based on the study of Brock and Hommes (1998) on discrete choice theory. According to this approach, speculators are seen as using
fundamentalist and chartist approaches to evaluate the forecasting rules. The weights assigned to a forecasting rule are based on the forecasting error of the strategy in the previous $k > 0$ months. Hence:

$$SFE^E_t = -\sum_{k=1}^{K} [E^E_{t-k-1}S_{t-k} - S_{t-k}]^2$$ \hspace{1cm} (8.5)$$

$$SFE^C_t = -\sum_{k=1}^{K} [E^C_{t-k-1}S_{t-k} - S_{t-k}]^2$$ \hspace{1cm} (8.6)$$

where $E^E_{t-k-1}S_{t-k}$ represents period $t - k - 1$ expectations of the spot price at period $t - k$ using a fundamentalist rule and $E^C_{t-k-1}S_{t-k}$ represents period $t - k - 1$ expectations of the spot price at time $t - k$ using a chartist rule. $SFE^E_t$ represents the cumulative forecasting error of the fundamentalist forecasting strategies and $SFE^C_t$ represents the cumulative forecasting error of the chartist method. Rational investors choose the forecasting method that leads to the forecast that is closest to the realised value, or where $SFE$ is the lowest. Ellen and Zwinkels (2010) then construct a multinomial switching rule to determine the fraction of fundamentalist traders in the market relative to chartist traders. This switching rule is given by:

$$W_t = \left[1 + \exp \left(-\gamma \frac{SFE^E_t - SFE^C_t}{SFE^E_t + SFE^C_t}\right)\right]^{-1}$$ \hspace{1cm} (8.7)$$

where $0 < W_t < 1$ is the weight assigned to the expectations of future price using a fundamentalist rule and $(1 - W_t)$ is the fraction of weights assigned to spot price expectations formed using a chartist rule. The parameter $\gamma$ in Equation (8.7) represents the extent of adoption of a strategy based on its past forecasting performance. The parameter could also be interpreted as the level of status quo bias (Samuelson and Zeckhauser, 1988). If $\gamma > 0$ then the strategy with the better past performance will receive more weight in forming the next period’s forecast. When $\gamma \to \infty$, all speculators would adopt the strategy.
with the best forecasting performance (smallest forecast error magnitude over the past $k$ periods). In this case, $W_t$ would be equal to 1 if $SFE^F_t > SFE^C_t$ and 0 if $SFE^F_t < SFE^C_t$. On the other hand, if $\gamma = 0$, then speculators will be insensitive to the relative performance of the strategy such that $W_t = 0.5$. The impact of non-switching speculators is tested as a static HAM by restricting $\gamma$ to 0.

The expectations for the spot price of oil prevailing at the maturity of the futures contract can be specified as the weighted average of the expectations formed by the chartists and fundamentalists:

$$E_t S_{t+n} - S_t = W_t E^F_t S_{t+n} + (1 - W_t)E^C_t S_{t+n} \quad (8.8)$$

Substituting the value of $E^F_t S_{t+n}$ and $E^C_t S_{t+n}$ into Equation (8.8) leads to the following sequence of equations:

$$E_t S_{t+n} - S_t = \delta_0 + W [\alpha_1 (S_t - \bar{S})^+ + \alpha_2 (S_t - \bar{S})^-] + (1 - W) [\beta_1 (S_t - S_{t-1})^+ + \beta_2 (S_t - S_{t-1})^-] + \epsilon_t \quad (8.9)$$

$$W_t = \left[1 + \exp \left(-\gamma \frac{SFE^F_t - SFE^C_t}{SFE^F_t + SFE^C_t} \right) \right]^{-1} \quad (8.10)$$

$$SFE^F_t = -\sum_{k=1}^{K} [E^F_{t-k}S_{t-k} - S_{t-k}]^2 \quad (8.11)$$

$$SFE^C_t = -\sum_{k=1}^{K} [E^C_{t-k}S_{t-k} - S_{t-k}]^2 \quad (8.12)$$

In Equation (8.9), $\alpha_1$ and $\alpha_2$ represent the price impact of fundamentalists’ expectations when oil is mispriced relative to its equilibrium value, $\bar{S}$ and $\beta_1$ and $\beta_2$ represent the price impact of chartists’ expectations when the oil price in the current period is higher or lower than in the previous period.
8.4 Methodology

This section outlines the empirical framework for this study. The first step involves estimating the mean reversion component, $S_t - \bar{S}$ in Equation (8.2). For this purpose, residuals obtained from the estimation of Equation (7.1) are used. The crude oil price expectations are then generated from Equation (8.9). According to the prospect theory, the risk aversion of fundamentalists may vary according to the degree of over- and undervaluation of the spot price. As a result, undervaluation and overvaluation in Equation (8.9) are identified by a dummy variable, which takes the value of 1 if $(S_t - \bar{S}) \geq 0$ and 0 otherwise. If this condition is true then $(1 - \text{dummy variable}) = 0$, which allows distinction between undervaluation and overvaluation. Similarly, we employ a dummy variable to identify positive and negative trends.

To obtain parameter values for Equation (8.9), we follow Ellen and Zwinkels (2010) by employing the quasi-maximum likelihood (QML) estimation approach. This approach involves estimating a finite number of parameters in a manner such that the probability (likelihood) of obtaining the observed values of the endogenous variable is maximised (Wooldridge, 1991). Thus, the QML approach is based on the same principle as maximum likelihood (ML) estimation except that ML estimation relies on the assumption that the error terms are normally distributed. If the error terms are non-normal or heteroscedastic, the variance–covariance matrix of the ML estimator will no longer be appropriate (Goncalves and White, 2004), which biases the ML estimate of the true parameter. The QML approach addresses this issue by using White and Newey-West estimators to provide a robust estimate of the asymptotic covariance matrix of the estimator (Greene, 2003). A simple method of performing a QML estimation is to conduct a regular ML estimation and adjust the error terms using the covariance matrix. The covariance matrix in this case equals the outer product of the inverse Hessian or the gradient:
\[ \text{Cov} = H^{-1}gg'H^{-1} \] (8.13)

where \( \text{Cov} \) is the covariance matrix, \( H \) is the Hessian and \( g \) is the gradient. Equation (8.13) is also known as the ‘sandwich estimator’.

To obtain a forecast of the spot price of oil, Equation (8.9) is estimated recursively over a portion of the sample period using past data (the recursive estimation procedure is discussed in detail in Section 7.4). It must be noted that because the forecast is obtained by averaging the monthly forecasted values, the expected spot price sample is reduced by \( n - 1 \) where \( n \) is the forecasting horizon. The initial estimation window consists of the first 50 observations, which include the period January 1994–February 1998.\(^{51}\) Once the expected spot price series is constructed, the expected profit for a speculative strategy can be estimated as the difference between the expected spot price and the observed futures price at time \( t \).

One issue that arises in the estimation of the forecasting error in Equations (8.5) and (8.6) is related to determining the value of \( k \)—the number of months over which the forecasting accuracy of the chartists and fundamentalists are evaluated. Following Ellen and Zwinkels (2010), the number of previous \( k \)-months is determined empirically by selecting the model with the highest log-likelihood score.\(^{52}\) This approach is similar to the Box–Jenkins method, which is applied to find the best fit of a time series to make forecasts.

To evaluate predictive accuracy, the forecasts obtained from the HAMs (switching and non-switching case) are compared to that of the random walk model without drift. The random walk without drift is selected as it has been used as a natural benchmark for evaluating crude

---

\(^{51}\) The selection of the initial estimation window is limited by OPEC’s spare capacity data, which are available from January 1994.

\(^{52}\) \( k=5 \) results in the highest log-likelihood value on average for all forecasting horizons. Ellen and Zwinkels (2010) use 6 months.
oil price forecasts in past studies.\textsuperscript{53} The AGS test is conducted to determine any statistical
difference between the forecasting errors of HAMs and the random walk model without drift.
The DA of the expectation formation methods is also examined. For this purpose, a test of
significance is conducted to determine if the DA of the relevant expectation formation
mechanism is significantly different from 0 and 50%.

Forecast evaluation is followed by cointegration analysis to determine if a long-run
relationship exists between the observed futures price, the theoretical futures price and the
expected spot price. For this purpose, a cointegration regression is estimated for Equation
(5.1) using the FMOLS technique. Prior to undertaking the cointegration regression, ADF
and PP unit root tests are conducted to ensure that $F_{t,t+n}$, $\bar{F}_{t,t+n}$ and $E_t S_{t+n}$ are integrated of
the same order. The EG and PO tests are conducted on the regression residuals to confirm the
presence of cointegration.

8.5 Data Description and Empirical Results

The study uses monthly spot and futures prices of the WTI crude oil traded on the NYMEX
from January 1994 to April 2013. Although several grades of crude oil are produced in the
US, WTI has significant importance in global oil and financial markets because it underlies
one of most actively traded commodity futures contracts. Because of their liquidity and
transparency, light sweet crude oil futures also provide excellent means of managing energy
risk. The sample comprises futures contracts with maturities of 1, 3, 6 and 9 months.\textsuperscript{54}

8.5.1 Crude Oil Price Forecast Evaluation

The evaluation of oil price forecasts commences with a visual comparison of the forecast
\textsuperscript{53} A few notable studies include Meese and Rogoff (1983), MacDonald and Marsh (1993), Diebold and Mariano
\textsuperscript{54} Refer to Chapter 5, Section 5.3 for a detailed description of sample and time series characteristics of spot and
futures prices.
obtained from the HAM with that of the random walk. The HAM is estimated by using constant as well as time-varying weights. When the weights are held constant, speculators do not switch between chartist and fundamentalist forecasting rules, which gives $\gamma = 0$ and $W_t = W = 0.5$. The resulting model is referred to as a ‘non-switching HAM’.

When the weights are allowed to vary, the speculators switch freely between chartist and fundamentalist forecasting rules such that the intensity of switching is based on the predictive accuracy (as measured by forecasting error) of the forecasting rule. The model obtained in this case is referred to as a ‘switching HAM’. A forecast of 1, 3, 6 and 9-month-ahead expected spot price is generated for each of the switching and non-switching cases. The oil price forecasts are evaluated over the period March 1998–April 2013.

The time series plots of the forecasts are presented in Figure 8.1. Over the 1-month horizon, the time plot of the forecasted price series is indistinguishable from that of the actual price. The strong similarities in the plots can be attributed primarily to the inclusion of a lagged spot price as an independent variable in all of the expectation generation mechanisms. At the 3-month horizon, time series of the forecasts systematically lag behind the actual price. The forecasts of switching and non-switching HAMs closely track the forecasts obtained from the random walk with no noticeable difference. The time series of 1 and 3-month-ahead forecasts closely resemble the forecasts obtained earlier using chartist and fundamentalist rules (see Figure 7.6).

The graphs for 6 and 9-month-ahead forecasts show that the difference between actual and predicted series increases with the forecast horizon. A similar observation is made by Alquist et al. (2011) and Baumeister and Kilian (2012) in their analysis of Brent and WTI crude oil prices. Over 6 and 9-month horizons, the performance of the switching HAM appears to be
identical to that of the random walk. In contrast, the non-switching HAM shows less stability over the evaluation period, inferred by the clear deviation of the forecast of non-switching HAM from other forecast series from 2004 to 2007 and from 2010 to 2012. It should be noted that over these periods the WTI crude oil price exhibits higher volatility.

Figure 8.1: Oil Price Forecasts —Heterogeneous Agent Model

Figure 8.1 also illustrates that at forecasting horizons greater than 1 month, even the random walk under-predicts the WTI crude oil price between 2006 and mid-2008, and then over-predicts after the decline in the crude oil price at the end of 2008, which is again followed by an under-prediction in the recent period. The forecast series tracks the actual price in a lagged
manner, which suggests that expectation formation mechanisms fail to provide a timely prediction of sharp changes in price.

Visual inspection of the time plots suggests that the expectation formation mechanisms perform equally well in forecasting WTI crude oil price at the 1 and 3-month horizons. Over the 6 and 9-month horizons, the forecasting accuracy improves considerably in comparison with the non-switching HAM in which forecasters switch between the chartist and fundamentalist expectation formation mechanisms. However, forecasting accuracy cannot be evaluated on the basis of visual inspection alone; thus, the magnitude of errors is tested to determine if the differences in accuracy are statistically significant.

Table 8.1 presents the RMSEs for the switching HAM and the random walk. The results of the AGS test are also presented in the table. With the exception of the 9-month horizon, the RMSEs for the switching HAM are numerically higher than the RMSE of the random walk model. In line with the visual analysis of the graphs performed earlier, the RMSEs increase with the length of the forecasting horizon. The RMSE of the random walk increases by a smaller amount over the forecasting horizons than that of the switching HAM.

According to the AGS test results, the intercept term $\alpha_0$ is significantly positive for all forecasting horizons whereas the slope $\alpha_1$ is not significantly different from 0, with the exception of the 9-month forecast horizon where it is significantly positive at the 5% level. The Wald chi-square test statistics suggest that the null hypothesis $\alpha_0 = \alpha_1 = 0$ can be rejected at the 5% level, which means that the random walk generates a statistically smaller forecasting error than does the switching HAM.
Table 8.1: Root Mean Squared Errors for the Heterogeneous Agent Model with Time-varying Weights

<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE (switching HAM)</td>
<td>2.82</td>
<td>6.12</td>
<td>9.54</td>
<td>10.09</td>
</tr>
<tr>
<td>RMSE (random walk)</td>
<td>2.73</td>
<td>4.92</td>
<td>7.73</td>
<td>10.86</td>
</tr>
<tr>
<td>$t(a_0)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.219)**</td>
<td>(0.512)**</td>
<td>(0.887)**</td>
<td>(1.031)**</td>
</tr>
<tr>
<td>$t(a_1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.011</td>
<td>0.032</td>
<td>0.047</td>
<td>(0.0462)**</td>
</tr>
<tr>
<td>$a_0 = a_1 = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(22.41)**</td>
<td>(26.74)**</td>
<td>(38.22)**</td>
<td>(43.19)**</td>
</tr>
</tbody>
</table>

Note: $t(a_0)$ and $t(a_1)$ represent the $t$-values for the intercept ($a_0$) and slope ($a_1$) estimates. The test statistic follows a $\chi^2(2)$ distribution, * and ** indicate significance at the 10 and 5% levels.

The RMSEs and the AGS test results for forecast equivalence between the non-switching HAM and the random walk model are presented in Table 8.2. The RMSEs for the non-switching HAM exceed those of the random walk over all forecasting horizons. The RMSE of the non-switching HAM increases as the forecast horizon becomes longer. This is consistent with the analysis of the time plots for the forecasts, which shows that the magnitude of the difference between the forecasts of the non-switching HAM and the random walk increases with forecasting horizon. The AGS test results show that the null hypothesis of the forecast equivalence is rejected in all cases. This implies that the RMSE of the random walk is statistically lower than that of the non-switching HAM.

The above analysis demonstrates that the use of HAMs for forecasting crude oil prices in general does not yield any significant gains in accuracy over the random walk. Although the switching mechanism does lead to an increase in forecasting performance at longer horizons, the overall evidence favours the random walk over HAMs.
Table 8.2: Root Mean Squared Errors for the Heterogeneous Agent Model with Constant Weights

<table>
<thead>
<tr>
<th></th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE (regressive)</td>
<td>2.91</td>
<td>6.11</td>
<td>10.95</td>
<td>14.10</td>
</tr>
<tr>
<td>RMSE (random walk)</td>
<td>2.79</td>
<td>5.79</td>
<td>9.03</td>
<td>10.86</td>
</tr>
<tr>
<td>$t(a_0)$</td>
<td>(0.331)*</td>
<td>(0.510)**</td>
<td>(0.965)**</td>
<td>(1.588)**</td>
</tr>
<tr>
<td>$t(a_1)$</td>
<td>–0.016</td>
<td>–0.003</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>$a_0 = a_1 = 0$</td>
<td>(37.81)**</td>
<td>(42.07)**</td>
<td>(71.30)**</td>
<td>(26.16)**</td>
</tr>
</tbody>
</table>

Note: $t(\alpha_0)$ and $t(\alpha_1)$ represent the t-values for the intercept ($\alpha_0$) and slope ($\alpha_1$) estimates. The test statistic follows a $\chi^2(2)$ distribution, * and ** indicate significance at the 10 and 5% levels.

These results are inconsistent with the findings of Ellen and Zwinkels (2010), who show that the switching HAM outperforms the random walk without drift over the 1, 3, 6 and 9-month horizons. The contradictory findings obtained in this study can be attributed to the method of estimating the equilibrium oil price. Ellen and Zwinkels proxy the equilibrium oil price by the moving average of past oil prices, whereas in this study the equilibrium is estimated empirically using oil market and macroeconomic variables.

The DA of the forecasts is also assessed by testing whether the proportion of successful predictions is significantly different from 0 and 50%. The results of the DA tests are presented in Table 8.3 and show that the DA of switching and non-switching HAMs is significantly different from 0 at all forecasting horizons. The switching HAM has a DA between 71 and 49% whereas the accuracy of non-switching ranges from 69 to 41%, depending on the forecasting horizon. Consistent with the RMSE results, DA generally declines with the forecasting horizons.
Table 8.3: Directional Accuracy of the Forecasts

<table>
<thead>
<tr>
<th></th>
<th>$t+1$</th>
<th>$t+3$</th>
<th>$t+6$</th>
<th>$t+9$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0: DA=0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching HAM</td>
<td>0.71</td>
<td>0.66</td>
<td>0.49</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>(13.74)**</td>
<td>(9.58)**</td>
<td>(8.80)**</td>
<td>(9.44)**</td>
</tr>
<tr>
<td>Non-switching HAM</td>
<td>0.69</td>
<td>0.52</td>
<td>0.46</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(12.60)**</td>
<td>(10.3)**</td>
<td>(9.34)**</td>
<td>(8.67)**</td>
</tr>
<tr>
<td>$H_0: DA=0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching HAM</td>
<td>0.71</td>
<td>0.66</td>
<td>0.49</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>(7.38)**</td>
<td>(2.25)**</td>
<td>(–0.49)</td>
<td>(–1.13)</td>
</tr>
<tr>
<td>Non-switching HAM</td>
<td>0.69</td>
<td>0.52</td>
<td>0.46</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(2.06)**</td>
<td>(1.55)</td>
<td>(–0.63)</td>
<td>(–2.39)</td>
</tr>
</tbody>
</table>

Note: $T$-values are reported in parenthesis; * and ** indicate significance at the 10 and 5% levels.

However, in the case of the switching HAM and the random walk, the DA improves at 9-month horizons relative to the previous forecasting period. The null hypothesis that $DA=0.5$ cannot be rejected on most occasions. However, the DA of the switching HAM is significant at the 1 and 3-month horizons, and that of the non-switching HAM is significant at the 1-month horizon. These results indicate that switching and non-switching specifications of the HAM only provide modest gains in predicting the change in direction more than 50% of the time.

Table 8.4: Directional Accuracy of the Forecasts Using the Pesaran–Timmermann (PT) Test

<table>
<thead>
<tr>
<th></th>
<th>$t+1$</th>
<th>$t+3$</th>
<th>$t+6$</th>
<th>$t+9$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching HAM</td>
<td>0.71</td>
<td>0.66</td>
<td>0.49</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>(6.09)**</td>
<td>(4.23)**</td>
<td>(1.97)</td>
<td>(2.81)*</td>
</tr>
<tr>
<td>Non-switching HAM</td>
<td>0.69</td>
<td>0.52</td>
<td>0.46</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(5.36)**</td>
<td>(2.39)</td>
<td>(2.01)</td>
<td>(1.82)</td>
</tr>
</tbody>
</table>

Note: The PT test statistic follows a $X^2(2)$ distribution and has critical values of 3.84 and 2.71 for significance at the 5(**) and 10%(*) levels.

The results of the PT test are reported in Table 8.4. DA, in most cases, declines as the forecasting horizon becomes longer. The results also show that no predictable relationship is obtained between the forecast and realised changes (the null hypothesis cannot be rejected) at
the 6-month horizon for the switching HAM. In the case of the non-switching HAM, the null hypothesis can only be rejected at the 1-month horizon, which implies that in most cases the predicted and realised changes are not independent. The results support the findings of the DA test discussed earlier in suggesting that DA deteriorates at longer horizons.

The conclusion that can be drawn from the forecast evaluation is that, in most cases, the HAMs do not outperform the random walk model without drift in out-of-sample forecasting, although the HAMs do perform better than the random walk model in terms of the DA. The reason why the random walk ranks lower in terms of DA is that it always predicts no change; that is, it has a DA of zero on each occasion. However, the higher DA of the HAMs comes at the expense of lower statistical accuracy in terms of error magnitude.

It must be emphasised here that both magnitude and direction matter to speculators in the futures market. For instance, the expected profit for a speculator with a long position in the futures market equals $E_tS_{t+n} - F_{t,t+n}$, whereas the expected profit for a speculator with a short position is given by $F_{t,t+n} - E_tS_{t+n}$. In both cases, the speculator must predict accurately the direction of change in the spot price to earn a positive return. At the same time, expected profit must exceed a certain threshold (transaction cost, return on a benchmark, etc.) for the trade to be profitable. This suggests that the magnitude of the forecasting error is important to the success of the underlying speculative strategy.

8.5.2 The Heterogeneous Agent Model Estimation Results

The results of the switching and non-switching HAM estimations are presented in Table 8.5, showing that the likelihood scores are similar for the switching and non-switching models.
Table 8.5: Heterogeneous Agent Model

<table>
<thead>
<tr>
<th></th>
<th>Static</th>
<th>Switching</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta_0 )</td>
<td>( -0.0257 )</td>
<td>( -0.0159 )</td>
</tr>
<tr>
<td></td>
<td>( (-0.25) )</td>
<td>( (-0.86) )</td>
</tr>
<tr>
<td><strong>Fundamentalists</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.1108</td>
<td>-0.168</td>
</tr>
<tr>
<td></td>
<td>(0.94)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.1565</td>
<td>-0.4462*</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(1.7109)</td>
</tr>
<tr>
<td><strong>Chartists</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.2924**</td>
<td>0.1807**</td>
</tr>
<tr>
<td></td>
<td>(4.2034)</td>
<td>(5.090)</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.5418**</td>
<td>-0.3855**</td>
</tr>
<tr>
<td></td>
<td>(4.1182)</td>
<td>(7.1810)</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>1.3179**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.3204)</td>
<td></td>
</tr>
<tr>
<td>LogL</td>
<td>122.45</td>
<td>133.17</td>
</tr>
<tr>
<td>N</td>
<td>181</td>
<td>181</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

The intensity of choice \( \gamma \) is positive and significantly different from 0, which indicates that speculators switch to the strategy that performs better in terms of the magnitude of the forecasting errors. The coefficients \( \alpha_1 \) and \( \alpha_2 \) carry the hypothesised signs for both static and switching cases, which implies that speculators do expect a correction when the price is above or below its equilibrium value. This is consistent with the assertion by Jong et al. (2009) that the probability of adjustment back to equilibrium increases with the discrepancy between the actual and fundamental values. However, neither \( \alpha_1 \) and \( \alpha_2 \) are significantly different from 0 in the static case, whereas \( \alpha_2 \) is significant when speculators are allowed to switch between forecasting strategies. This identifies an influence of fundamentalist strategies on expectations even when price misalignment is limited. The limited explanatory value of the fundamentalist strategy could be in part due to the strategy’s performance being assessed only on the basis of forecasting error magnitude. It may be worthwhile to determine...
if the performance of the strategy improves when agents also include other performance measures such as successful prediction of the sign of change in price of crude oil.

In comparison, Ellen and Zwinkels (2010) find $\alpha_1$ and $\alpha_2$ to be significantly different from 0 for the static and non-switching cases. However, they compute the fundamental value as a moving average of the spot price over a period of 24 months, so stabilisation is more likely to occur towards the fundamental value. According to Reitz and Slopek (2009), the fundamental value must account for the informational advantage that accrues from the incorporation of oil market fundamentals.

In contrast to the findings of Ellen and Zwinkels (2010), the results here show that the coefficient of undervaluation is larger than the coefficient of overvaluation, which indicates that fundamentalist speculators respond more to undervaluation than to overvaluation. The chartist coefficients $\beta_1$ and $\beta_2$ are significant at the 5% level for both static and switching investors. With the exception of the coefficient $\beta_2$ for chartists that switch strategies, the coefficients $\beta_1$ and $\beta_2$ are positively signed, which indicates that chartists on average demonstrate momentum behaviour and as a result exert a destabilising influence on oil prices. The negative $\beta_2$ coefficient for the switching case, on the other hand, suggests contrarian behaviour and a stabilising influence. Also consistent with Ellen and Zwinkels is the finding that $\beta_2 > \beta_1$, which suggests that (on average) when the price declines in one period, chartists expect the price to decline in the next period as well. The general chartist reaction is significant, which suggests that they extrapolate past price movements to form expectations about futures price behaviour. Similarly, Vansteenkiste (2011) finds that the increase in speculative activity during periods of high volatility is consistent with the increase in the probability of adopting a chartist trading regime.
8.5.3 Cointegration Testing

Prior to conducting the cointegration test, the unit root tests are performed to determine the order to integration of the expected spot price. Table 8.6 displays the results of unit root test at levels.

Table 8.6: Unit Root Test for Expected Spot Price at Log Levels

<table>
<thead>
<tr>
<th></th>
<th>Augmented Dickey– Fuller</th>
<th>Phillips– Perron test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>intercept</td>
</tr>
<tr>
<td>$E_tS_{t+1}$</td>
<td>-1.4849</td>
<td>-3.0106</td>
</tr>
<tr>
<td>$E_tS_{t+3}$</td>
<td>-1.5209</td>
<td>-3.1001</td>
</tr>
<tr>
<td>$E_tS_{t+6}$</td>
<td>-1.6563</td>
<td>-3.0839</td>
</tr>
<tr>
<td>$E_tS_{t+9}$</td>
<td>-1.6161</td>
<td>-2.9548</td>
</tr>
</tbody>
</table>

Note: All variables are expressed as natural logs.

The null hypothesis of a unit root cannot be rejected at the 10% level on all occasions with or without a trend component using the ADF and PP tests. The unit root tests are then conducted on the first differences of the expected spot price. The results are reported in Table 8.7. The null hypothesis of non-stationarity is rejected at the 5% level on all occasions with and without a trend. These results confirm that the expected spot price variable is integrated of order 1 or $I(1)$ and a long-run relationship can be estimated between the observed futures price, the theoretical futures price and the expected spot price.

Table 8.7: Unit Root Test for Expected Spot Price at First Difference

<table>
<thead>
<tr>
<th></th>
<th>Augmented Dickey Fuller</th>
<th>Phillips Perron Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with intercept</td>
<td>with trend and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>intercept</td>
</tr>
<tr>
<td>$E_tS_{t+1}$</td>
<td>-11.4040**</td>
<td>-11.3711**</td>
</tr>
<tr>
<td>$E_tS_{t+3}$</td>
<td>-11.7012**</td>
<td>-11.6660**</td>
</tr>
<tr>
<td>$E_tS_{t+6}$</td>
<td>-11.7715**</td>
<td>-11.7380**</td>
</tr>
<tr>
<td>$E_tS_{t+9}$</td>
<td>-11.7620**</td>
<td>-11.7253**</td>
</tr>
</tbody>
</table>

Note: All variables are expressed as natural logs. * and ** indicate significance at the 10 and 5% levels.
After having ascertained the order of integration of the expected spot price, the analysis proceeds with the cointegration tests. The results of the cointegration regression are presented in Table 8.8. The expected spot price prevailing on the futures contract maturity is estimated using the switching HAM. The EG \( \hat{\tau} \) and \( \hat{z} \) statistics, as well as PO \( Z_a \) and \( Z_t \) statistics, confirm the existence of a cointegration relationship between the observed futures price, the cost of carry price and expected spot price in all cases. The demand for futures by the arbitragers, which is proxied by the cost of carry price, is significant at the 5% level and positively related to the futures price. The expected spot price is also significant and positively related to the futures price. The coefficient for the cost of carry price is higher than the expected spot price coefficient. This indicates that the impact of arbitrage dominates the impact of speculation.

The Wald coefficient restriction tests show that the intercept is not significantly different from 0. The restrictions \( \beta_1 = 1 \) and \( \beta_2 = 0 \) are rejected at the 5% level for all maturities, which confirms the significance of both arbitrage speculation. However, the restriction \( \beta_1 + \beta_2 = 1 \) is only accepted for the 1 and 3-month maturities. The rejection of the Wald coefficient restriction for the 6 and 9-month maturities implies that arbitrage and speculation are not sufficient to explain the prices of future contract over these periods, and there exists another factor (in addition to arbitrage and speculation) that determines the price of crude oil futures. A possible explanation for this result is the exclusion of the expected inflation rate from spot price expectations. This can be validated by scaling the forecast of crude oil price by the inflation rate expected to prevail over the forecast horizon.
Table 8.8: Cointegration Regression—Heterogeneous Agent Model

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>$F_{t,t+n}$ ($n=1$)</th>
<th>$F_{t,t+n}$ ($n=3$)</th>
<th>$F_{t,t+n}$ ($n=6$)</th>
<th>$F_{t,t+n}$ ($n=9$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>-0.006</td>
<td>-0.018</td>
<td>-0.0568</td>
<td>-0.087</td>
</tr>
<tr>
<td></td>
<td>(-0.28)</td>
<td>(-0.48)</td>
<td>(-0.96)</td>
<td>(-1.25)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9314</td>
<td>0.859</td>
<td>0.7492</td>
<td>0.6672</td>
</tr>
<tr>
<td></td>
<td>(34.06)**</td>
<td>(21.23)**</td>
<td>(8.43)**</td>
<td>(5.94)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0729</td>
<td>0.1178</td>
<td>0.2318</td>
<td>0.2869</td>
</tr>
<tr>
<td></td>
<td>(1.91)*</td>
<td>(6.09)**</td>
<td>(6.75)**</td>
<td>(7.67)**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0003</td>
<td>0.001</td>
<td>0.0014</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>(2.48)**</td>
<td>(4.21)**</td>
<td>(4.10)**</td>
<td>(4.57)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9973</td>
<td>0.9945</td>
<td>0.9895</td>
<td>0.9861</td>
</tr>
<tr>
<td>Adj.$R^2$</td>
<td>0.9972</td>
<td>0.9944</td>
<td>0.9893</td>
<td>0.9859</td>
</tr>
<tr>
<td>$t$</td>
<td>-11.22**</td>
<td>-7.12**</td>
<td>-5.86**</td>
<td>-5.50**</td>
</tr>
<tr>
<td>$z$</td>
<td>-178.28**</td>
<td>-86.93**</td>
<td>-61.69**</td>
<td>-54.78**</td>
</tr>
<tr>
<td>$Z_{xt}$</td>
<td>-11.57**</td>
<td>-7.15**</td>
<td>-5.85**</td>
<td>-5.48**</td>
</tr>
<tr>
<td>$Z_t$</td>
<td>-196.06**</td>
<td>-87.37**</td>
<td>-61.11**</td>
<td>-54.23**</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>0.08</td>
<td>0.23</td>
<td>0.93</td>
<td>1.56</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>7.42**</td>
<td>59.88**</td>
<td>71.43**</td>
<td>93.87**</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>3.64*</td>
<td>37.04**</td>
<td>45.60**</td>
<td>58.90**</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>1.56</td>
<td>2.16</td>
<td>4.19**</td>
<td>6.23**</td>
</tr>
<tr>
<td>$N$</td>
<td>181</td>
<td>179</td>
<td>176</td>
<td>173</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

Table 8.9 displays the average expected profit obtained by speculators. According to the table, speculators (with long and short positions in the futures market) earn significant expected profit across all maturities.

Table 8.9: Summary Statistics for Expected Profit—Heterogeneous Agent Model

<table>
<thead>
<tr>
<th></th>
<th>$EP_{t,t+n}$ ($n=1$)</th>
<th>$EP_{t,t+n}$ ($n=3$)</th>
<th>$EP_{t,t+n}$ ($n=6$)</th>
<th>$EP_{t,t+n}$ ($n=9$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.089**</td>
<td>0.095**</td>
<td>0.127**</td>
<td>0.167**</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.072</td>
<td>0.087</td>
<td>0.120</td>
<td>0.155</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.
Figure 8.2 shows that expected profits obtained by speculators are highest in 2008, when crude oil price reached a peak level and subsequently declined. This period also corresponds to the decline in the predictive accuracy of chartist and fundamentalist expectation formation methods (see Figures 7.6 and 7.7).

**Figure 8.2: Expected Profit—Switching Heterogeneous Agent Model**

![Graph showing expected profits over time]

Note: \( EP_{t+1}, EP_{t+3}, EP_{t+6} \) and \( EP_{t+9} \) represent expected profit at 1, 3, 6 and 9-month maturities.

Although these results do show a significant role for speculation, it is not clear whether the HAM provides any additional value in explaining the impact of speculation on oil price. The results obtained from cointegration analysis appear to be similar to the findings obtained earlier using chartist expectation formation mechanisms. However, in contrast with fundamentalist expectation mechanisms, the impact of speculation varies significantly at longer maturities (as measured by significance of the coefficient \( \beta_2 \)) when expectations are formed using a HAM.

However, this cannot be concluded simply on the basis of the numerical value of coefficients. As a result, a \( z \)-test is conducted to determine if the difference between the two coefficients is
significantly different from 0. For this purpose we test the null hypothesis $H_0: \beta_x = \beta_y$ against the alternative, $\beta_x \neq \beta_y$, where $\beta_x$ and $\beta_y$ represent the coefficients of the expected spot price generated using regressive expectations and the switching HAM. A similar test is conducted for the chartist expectation formation mechanisms. The results of the test are presented in Table 8.10.

**Table 8.10: Test of Coefficient Equality**

<table>
<thead>
<tr>
<th>Fundamentals</th>
<th>$t + 1$</th>
<th>$t + 3$</th>
<th>$t + 6$</th>
<th>$t + 9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressive expectations</td>
<td>0.41</td>
<td>-0.57</td>
<td>(2.23)**</td>
<td>(2.31)**</td>
</tr>
<tr>
<td>Extrapolative expectations</td>
<td>(1.14)</td>
<td>(0.53)</td>
<td>(0.81)</td>
<td>(-0.26)</td>
</tr>
<tr>
<td>Adaptive expectations</td>
<td>(0.97)</td>
<td>(0.8)</td>
<td>(0.10)</td>
<td>(-0.01)</td>
</tr>
<tr>
<td>Rational expectations</td>
<td>(1.08)</td>
<td>(0.33)</td>
<td>(-0.06)</td>
<td>(-0.01)</td>
</tr>
<tr>
<td>Random walk without drift</td>
<td>(0.82)</td>
<td>(0.15)</td>
<td>(-0.07)</td>
<td>(-0.02)</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

The null hypothesis is rejected at 6 and 9-month maturities. This indicates that at longer horizons, the impact of speculative activity on futures price when expectations are formed heterogeneously is stronger than in the situation when the expectations are based on market fundamentals alone. At the same time, the null hypothesis of equality between the expected spot price coefficient obtained from the HAM and that obtained from the chartist expectation formation mechanism cannot be rejected. This shows that the HAM does not provide any incremental value over the chartist expectation mechanism in terms of explaining the impact of speculation on crude oil futures price.

To determine whether this result is motivated by the choice of the sample period the regression is re-estimated over the period January 2003–April 2013. The year 2003 marks the
start of a period of relatively high volatility in crude oil prices. The results of this regression are presented in Table 8.11.

Table 8.11: Cointegration regression for the Heterogeneous Agent Model (January 2003–April 2013)

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>$F_{t,t+n}$ $(n=1)$</th>
<th>$F_{t,t+n}$ $(n=3)$</th>
<th>$F_{t,t+n}$ $(n=6)$</th>
<th>$F_{t,t+n}$ $(n=9)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>−0.037</td>
<td>−0.0444</td>
<td>−0.0976</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>(−0.86)</td>
<td>(−0.55)</td>
<td>(−0.85)</td>
<td>0.01</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.9097</td>
<td>0.7397</td>
<td>0.6862</td>
<td>0.4175</td>
</tr>
<tr>
<td></td>
<td>(24.83)**</td>
<td>(10.47)**</td>
<td>(5.58)**</td>
<td>(3.31)**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0903</td>
<td>0.2326</td>
<td>0.2636</td>
<td>0.4543</td>
</tr>
<tr>
<td></td>
<td>(2.51)**</td>
<td>(3.49)**</td>
<td>(2.46)**</td>
<td>(4.22)**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.0001</td>
<td>0.0006</td>
<td>0.0012</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>−0.76</td>
<td>(1.82)*</td>
<td>(2.57)**</td>
<td>(3.96)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9922</td>
<td>0.9825</td>
<td>0.9726</td>
<td>0.966</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.992</td>
<td>0.982</td>
<td>0.9719</td>
<td>0.9651</td>
</tr>
<tr>
<td>$\hat{\epsilon}$</td>
<td>−7.85**</td>
<td>−5.15**</td>
<td>−5.22**</td>
<td>−5.41**</td>
</tr>
<tr>
<td>$\hat{\eta}$</td>
<td>−81.91**</td>
<td>−43.43**</td>
<td>−43.89**</td>
<td>−46.29**</td>
</tr>
<tr>
<td>$\hat{\zeta}_\alpha$</td>
<td>−7.87**</td>
<td>−5.01**</td>
<td>−5.09**</td>
<td>−5.35**</td>
</tr>
<tr>
<td>$\hat{\zeta}_\beta$</td>
<td>−81.24**</td>
<td>−39.82**</td>
<td>−40.45**</td>
<td>−44.45**</td>
</tr>
<tr>
<td>$\beta_0 = 0$</td>
<td>−0.86</td>
<td>−0.55</td>
<td>−0.85</td>
<td>0.01</td>
</tr>
<tr>
<td>$\beta_1 = 1$</td>
<td>−2.47**</td>
<td>−3.68**</td>
<td>−2.55**</td>
<td>−4.62**</td>
</tr>
<tr>
<td>$\beta_2 = 0$</td>
<td>2.51**</td>
<td>3.49**</td>
<td>2.46**</td>
<td>4.22**</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 = 1$</td>
<td>0</td>
<td>−0.93</td>
<td>−1.25</td>
<td>−2.95**</td>
</tr>
<tr>
<td>$N$</td>
<td>124</td>
<td>122</td>
<td>119</td>
<td>116</td>
</tr>
</tbody>
</table>

Note: * and ** indicate significance at the 10 and 5% levels.

Based on the EG and PO test statistics, the null hypothesis of non-stationarity is rejected, indicating that all the price series are co-integrated. The restrictions $\beta_1 = 1$ and $\beta_2 = 0$ are rejected for all horizons, whereas the restriction $\beta_1 + \beta_2 = 1$ is rejected only for the 9-month futures contract. Both arbitrage and speculation have significant impacts on the futures price but the impact of arbitrage is smaller over this period. This result lends support to our hypothesis that speculation played a role in the determination of crude oil futures price over
the alleged oil bubble period. Moreover, the results suggest that policymakers should also take into account the non-linear behaviour of crude oil price expectations when building and estimating models for predicting energy prices. Finally, for a firm attempting to hedge the risk by future contracts, appropriate consideration must be given to the length of the futures contract to limit prediction errors.

8.6 Summary and Conclusion

In this chapter, we extend the previous analysis by considering whether or not speculators take into account the predictive accuracy of the expectation formation mechanism in previous periods in forming expectations. For this purpose, we use the HAM proposed by Ellen and Zwinkels (2010). In this model, speculators switch between the chartist and fundamentalist expectation formation mechanism based on past forecasting errors.

The results presented in this chapter show that the past forecasting performance of the expectation formation mechanisms is important for the determination of oil price expectations. Although speculators do switch between chartist and fundamentalist expectation formation mechanisms, the role of market fundamentals in explaining speculative pressure on the futures price becomes less important at longer maturities. Most importantly, and in line with observations in the literature, the findings show that heterogeneous speculators have played an active role in driving the prices of crude oil futures.
9.1 Recapitulation

This study emphasises the role of financialisation in explaining the price of crude oil futures. It examines the proposition that the impact of financialisation on the crude oil price can be attributed to the profitability of trading strategies adopted by financial traders. Although a number of studies investigate the impact of financial trading on crude oil prices, they achieve this predominantly by examining the link between the ratio of commercial to non-commercial traders and the corresponding changes in the crude oil price. The findings obtained by these studies do not provide clear evidence for the proposition that financial trading has systematically driven the price of crude oil over and above the level warranted by the supply and demand fundamentals. In this study, we employ an alternative approach to identify financial trading activity, based on the study of Moosa and Al-Loughani (1995). A detailed analysis is conducted to extend our understanding of the role of financialisation in determining the crude oil price by examining futures demand by different types of financial traders.

To identify potential gaps, a review of the literature surrounding financialisation of commodity prices is presented in Chapter 2. The chapter commences by establishing a link between commodity market returns and the presence of financial investors. It then reviews the empirical approaches employed for investigating the role of financialisation. The review reveals a number of limitations inherent in past studies. For instance, studies examining the impact of speculation in terms of inventory response have been criticised due to their implicit
assumption that speculators can successfully distinguish price changes induced by speculative trading from those resulting from shifts in inventory demand. Similarly, concern has been raised regarding studies that use commercial and non-commercial positions to differentiate between conventional hedgers and financial traders, such as speculators and arbitragers.

What is absent (primarily by assumption) from this strand of literature is the notion that the trading patterns and beliefs of investors, which stem from public information and market prices, may also influence commodity prices. Heterogeneity in expectations, capital constraints and agency-related frictions can lead prices away from their fundamental values, thus producing boom and bust cycles (Singleton, 2011). However, empirical work examining the dynamics, measurement and implications of investor expectations in the crude oil futures market remains limited. For many years, the REH has been a dominant framework in understanding expectation formation across different markets. Though theoretically appealing, the REH has received tremendous criticism for the requirements of excessively sophisticated learning abilities of market participants (Branch, 2004) and for being excessively demanding (Nachbar, 1997, 2001, 2005; Foster and Young, 2001). Empirical tests also show that expectations of individuals across different markets are on average inconsistent with the REH (Lovell, 1986; Frankel and Froot, 1987; Ito, 1990; Macdonald and Taylor, 1993; Pilbeam, 1995; MacDonald and Marsh, 1993; Ellen and Zwinkels, 2010; Prat and Uctum, 2011). Instead, the findings indicate that market participants choose from a wide variety of expectations formation methods. In other words, expectations are formed heterogeneously.

The main conclusion emerging from our review of the literature is that the expectations of traders, which are essentially absent from the representative agent models, play an important role in the determination of commodity prices. Much of the literature on financialisation
abstracts from the dynamic interactions of the trading activities of investors who employ different expectation formation mechanisms and in turn generate price changes that are independent of the underlying supply and demand fundamentals. Complementing this view, Moosa and Al-Loughani (1995) explain the movement of crude oil futures price in terms of arbitrage and speculation. They show that both arbitrage and speculation have a significant influence on the price of crude oil futures. Their study adopts a unique approach to assess the role of the financialisation by combining two separate strands of literature: the theory of storage and behavioural finance. Based on our review of the extant literature, we establish that further insights into the role of financialisation can be gained by making two important extensions to the approach of Moosa and Al-Loughani. The first extension involves generalising their model by relaxing the assumption of the zero convenience yield. The second extension pertains to introducing heterogeneity and non-linear dynamics into the expectation formation mechanism of speculators.

Having established the link between financial trading and movements in commodity prices, in Chapter 3 we outline a theoretical framework for evaluating the effectiveness of arbitrage and speculation in explaining the price of crude oil futures. In that chapter the futures price is expressed in terms of excess demand for futures contracts by arbitragers and speculators. The demand for futures contracts by arbitragers is represented as a positive function of the difference between the actual futures price and the theoretical futures price. Speculative demand for futures, on the other hand, is a positive function of the difference between the futures price and expected spot price at the maturity of the futures contract. Using the respective demand functions we specify a testable model for analysing the impact of both arbitrage and speculation on crude oil futures price.
Chapter 4 discusses the literature surrounding the efficiency of the futures market within the context of the theory of storage. Discussion of the literature leads to the conclusion that net storage cost is important in evaluating the efficiency of the futures market. Given the limitations of the unbiasedness approach, market efficiency can be defined in terms of no-arbitrage profit. However some studies show that arbitrage alone is not sufficient for futures market equilibrium. According to some commentators, this result can be attributed to the missing role of speculation.

Chapter 5 examines the long-run relationship between the observed futures price, the theoretical futures price and the expected spot price using cointegration analysis. In doing so, the chapter demonstrates the effectiveness of arbitrage and speculation in explaining the price of WTI crude oil futures. The analysis extends the study of Moosa and Al-Loughani by relaxing the assumption of zero convenience in estimating the arbitrage price. The option pricing framework outlined by Heaney (2002a) is employed to approximate the convenience yield arising on the crude oil stocks. The empirical results indicate that in addition to arbitrage, speculation plays a significant role in the determination of futures price. Coefficient restriction tests show that both arbitrage and speculation are sufficient for achieving futures market equilibrium. In addition, we find that the inclusion of convenience yield does not lead to a significant change in the impact of arbitrage on crude oil futures price.

Chapter 6 examines the link between the behaviour of traders and movements in asset prices. It briefly reviews the theory of bounded rationality and heterogeneous expectations. The theory provides critical assumptions regarding the expectation formation process of market participants. The important conclusion drawn is that market participants are heterogeneous in that they can choose from a wide variety of prediction rules to form expectations. This conclusion is vital for understanding and identifying alternative expectation formation
mechanisms that do not rely on the assumptions of rationality. This helps create a scenario where the behaviour of market participants is closer to reality and allows a better reproduction of the main stylised facts.

Chapter 7 builds upon the analysis undertaken in Chapter 5 by introducing the role of heterogeneous expectations, where spot price expectations are proxied using adaptive, extrapolative and regressive expectations formation mechanisms. These mechanisms are selected because they are frequently used in the literature, rather than on the basis of an exhaustive review of all possible expectation formation techniques. To evaluate the predictive accuracy of the expectation formation mechanisms, the random walk model without drift is used as a benchmark.

Cointegration regressions involving the observed futures price, the theoretical futures price and the expected spot price are estimated to investigate the role of arbitrage and speculation. The main finding emerging from the analysis is that speculation plays a significant role in the determination of futures prices, even when speculators form their expectations heterogeneously. Also, the predictive performance of expectation formation mechanisms and the profitability of speculation vary considerably over the sample period, with most variation occurring during the oil price bubble of 2008. Finally, in line with the behavioural finance literature, support for technical trading behaviour among speculators is observed.

Chapter 8 extends the previous analysis by implementing a heterogeneous agent model (HAM) proposed by Ellen and Zwinkels (2010). In this model, the speculators are grouped as either chartists or fundamentalists and are assumed to incorporate a learning mechanism into their trading strategy whereby they switch between the chartist and fundamentalist expectation formation mechanisms, based on the past predictive performance. The results of
HAM estimation show that fundamentalist expectations exhibit mean reverting behaviour and do not exert a destabilising influence on the price. Chartist expectation formation mechanisms, on the other hand, show momentum behaviour, which implies a destabilising influence on the price.

Consistent with the findings obtained from the preceding analysis, we observe a significant influence of both arbitrage and speculation on crude oil price. The results also indicate that although speculators demonstrate heterogeneous behaviour, the HAM does not provide any additional value over chartist expectation formation mechanisms, in explaining the impact of speculative activity. However, we find that in the long run, the HAM does provide a better explanation of the extent of speculative activity in comparison to expectation formation mechanisms based on fundamentals. The conclusion emerging from these results is that speculators do incorporate past predictive performance when selecting their trading strategy. Moreover, the findings lend support to the assertion that speculative activity in addition to market fundamentals is responsible for driving the price of WTI crude oil futures.

9.2 Contribution

In undertaking this analysis, we make the several contributions to the existing literature. First, the study addresses an important gap in the literature by updating the estimate of the theoretical futures price to examine the effectiveness of arbitrage. The analysis also extends previous work by generalising the empirical model for different contract maturities. The addition of different maturity periods demonstrates the importance of oil price expectations when analysing the effectiveness of arbitrage and speculation.

Second, this study departs from the traditional assumption of a representative rational agent to highlight the explicit role of expectations of speculators in the determination of futures
prices. The study investigates price formation in the futures market by heterogeneous agents. This also provides a valuable contribution to our understanding of the individual impact of chartist and fundamentalist trading rules on the price of crude oil futures.

Third, this thesis contributes to the understanding of heterogeneous behaviour among futures market traders by employing a HAM in which agents update their beliefs based on past forecasting performance. The empirical results help establish evidence for behavioural heterogeneity among futures traders and provide significant evidence for switching between chartist and fundamentalist forecasting rules.

9.3 Limitations and Extensions

In this thesis we highlight the importance of arbitrage and speculation in the determination of price of WTI crude oil futures contracts. However, in undertaking this analysis, several limitations do arise, which must be summarised.

Although the analysis conducted in Chapter 4 suggests the presence of arbitrage profits, it does not incorporate the transaction costs incurred by arbitragers. To determine if arbitrage is profitable, an estimate of transaction costs is required. Absence of a reliable measure of transaction costs can lead to overstatement of the frequency of arbitrage opportunities and the size of arbitrage profit. This shortcoming can be addressed by evaluating the effectiveness of arbitrage once transaction costs are incorporated.

In the estimated model, speculators are constrained to spot speculation. However, as shown in Section 3.6, the empirical framework can be extended to incorporate futures speculation as well. This allows for a broader set of variables on which speculators can base speculative decisions. Another reason for incorporating futures speculation is that it allows speculators to
unwind their positions without the need to hold the spot commodity as is required under spot speculation.

Another caveat is that we examine the intensity of choice between chartist and fundamentalist expectation formation mechanisms purely on the basis of the magnitude of past forecasting errors. As a result, the model suffers from the important weakness that it does not account for the direction of change in the crude oil price. Hence, future work should incorporate the direction of change as well as the forecasting error magnitude in the selection mechanism.
REFERENCES


191.


Feedback Investment Strategies and Destabilizing Rational Speculation, *Journal of


Keynes, J.M. (1923), Some Aspects of Commodity Markets, Manchester Guardian Commercial, European Reconstruction Series, Section 13, 784-786.


228


